EDUCATIONAL TECHNIQUES AND METHODOLOGY

Dr. Anil Kumar

Educational Techniques and Methodology

Educational Techniques and Methodology

Dr. Anil Kumar (Associate Professor)
M.Ed., Ph.D. (Education), M.A. (Political Science),
NET (Education), Faculty of Education, Tantia University, Sri Ganganagar

ACADEMIC
UNIVERSITY PRESS

Educational Techniques and Methodology

Dr. Anil Kumar

ACADEMIC
UNIVERSITY PRESS

4378/4-B, Murarilal Street, Ansari Road, Daryaganj, New Delhi-110002.
Ph. No: +91-11-23281685, 41043100, Fax: +91-11-23270680
E-mail: academicuniversitypress@gmail.com

Year of Publication 2024-25

ISBN : 978-93-6284-300-5

Printed and bound by: Global Printing Services, Delhi
10 9 8 7 6 5 4 3 2 1

Contents

Preface

"Educational Techniques and Methodology" serves as a comprehensive guide for educators seeking to enhance their instructional practices and engage students effectively in the learning process. The book begins by exploring various active learning strategies aimed at promoting student engagement and participation. From collaborative group discussions to problem-solving activities and experiential learning opportunities, educators discover innovative ways to facilitate meaningful learning experiences that encourage students to take an active role in constructing their understanding of course material.

Furthermore, the text delves into the concept of differentiated instruction, which emphasizes the importance of tailoring teaching approaches to meet the diverse needs and abilities of students. Educators learn how to modify instructional methods, materials, and assessments to accommodate individual learning styles, preferences, and readiness levels. By embracing differentiated instruction, educators can create inclusive learning environments that empower all students to succeed.

In addition to active learning and differentiated instruction, "Educational Techniques and Methodology" addresses the vital role of assessment in the teaching and learning process. Educators explore a variety of assessment methods and techniques, including formative assessments, summative assessments, performance assessments, and authentic assessments. Through practical examples and case studies, educators gain insights into how to design assessment tasks that effectively measure student learning outcomes and provide meaningful feedback to guide instructional decision-making.

Moreover, the book examines the integration of technology in educational practices, highlighting how digital tools and resources can enhance teaching and

learning experiences. Educators discover how to leverage technology to create interactive lessons, facilitate collaborative learning activities, and provide access to educational resources beyond the confines of the traditional classroom. By incorporating technology into their instructional practices, educators can engage digital-native students and prepare them for success in an increasingly digital world.

"Educational Techniques and Methodology" explores strategies for promoting critical thinking, creativity, and problem-solving skills among students. Educators learn how to design learning activities that challenge students to think critically, analyze information, and apply their knowledge to real-world situations. By fostering these essential skills, educators prepare students to navigate the complexities of the modern world and become lifelong learners.

The book begins by examining active learning strategies that encourage students to actively participate in the learning process. From group discussions and problem-solving activities to hands-on experiments and simulations, educators discover innovative ways to promote student engagement and facilitate meaningful learning experiences. Additionally, the text explores the concept of differentiated instruction, emphasizing the importance of tailoring teaching approaches to meet the individual needs and abilities of diverse learners. By embracing differentiated instruction, educators can create personalized learning pathways that empower all students to succeed.

Delving into effective pedagogical practices, this book offers a comprehensive exploration of educational techniques and methodologies, providing educators with valuable insights and practical strategies for enhancing teaching and learning.

–Author

1

Introduction

MEANING OF RESEARCH

Research is a careful and detailed study into a specific problem, concern, or issue using the scientific method. It's the adult form of the science fair projects back in elementary school, where you try and learn something by performing an experiment. This is best accomplished by turning the issue into a question, with the intent of the research to answer the question.

Research can be about anything, and we hear about all different types of research in the news. Cancer research has 'Breakthrough Cancer-Killing Treatment Has No Side Effects in Mice,' and 'Baby Born with HIV Cured.' Each of these began with an issue or a problem (such as cancer or HIV), and they had a question, like, 'Does medication X reduce cancerous tissue or HIV infections?'

But all I've said so far is what research has done (sort of like saying baking leads to apple pie; it doesn't really tell you anything other than the two are connected). To begin researching something, you have to have a problem, concern, or issue that has turned into a question. These can come from observing the world, prior research, professional literature, or from peers. Research really begins with the right question, because your question must be answerable. Questions like, 'How can I cure cancer?' aren't really answerable with a study. It's too vague and not testable. Having a question creates an internal state of 'I need to know something.' To continue the baking example, this internal state of wanting something is like having a hankering for apple pie. Since you are reading this in a psychology section, we will put a psychological slant on this, and hopefully lose some of the baking metaphors.

MEANING AND INFORMATION OF RESEARCH

Research may be very broadly defined as systematic gathering of data and information and its analysis for advancement of knowledge in any subject. Research attempts to find answer intellectual and practical questions through application of systematic methods. Webster's Collegiate Dictionary defines research as "studious inquiry or examination; esp: investigation or experimentation aimed at the discovery and interpretation of facts, revision of accepted theories or laws in the light of new facts, or practical application of such new or revised theories or laws". Some people consider research as a movement, a movement from the known to the unknown.

It is actually a voyage of discovery. We all possess the vital instinct of inquisitiveness for, when the unknown confronts us, we wonder and our inquisitiveness makes us probe and attain full and fuller understanding of the unknown. This inquisitiveness is the mother of all knowledge and the method, which man employs for obtaining the knowledge of whatever the unknown, can be termed as research. Research is an academic activity and as such the term should be used in a technical sense. According to Clifford Woody research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organizing and evaluating data; making deductions and reaching conclusions; and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis. D. Steiner and M. Stephenson in the Encyclopedia of Social Sciences define research as "the manipulation of things, concepts or symbols for the purpose of generalizing to extend, correct or verify knowledge, whether that knowledge aids in construction of theory or in the practice of an art."

Research is, thus, an original contribution to the existing stock of knowledge making for its advancement. It is the pursuit of truth with the help of study, observation, comparison and experiment. In short, the search for knowledge through objective and systematic method of finding solution to a problem is research. The systematic approach concerning generalization and the formulation of a theory is also research. As such the term 'research' refers to the systematic method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analyzing the facts and reaching certain conclusions either in the form of solutions(s) towards the concerned problem or in certain generalizations for some theoretical formulation. Research in simple terms refers to search for knowledge. It is a scientific and systematic search for information on a particular topic or issue. It is also known as the art of scientific investigation. Several social scientists have defined research in different ways.

In the *Encyclopedia of Social Sciences*, D. Slesinger and M. Stephension (1930) defined research as "the manipulation of things, concepts or symbols for the purpose of generalizing to extend, correct or verify knowledge, whether that knowledge aids in the construction of theory or in the practice of an art".

According to Redman and Mory (1923), research is a "systematized effort to gain new knowledge". It is an academic activity and therefore the term should be used in a technical sense. According to Clifford Woody, research comprises

"defining and redefining problems, formulating hypotheses or suggested solutions; collecting, organizing 3 and evaluating data; making deductions and reaching conclusions; and finally, carefully testing the conclusions to determine whether they fit the formulated hypotheses".

Thus, research is an original addition to the available knowledge, which contributes to its further advancement. It is an attempt to pursue truth through the methods of study, observation, comparison and experiment. In sum, research is the search for knowledge, using objective and systematic methods to find solution to a problem.

MEANING AND TERM OF RESEARCH

The term 'research' has received a number of varied meanings and explanations. In its ordinary sense, the term refers to a search for knowledge. The *Advanced Learner's Dictionary of Current English* spells out the meaning of 'research' as 'a careful investigation or enquiry specifically through search for new facts in any branch of knowledge'.7 Redman and Mory, in a similar tone, define research as a 'systematized effort to gain new knowledge'.8 According to the *Webster's International Dictionary*, 'research' is 'a careful, critical enquiry or explanation in seeking facts or principles; diligent investigation in order to ascertain something'.

While *Webster Dictionary* explains the term 'research' to mean 'a systematic investigation towards increasing the sum of knowledge'. D Slesinger and M Stephenson perceived the term 'research' as 'the manipulation of things, concepts or symbols for the purpose of generalizing to extend, correct or verify knowledge, whether that knowledge aids in construction of theory or in the practice of an art'.9 The 1911 Cambridge edition of the *Encyclopaedia Britannica* defines research as:

The act of searching into a matter closely and carefully, enquiry directed to the *discovery of truth* and in particular, *the trained scientific investigation* of the principles and facts of any subject, based on original *and first hand study of authorities* or experiment.

Investigations of every kind which has been based on original sources of knowledge may be styled research and it may be said that without 'research' no authoritative works have been written, no scientific discoveries or inventions made, no theories of any value propounded – A combined reading of all the above-mentioned 'explanations' of the term 'research' reveals that 'research' is the 'careful, diligent and exhaustive investigation of a specific subject matter' with a view to knowing the truth and making original contribution in the existing stock of knowledge.

It is, in short, 'systematic search' in 'pursuit of knowledge' of the researcher. Mere aimless, unrecorded, unchecked search is not research which can never lead to valid conclusions. But diligent, intelligent, continued search for something is research. It refers to the process and means to acquire knowledge about any natural or human phenomenon. It involves a systematic enquiry into a phenomenon of interest. It is the process of discovering or uncovering new facts. It aims to contribute to the thitherto known information of the phenomenon.

Therefore, only systematic intensive investigation into, or enquiry of, fact qualifies to get the label of 'research'. And a 'search' becomes 'systematic' when a researcher, in his quest for knowledge and pursuit of truth, attempts to collect the required information from various sources and in a variety of ways systematically and exposes data to a severe and intensive scrutiny.

Research, thus, involves systematic scientific investigation of facts (or their hidden or unknown facets) with a view to determining or ascertaining something, which may satisfy the curiosity of the investigator and carry forward (his) knowledge. Such research involves identification of a research problem, the ascertainment of facts, their logical ordering and classification, the use of (inductive and deductive) logic to interpret the collected and classified facts and the assertion of conclusions premised on, and supported by, the collected information. 'Research', therefore, means a scientific collection and inspection of facts with a view to determining (or searching) something, which may satisfy the curiosity of the investigator and carry forward his knowledge. It requires a sound design for investigation, the appropriate methods of data collection and a mode of analysis. The prefix 're' in the word 'research', according to the *Concise Oxford Dictionary*, means 'repeated, frequent or intensive'. 'Research', therefore, implies a continued 'frequentative' 'intensive' 'search' for truth and/ or an enquiry for the verification of a fresh theory or for supplementing a prevailing theory. Research is, thus, a continuum.

Objectives of Research

The purpose of research, thus, is to acquire knowledge or to know about 'something' in a scientific and systematic way. Its purpose may, however, be to find solution to the identified problem. The former is referred to as 'basic' or 'pure' or 'fundamental' research while the latter takes the label of 'applied' or 'action' research. Fundamental research is mainly concerned with generalizations and with formulation of a theory (or re-confirmation of the existing theory). Its main aim is to acquire knowledge for the sake of acquiring it. Applied research, on the other hand, aims at finding or discovering solutions or answers to the identified 'problem(s)' or 'question(s)'. Obviously, every research study has its own goal(s) or objective(s). Nevertheless, 'research objective' of a given research study may fall under either of the following broad categories of 'research objectives':

1. To gain familiarity with a phenomenon or to achieve new insights into it.
2. To portray accurately the characteristics of a particular individual, situation or a group.
3. To determine the frequency with which something occurs or with which it is associated.
4. To test causal relationship between two or more than two facts or situations. 10
5. To 'know' and 'understand' a phenomenon with a view to formulating the problem precisely.
6. To 'describe' accurately a given phenomenon and to test hypotheses about relationships among its different dimensions.

CHARACTERISTICS OF EDUCATION RESEARCH

While educational research can take numerous forms and approaches, several characteristics define its process and approach. Some of them are listed below:

1. It sets out to solve a specific problem.
2. Educational research adopts primary and secondary research methods in its data collection process. This means that in educational research, the investigator relies on first-hand sources of information and secondary data to arrive at a suitable conclusion.
3. Educational research relies on empirical evidence. This results from its largely scientific approach.
4. Educational research is objective and accurate because it measures verifiable information.
5. In educational research, the researcher adopts specific methodologies, detailed procedures, and analysis to arrive at the most objective responses
6. Educational research findings are useful in the development of principles and theories that provide better insights into pressing issues.
7. This research approach combines structured, semi-structured, and unstructured questions to gather verifiable data from respondents.
8. Many educational research findings are documented for peer review before their presentation.
9. Educational research is interdisciplinary in nature because it draws from different fields and studies complex factual relations.

DEFINITION OF DESCRIPTIVE RESEARCH

Descriptive research does not fit neatly into the definition of either quantitative or qualitative research methodologies, but instead it can utilize elements of both, often within the same study. The term descriptive research refers to the type of research question, design, and data analysis that will be applied to a given topic. Descriptive statistics tell what is, while inferential statistics try to determine cause and effect.

The type of question asked by the researcher will ultimately determine the type of approach necessary to complete an accurate assessment of the topic at hand. Descriptive studies, primarily concerned with finding out "what is," might be applied to investigate the following questions: Do teachers hold favourable attitudes towards using computers in schools? What kinds of activities that involve technology occur in sixth-grade classrooms and how frequently do they occur? What have been the reactions of school administrators to technological innovations in teaching the social sciences? How have high school computing courses changed over the last 10 years? How do the new multimediated textbooks compare to the print-based textbooks? How are decisions being made about using Channel One in schools, and for those schools that choose to use it, how is Channel One being implemented? What is the best way to provide access to computer equipment in schools? How should instructional designers improve software design to make the software more appealing to students? To what degree are special-education teachers well versed concerning assistive technology? Is there a relationship between experience with multimedia computers and problem-solving skills? How successful is a certain satellite-delivered Spanish course in terms of motivational value and academic achievement? Do teachers actually implement technology in the way they perceive? How many people use the AECT gopher server, and what do they use if for?

Descriptive research can be either quantitative or qualitative. It can involve collections of quantitative information that can be tabulated along a continuum in numerical form, such as scores on a test or the number of times a person chooses to use a-certain feature of a multimedia programme, or it can describe categories of information such as gender or patterns of interaction when using technology in a group situation. Descriptive research involves gathering data that describe events and then organizes, tabulates, depicts, and describes the data collection. It often uses visual aids such as graphs and charts to aid the reader in understanding the data distribution. Because the human mind cannot extract the full import of a large mass of raw data, descriptive statistics are very important in reducing the data to manageable form. When in-depth, narrative descriptions of small numbers of cases are involved, the research uses description as a tool to organize data into patterns that emerge during analysis. Those patterns aid the mind in comprehending a qualitative study and its implications.

Most quantitative research falls into two areas: studies that describe events and studies aimed at discovering inferences or causal relationships. Descriptive

studies are aimed at finding out "what is," so observational and survey methods are frequently used to collect descriptive data (Borg & Gall, 1989). Studies of this type might describe the current state of multimedia usage in schools or patterns of activity resulting from group work at the computer. An example of this is Cochenour, Hakes, and Neal's (1994) study of trends in compressed video applications with education and the private sector.

Descriptive studies report summary data such as measures of central tendency including the mean, median, mode, deviance from the mean, variation, percentage, and correlation between variables. Survey research commonly includes that type of measurement, but often goes beyond the descriptive statistics in order to draw inferences. Thick, rich descriptions of phenomena can also emerge from qualitative studies, case studies, observational studies, interviews, and portfolio assessments. Robinson's (1994) case study of a televised news programme in classrooms and Lee's (1994) case study about identifying values concerning school restructuring are excellent examples of case studies.

Descriptive research is unique in the number of variables employed. Like other types of research, descriptive research can include multiple variables for analysis, yet unlike other methods, it requires only one variable (Borg & Gall, 1989). For example, a descriptive study might employ methods of analyzing correlations between multiple variables by using tests such as Pearson's Product Moment correlation, regression, or multiple regression analysis. Good examples of this are the Knupfer and Hayes (1994) study about the effects of the Channel One broadcast on knowledge of current events, Manaev's (1991) study about mass media effectiveness, McKenna's (1993) study of the relationship between attributes of a radio programme and it's appeal to listeners, Orey and Nelson's (1994) examination of learner interactions with hypermedia environments, and Shapiro's (1991) study of memory and decision processes.

On the other hand, descriptive research might simply report the percentage summary on a single variable. Examples of this are the tally of reference citations in selected instructional design and technology journals by Anglin and Towers (1992); Barry's (1994) investigation of the controversy surrounding advertising and Channel One; Lu, Morlan, Lerchlorlarn, Lee, and Dike's (1993) investigation of the international utilization of media in education (1993); and Pettersson, Metallinos, Muffoletto, Shaw, and Takakuwa's (1993) analysis of the use of verbo-visual information in teaching geography in various countries.

Descriptive statistics utilize data collection and analysis techniques that yield reports concerning the measures of central tendency, variation, and correlation. The combination of its characteristic summary and correlational statistics, along with its focus on specific types of research questions, methods, and outcomes is what distinguishes descriptive research from other research types.

Three main purposes of research are to describe, explain, and validate findings. Description emerges following creative exploration, and serves to organize the findings in order to fit them with explanations, and then test or validate those explanations (Krathwohl, 1993). Many research studies call for the description of

natural or man-made phenomena such as their form, structure, activity, change over time, relation to other phenomena, and so on. The description often illuminates knowledge that we might not otherwise notice or even encounter. Several important scientific discoveries as well as anthropological information about events outside of our common experiences have resulted from making such descriptions. For example, astronomers use their telescopes to develop descriptions of different parts of the universe, anthropologists describe life events of socially atypical situations or cultures uniquely different from our own, and educational researchers describe activities within classrooms concerning the implementation of technology. This process sometimes results in the discovery of stars and stellar events, new knowledge about value systems or practices of other cultures, or even the reality of classroom life as new technologies are implemented within schools.

Educational researchers might use observational, survey, and interview techniques to collect data about group dynamics during computer-based activities. These data could then be used to recommend specific strategies for implementing computers or improving teaching strategies. Two excellent studies concerning the role of collaborative groups were conducted by Webb (1982), and Rysavy and Sales (1991). Noreen Webb's landmark study used descriptive research techniques to investigate collaborative groups as they worked within classrooms. Rysavy and Sales also apply a descriptive approach to study the role of group collaboration for working at computers. The Rysavy and Sales approach did not observe students in classrooms, but reported certain common findings that emerged through a literature search.

Descriptive studies have an important role in educational research. They have greatly increased our knowledge about what happens in schools. Some of the important books in education have reported studies of this type: Life in Classrooms, by Philip Jackson; The Good High School, by Sara Lawrence Lightfoot; Teachers and Machines: The Classroom Use of Technology Since 1920, by Larry Cuban; A Place Called School, by John Goodlad; Visual Literacy: A Spectrum of Learning, by D. M. Moore and Dwyer; Computers in Education: Social, Political, and Historical Perspectives, by Muffoletto and Knupfer; and Contemporary Issues in American Distance Education, by M. G. Moore. Henry J. Becker's (1986) series of survey reports concerning the implementation of computers into schools across the United States as well as Nancy Nelson Knupfer's (1988) reports about teacher's opinions and patterns of computer usage also fit partially within the realm of descriptive research. Both studies describe categories of data and use statistical analysis to examine correlations between specific variables. Both also go beyond the bounds of descriptive research and conduct further statistical procedures appropriate to their research questions, thus enabling them to make further recommendations about implementing computing technology in ways to support grassroots change and equitable practices within the schools. Finally, Knupfer's study extended the analysis and conclusions in order to yield suggestions for instructional designers involved with educational computing.

RESEARCH APPROACHES

The above description of the types of research brings to light the fact that there are two basic approaches to research, viz., quantitative approach and the qualitative approach. The former involves the generation of data in quantitative form which can be subjected to rigorous quantitative analysis in a formal and rigid fashion. This approach can be further sub-classified into inferential, experimental and simulation approaches to research. The purpose of inferential approach to research is to form a data base from which to infer characteristics or relationships of population.

This usually means survey research where a sample of population is studied (questioned or observed) to determine its characteristics, and it is then inferred that the population has the same characteristics. Experimental approach is characterized by much greater control over the research environment and in this case some variables are manipulated to observe their effect on other variables. Simulation approach involves the construction of an artificial environment within which relevant information and data can be generated. This permits an observation of the dynamic behaviour of a system (or its sub-system) under controlled conditions.

CONTRIBUTION TO HIGHER EDUCATION AND RESEARCH

In recent years, with the advent of liberalization, and the growing fascination with all things American, there has been a concerted attempt to discount the role of enlightened state intervention in the economy. Certain idealogues-particularly those affiliated with the BJP, and other neoliberal have been slamming Nehruvian "socialism" as the single main cause for India's poverty and underdevelopment.

They have been contending that had the Indian government left everything to the Private Sector after independence, India would have grown much faster, and would have been like a "developed" nation today.

Although such an argument is entirely untenable when one looks at the concrete experience of most other developed and developing nations-particularly the Asian nations that have developed the most (such as S. Korea and Japan), this argument has been repeated so often, that it has now been accepted as gospel truth by many younger Indians—especially young software engineers and technology managers.

Unfamiliar with how both the Japanese and S. Korean governments have invested enormous sums of money in both higher education and cutting-edge research, they cling to the naive and misinformed belief that privatization is the magic wand for all of India's ills. In fact, most advanced nations in the world have developed thanks to much higher levels of state intervention in higher education and scientific and technological research than India.

For instance, it might be worth mentioning, that to this day, Moscow-with its plethora of Soviet era educational institutes and scientific labs remains a

world leader in cutting edge research. In many areas of modern Physics and other physical sciences, the research output from Moscow exceeds that of long-established US centres of privately-funded higher education such as Stanford or Yale.

So prolific are Moscow's scientists that even GM has now decided that it stands to gain more from building its newest research centre in Moscow rather than in the US. In fact, a scan of the leading online scientific journals reveals that many former Eastern-Bloc capitals (with state-funded universities and Science Academies modelled along Soviet lines) such as Budapest, Warsaw or Prague—are all significant contributors when it comes to cutting-edge research. In both S. Korea and Japan (Asia's technologically most advanced nations), scientific and technological research at publicly funded universities generally outstrips research at privately funded institutions.

And it cannot be emphasized enough that in the past half-century, Korea has grown much faster than the US. In Korea, there has been greater state intervention in the economy-not less. This is not to discount the leading role played by US universities in the world of scientific and technological research.

But it is often forgotten that many of them are state-funded. In any case, Indians ought to know that US universities are at best mediocre when it comes to undergraduate education, and half of all Ph.Ds in the Natural sciences and Engineering are awarded to scholars who were not born and educated in the US.

Although in the 19th century, the US became a world leader in science and technology largely on its own merits, in the past half century, the reputation of its universities has been maintained as much or more by immigrants than US-born citizens. Soon after World War II, US universities became a magnet for Japanese, Korean and European scientists since the war had greatly diminished opportunities at home. Now that Western Europe, Japan and Korea have developed (or redeveloped) their economies, US universities are enjoying the benefit of the best minds from India, China, Eastern Europe, Iran, the Middle East, and many other nations.

But even in the case of the US, it should be pointed out that when it comes to fundamental research in the physical sciences, it is the government labs and state universities such as UC Berkeley, UCLA, UC Santa Barbara, Minnesota or Ohio State-along with a hundred other state universities, that make up the backbone of basic scientific research in the physical sciences.

Private universities, who are better known for their management, architecture and engineering programmes complement state investment in higher education and research-but even in the US, they do not play the leading role.

In any case, no sustained engineering innovation is possible without a strong foundation in Maths, Physics, Chemistry and Geology. If it weren't for all the basic research emanating from state-funded universities and government labs, even private universities such as Stanford would be hard pressed to achieve what they have so far.

However, regardless of the situation prevailing abroad, the greatest indictment of India's Private Sector comes from a perusal of data pertaining to scientific and technological research. Facts always speak louder than ideological speculations and opinions, and when it comes to research output-whether in refereed internationally indexed journals, or domestic journals, or presentations at international or national conferences, India's Private Sector institutions simply don't make it anywhere near the top.

According to a search on Google Scholar (which appears to have access to about 70-80 per cent of Japanese and Western scientific journals, as well as all internationally-recognized Chinese, ASIAN and some Indian journals), the research output in the last 5 years in the Physical Sciences and Engineering from all the BITS institutions combined (in Pilani, Goa and Mesra) was about 510.

Compare that to IISc Bangalore's 4500, or IIT Delhi's 2700, or TIFR Bombay's 2670. In fact, BITS' combined research output not only trails the IITs in Bombay, Kanpur, Kharagpur, and Chennai (2700-2200)-but also lags Jadavpur University (1250), Delhi University (1100), ISI (1100), BHU Varanasi(900), Madras University (800), Calcutta University (740), Anna University (740), Pune University (740) and Hyderabad University (700.

The situation for other private institutions is even less impressive. Thapar's TIET Patiala logs in at 70, Manipal at 55, SASTRA (Thanjavur) and MEPCO Schlenck (Sivakasi) at 25 and Ahmedabad's NIRMA at 15.

Thapar's TIET equals little known state-funded SLIET (Longowal, Sangrur), but is behind Punjab's Amritsar University and Patiala University as well as NIT Kurukshetra or Kurukshetra University.

In fact, CUSAT Cochin, Allahabad University, Mysore University, Jaipur University-all do as well or better than BITS Pilani (or any other privately funded deemed university or autonomous engineering college). If BITS Pilani is excluded from the list, even state universities in smaller cities and towns such as in Tirupati, Kottayam, Burdwan, Jodhpur, Shillong, Jabalpur, Gorakhpur, Karaikudi, Sambalpur, Berhampur, Jalgaon, or Rajkot do better in terms of research activities.

In fairness to India's Private Engineering Colleges, it may be pointed out that Google Scholar generally fails to pick up on papers published in most Indian (or African) journals or those presented at Indian conferences. In addition, only a very small fraction of Indian Ph.D. theses get indexed in international scientific databases, so Google Scholar tends to significantly under-report India's research output. This is particularly the case with research relating to Chemistry, and all branches of Engineering.

At the private engineering colleges, papers presented at national or regional conferences can exceed journal papers by a factor of 4 or 5. (In the case of NIST, Berhampur, only 10 per cent of its research output is picked up by Google Scholar). This is not a bad thing, because national conferences have become an important venue for Indian scientists and engineers to get to know one another

and to become quickly familar with the latest research. It is especially useful for small colleges and universities to link up with larger and more established institutions.

But even after making appropriate adjustments to include papers published in Indian journals or presented at domestic conferences, the results will not be dramatically different. This is because Google's coverage of international engineering journals is also somewhat spotty. As a result, the data for the IITs and the NITs is also under-reported (by a factor of about 1.5 for the IITs, and by a factor of 2-2.5 for the NITs). For government-funded SLIET Sangrur, the Google results are just a third of the actual research output.

Even in Tamil Nadu-where the private colleges are most engaged with national engineering conferences, no unaided private engineering college would be able to match the overall research output of Anna University's Constituent Colleges or NIT Trichy (4-500), let alone IIT Madras.

This is not to entirely write off the Private Engineering colleges. The best private engineering colleges in Tamil Nadu, Karnataka, Maharashtra or Orissa-all do slightly better than the weak government engineering colleges who aren't especially well-staffed or well-funded. With the retreat of the state, private engineering colleges have begun to play a decisive role in undergraduate education.

But so far, fewer than 10 per cent of the private engineering colleges are equipped to offer Post-graduate programmes; less than half offer Post-graduate courses in more than one or two disciplines; and several are government-aided colleges. Moreover, almost all of these Post-graduate colleges are concentrated in just a few states: Tamil Nadu, Karnataka, Punjab and Haryana. Besides, at many of these Post-graduate colleges, only a few departments are well-developed; most departments have only a few active researchers.

Whereas the Google-indexed research output from all of Delhi's government-funded institutions (at about 8-9000) puts Delhi on par with Japan's number three city, Nagoya, (and several European capitals), BITS Pilani's output of about 450-500 puts it in the same league as Indonesia's Bandung, and somewhat behind Malaysia's Penang.

TIET logs in with Accra, Ghana, (or Peshawar, Pakistan), but behind Lagos, Nigeria.

Had Nehru left India's higher education entirely to the mercy of India's private trading companies and industrialists, it is much more likely, that today, India would look more like a struggling African country than a Singapore or South Korea as claimed by the detractors of Nehruvian socialism. It is thanks almost entirely to Nehruvian planning that India has a scientific and technological infrastructure in its metros that can match what is available in the European capitals.

In fact, contrary to the canard that CSIR (and all other government labs) are simply white elephants that ought to be disbanded, the data shows that government labs are playing an indispensable role in the country. While it may

well be true that structural rigidities, bureaucratic conservatism and inadequate interaction with Ph.D. students and Post-Docs may be hampering their productivity, the research output of the government labs continues to exceed that of the private sector.

For instance, IICT Hyderabad's log of 1000, BARC's 840, or PRL Ahmedabad's 625-all outpace BITS' combined output. In fact, if the research output from the government labs is included in the mix, Kolkata's overall research output is ten times that of Kuala Lumpur or Penang, and easily exceeds that of Bangkok.

With the inclusion of the research output from CSIR labs and all other government research institutions, several of India's tier-2 cities (such as Ahmedabad, Thiruvananthapuram, Lucknow, Bhubaneshwar, Chandigarh, Indore, Kochi and Allahabad) outrank the internationally reputed university towns in Malaysia.

If research presented at national conferences were also included, Tamil Nadu's Coimbatore, Madurai, Trichy and Karaikudi, Karnataka's Mysore, and other state capitals such as Jaipur, would also exceed the Malaysian university towns-not only in terms of Ph.D. guidance, but also in terms of papers published or presented. But because India's private media has largely ignored the research activities emanating from the government labs and small town universities, such work has received inadequate recognition and support.

Of course, what needs to happen is that a comparable structure needs to be in place in all the state capitals and other second-tier and even third-tier cities and towns. But since 1991, hardly anyone has paid any attention to augmenting the scientific and technological capabilities of the non-metros.

The prime entities for this neglect are the Indian media, CII and FICCI, and the BJP-who have done nothing but malign the nation's best state-funded institutions, yet done little to replace or even complement them.

Rather than berate the Nehru era for the problems of the present, they ought to be inquiring as to why more recent governments did so little to infuse fresh blood into India's key research institutions, and why no new research institutions were launched in the last two decades. If the older institutions have stagnated, why wasn't anything constructive done to re-energize them?

And above all, if India's private sector were the true alternative, why hasn't it been able to step in (in any significant way) in the past decade and a half?

The truth of the matter is that India's private sector has been so intoxicated with its rising profits, that it has largely failed to see the reality beyond its blinkered ideological compulsions.

Unlike industry associations in Korea and Japan, who have pushed their governments to increase state spending in higher education and research, and who themselves have encouraged the national press and electronic media to promote scientific and technological upgrades and innovations through dedicated columns on science and technology, and sought special media coverage of all major science and technology conferences, symposiums and workshops, India's

business leaders have often behaved as petty traders-preferring to market internationally-manufactured goods in India. Others have sought to import off-the-shelf technologies rather than spend any of their own money on R&D or encourage the government spend it for them. Hawkish neo-liberals like Chidambaram have only encouraged such behaviour with their reckless approach to slashing import duties and undercutting tax support for domestic research and capital improvements.

Unlike in other aspiring nations (such as Iran, Thailand or Malaysia), organizations like CII or FICCI have taken little interest in championing Indian science and technology. They have instituted few grant schemes for serious researchers, and their funding of academic research has been puny in relation to their rising profits.

Nor have they attempted to launch any journals where scientific and technological research performed at the country's government labs and universities could be easily disseminated to industry and concerned professionals. In most developed nations, there are not only industry journals dedicated to specific areas of science and technology but also to niche industries; and it is almost routine for major newspapers to have scientifically-trained professional journalists who regularly report science and technology news.

Scientists have a way of communicating with each other in academia. Their work needs to be interpreted for industrial use. Advanced industrial nations have put exactly such an infrastructure in place so that important scientific and technological breakthroughs can be communicated to industry in language that is intelligible to working engineers and technology managers. But the Indian Private Sector has done little to mimic such an apparatus in India.

Instead, India's leading English language dailies (such as Times of India, Hindustan Times or the Economic Times) are beginning to look more and more like tabloids than serious newspapers. Only the Hindu devotes some of its coverage to reporting on scientific and technological events, and even the Hindu's reporting tends to be limited to developments in Tamil Nadu or Southern India.

India's industrialists often complain that Indian scientists are not in touch with industry-but how often have they ever bothered to keep in touch with India's scientists?

The fact of the matter is that a nation's industrialists develop a keen interest in scientific and technological research when they begins to manufacture their own machine tools and capital goods. Some of India's small and medium industries are trying to do just that. But the captains of Indian industry aren't supportive of such moves. Instead, a dominant section of Indian industry merely wants to be an outsourcing agent for the transnationals, and prefers to shop for cheap second-hand assembly lines abroad. But that will never provide any real fillip for hard-core technological research in India.

For instance, Indian industry appears to be thrilled at the booming demand for airline tickets-but no one is talking about the escalating trade deficit that is a result of the 30 per cent growth in imported capital goods, transportation equipment, tools and machinery. No one is talking about manufacturing civilian

aircraft in India. This is ironical, because today, India is graduating 350,000 engineers. In five years, that number will grow to 450,000, and is projected to hit 600,000 in ten years. Within 5-10 years, India will have an experienced pool of engineers that is far greater than either the US or the EU.

Where India may lag (if corrective steps aren't taken soon) is in Post-graduates and Ph.D.s capable of engaging in advanced research and capable of designing virtually any assembly line, or any type of complex capital good.

With the right leadership, India has the intellectual potential to match any developed nation in producing virtually anything. What is missing is not the scientific or engineering willingness, it is the spirit of patriotism necessary to make it happen—the visionary entrepreneurial spirit that could marshall India's intellectual talents that are too quickly being sold to the highest international bidder.

The Infosys strategy of outsourcing Indian intellectual labour does not build anything solid and lasting for India. It brings in quick money with very limited long-term gains. It should be seen as a stop-gap measure, as something to utilize any engineering extras the nation may have. It should never be seen as the primary use of precious Indian talent.

India's best scientists and engineers should be designing and building things the country doesn't yet have: such as the latest deep-sea prospecting and drilling equipment. They should be designing and manufacturing a range of capital goods and machine tools, or new corrosion-resistant and energy-efficient materials; they should be experts at managing and developing new (and existing) energy resources; they should be experts at fuel-efficient mass transit; they should be as comfortable building civilian aircraft (or advanced submarines) as they are building the world's most fuel-efficient two-wheelers.

They should be capable of developing the latest environmental, sanitation or mass-transit solutions-instead of always seeking technological collaborations from others. There should simply be no holes in India's technological capabilities. A nation of a billion plus people should be able to take pride in its broad-based engineering design and manufacturing acumen, not just its ability to write smart banking software. Companies such as Infosys should be the last place where a good engineer should want to build his or her career-not the first choice. Outsourcing should be a back-up option for the Indian engineer, hardly the only game in town.

But for too many Indian business analysts and investors, outsourcing has become the sole or key driving option. Core engineering is treated as a mere adjunct—when, in fact, it should be quite the other way around. The truth is, India's private sector is unduly euphoric over the opportunities presented through outsourcing. A nation as large as India cannot become prosperous through outsourcing alone.

In the long run, India should aim to be a Germany, a Japan or a Korea, and let others fight over an exhausted outsourcing pie. Serious scientists and engineers should not be swayed by the hype, and seriously look beyond the Infosys-model of development.

Indian businessmen who constantly whine about the government not giving enough concessions for this or that should instead be held accountable by the Indian public and asked as to what they have done for the nation. What is their vision for India's advancement? What sacrifices have they made for India's progress?

The truth is, that notwithstanding the barrage of private sector oriented propaganda from the nation's press, the actual contribution of the India's private sector to higher education and scientific and technological research is eight parts hype and only one part substance.

Barring the more dedicated amongst India's private colleges who are indeed trying to do their best, the majority aren't even trying. But even the best of them are in no position to replace government efforts in the forseeable future.

Whereas India's scientists and engineers have not failed the Private Sector, India's Private Sector has indeed failed the nation-not only by failing to recognize the importance of the scientific and technological foundation that was laid in the very difficult years following freedom from colonial rule, but by doing all too little to augment what was done then.

In fact, India's Private Sector has yet to articulate a truly forward-looking vision for India's scientific and technological transformation. While it has profited greatly from India's state-funded institutions (such as the IISc, the IITs, and the NITs), it has yet to give anything significant back.

India's concerned citizens should have few illusions in this regard.

TYPES OF EDUCATIONAL RESEARCH

Educational research can be broadly categorized into 3 which are descriptive research, correlational research, and experimental research. Each of these has distinct and overlapping features.

DESCRIPTIVE EDUCATIONAL RESEARCH

In this type of educational research, the researcher merely seeks to collect data with regards to the status quo or present situation of things. The core of descriptive research lies in defining the state and characteristics of the research subject being understudied.

Because of its emphasis on the "what" of the situation, descriptive research can be termed an observational research method. In descriptive educational research, the researcher makes use of quantitative research methods including surveys and questionnaires to gather the required data.

Typically, descriptive educational research is the first step in solving a specific problem. Here are a few examples of descriptive research:

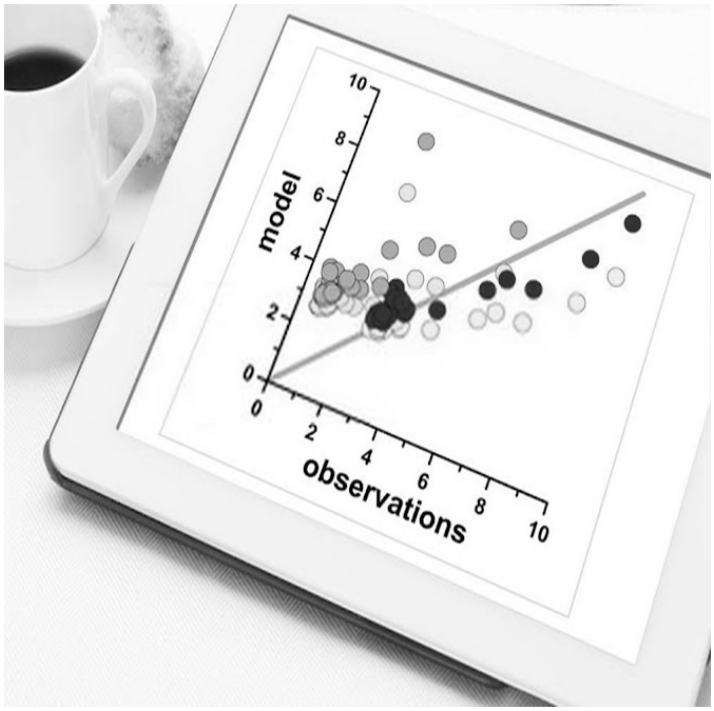
- A reading programme to help you understand student literacy levels.
- A study of students' classroom performance.
- Research to gather data on students' interests and preferences.

From these examples, you would notice that the researcher does not need to create a simulation of the natural environment of the research subjects; rather, he or she observes them as they engage in their routines. Also, the researcher is not concerned with creating a causal relationship between the research variables.

CORRELATIONAL EDUCATIONAL RESEARCH

This is a type of educational research that seeks insights into the statistical relationship between two research variables. In correlational research, the researcher studies two variables intending to establish a connection between them.

Correlational research can be positive, negative, or non-existent. Positive correlation occurs when an increase in variable A leads to an increase in variable B, while negative correlation occurs when an increase in variable A results in a decrease in variable B.



When a change in any of the variables does not trigger a succeeding change in the other, then the correlation is non-existent. Also, in correlational educational research, the research does not need to alter the natural environment of the variables; that is, there is no need for external conditioning.

Examples of educational correlational research include:

- Research to discover the relationship between students' behaviours and classroom performance.
- A study into the relationship between students' social skills and their learning behaviours.

EXPERIMENTAL EDUCATIONAL RESEARCH

Experimental educational research is a research approach that seeks to establish the causal relationship between two variables in the research environment. It adopts quantitative research methods in order to determine the cause and effect in terms of the research variables being studied.

Experimental educational research typically involves two groups – the control group and the experimental group. The researcher introduces some changes to the experimental group such as a change in environment or a catalyst, while the control group is left in its natural state.

The introduction of these catalysts allows the researcher to determine the causative factor(s) in the experiment. At the core of experimental educational research lies the formulation of a hypothesis and so, the overall research design relies on statistical analysis to approve or disprove this hypothesis.

Examples of Experimental Educational Research

- A study to determine the best teaching and learning methods in a school.
- A study to understand how extracurricular activities affect the learning process.

Based on functionality, educational research can be classified into fundamental research, applied research, and action research. The primary purpose of fundamental research is to provide insights into the research variables; that is, to gain more knowledge. Fundamental research does not solve any specific problems.

Just as the name suggests, applied research is a research approach that seeks to solve specific problems. Findings from applied research are useful in solving practical challenges in the educational sector such as improving teaching methods, modifying learning curricula, and simplifying pedagogy.

Action research is tailored to solve immediate problems that are specific to a context such as educational challenges in a local primary school. The goal of action research is to proffer solutions that work in this context and to solve general or universal challenges in the educational sector.

2

Technical Education

From the first Five Year Plan onwards India's emphasis was to develop a pool of scientifically inclined manpower. India's National Policy on Education (NPE) provisioned for an apex body for regulation and development of higher technical education, which came into being as the All India Council for Technical Education (AICTE) in 1987 through an act of the Indian parliament.

At the level of the centre the Indian Institutes of Technology and the Indian Institutes of Information Technology are deemed of national importance. The Indian Institutes of Management are also among the nation's premier education facilities.

Several Regional Engineering Colleges (REC) have been converted into National Institutes of Technology. The UGC has inter-university centres at a number of locations throughout India to promote common research, eg. the Nuclear Science Centre at the Jawaharlal Nehru University, New Delhi.

LITERACY

2001 government statistics hold the national literacy to be around 64.84 per cent. Government statistics of 2001 also hold that the rate of increase of literacy is more in rural areas than in urban areas. Female literacy was at a national average of 53.63 per cent whereas the male literacy was 75.26 per cent. Within the Indian states, Kerala has shown the highest literacy rates of 90.02 per cent whereas Bihar averaged lower than 50 per cent literacy, the lowest in India. The 2001 statistics also indicated that the total number of 'absolute non-literates' in the country was 304 million..

Attainment

World Bank statistics found that fewer than 40 per cent of adolescents in India attend secondary schools. *The Economist* reports that half of 10-year-old rural children could not read at a basic level, over 60 per cent were unable to do division, and half dropped out by the age 14.

Only one in ten young people have access to tertiary education. Out of those who receive higher education, *Mercer Consulting* estimates that only a quarter of graduates are "employable". An optimistic estimate is that only one in five job-seekers in India has ever had any sort of vocational training.

Private Education

According to current estimates, 80 per cent of all schools are government schools making the government the major provider of education. However, because of poor quality of public education, 27 per cent of Indian children are privately educated.

According to some research, private schools often provide superior results at a fraction of the unit cost of government schools. However, others have suggested that private schools fail to provide education to the poorest families, a selective being only a fifth of the schools and have in the past ignored Court orders for their regulation In their favour, it has been pointed out that private schools cover the entire curriculum and offer extra-curricular activities such as science fairs, general knowledge, sports, music and drama. The pupil teacher ratios are much better in private schools (1:31 to 1:37 for government schools and more teachers in private schools are female. There is some disgreement over which system has better educated teachers. According to the latest DISE survey, the percentage of untrained teachers (paratechers) is 54.91 per cent in private, compared to 44.88 per cent in government schools and only 2.32 per cent teachers in unaided schools receive inservice training compared to 43.44 per cent for government schools.

The competition in the school market is intense, yet most schools make profit. Even the poorest often go to private schools despite the fact that government schools are free. A study found that 65 per cent of schoolchildren in Hyderabad's slums attend private schools.

Private schools are often operating illegally. A 2001 study found that it takes 14 different licenses from four different authorities to open a private school in New Delhi and could take years if done legally. However, operation of unrecognized schools has been made illegal under the Children's Right to Free and Compulsory Education Act.

Rural Education

Following independence, India viewed education as an effective tool for bringing social change through community development. The administrative control was effectively initiated in the 1950s, when, in 1952, the government

grouped villages under a Community Development Block—an authority under national programme which could control education in up to 100 villages. A Block Development Officer oversaw a geographical area of 150 square miles which could contain a population of as many as 70000 people.

Setty and Ross elaborate on the role of such programmes, themselves divided further into *individual-based*, *community based*, or the *Individual-cum-community-based*, in which microscopic levels of development are overseen at village level by an appointed worker:

The community development programmes comprise agriculture, animal husbandry, cooperation, rural industries, rural engineering (consisting of minor irrigation, roads, buildings), health and sanitation including family welfare, family planning, women welfare, child care and nutrition, education including adult education, social education and literacy, youth welfare and community organisation. In each of these areas of development there are several programmes, schemes and activities which are additive, expanding and tapering off covering the total community, some segments, or specific target populations such as small and marginal farmers, artisans, women and in general people below the poverty line.

Despite some setbacks the rural education programmes continued throughout the 1950s, with support from private institutions. A sizable network of rural education had been established by the time the *Gandhigram Rural Institute* was established and 5, 200 Community Development Blocks were established in India. Nursery schools, elementary schools, secondary school, and schools for adult education for women were set up. The government continued to view rural education as an agenda that could be relatively free from bureaucratic backlog and general stagnation. However, in some cases lack of financing balanced the gains made by rural education institutes of India.

Some ideas failed to find acceptability among India's poor and investments made by the government sometimes yielded little results. Today, government rural schools remain poorly funded and understaffed. Several foundations, such as the Rural Development Foundation (Hyderabad), actively build high-quality rural schools, but the number of students served is small.

Issues

One study found out that 25 per cent of public sector teachers and 40 per cent of public sector medical workers were absent during the survey. Among teachers who were paid to teach, absence rates ranged from 15 per cent in Maharashtra to 71 per cent in Bihar. Only 1 in nearly 3000 public school head teachers had ever dismissed a teacher for repeated absence. A study on teachers by Kremer, *etc.*, found that 'only about half were teaching, during unannounced visits to a nationally representative sample of government primary schools in India.'.

Modern education in India is often criticized for being based on rote learning rather than problem solving. *BusinessWeek* denigrates the Indian curriculum

saying it revolves around rote learning. and ExpressIndia suggests that students are focused on cramming. A study of 188 government-run primary schools found that 59 per cent of the schools had no drinking water and 89 per cent had no toilets. 2003-04 data by National Institute of Educational Planning and Administration revealed that only 3.5 per cent of primary schools in Bihar and Chhattisgarh had toilets for girls.

In Madhya Pradesh, Maharashtra, Andhra Pradesh, Gujarat, Rajasthan and Himachal Pradesh, rates were 12-16 per cent.

Fake degrees are a problem. One raid in Bihar found 0.1 million fake certificates. In February 2009, the University Grant Commission found 19 fake institutions operating in India. Only 16 per cent of manufacturers in India offer in-service training to their employees, compared with over 90 per cent in China.

Initiatives

Following India's independence a number of rules were formulated for the backward Scheduled Castes and the Scheduled Tribes of India, and in 1960 a list identifying 405 Scheduled Castes and 225 Scheduled Tribes was published by the central government. An amendment was made to the list in 1975, which identified 841 Scheduled Castes and 510 Scheduled Tribes. The total percentage of Scheduled Castes and Scheduled Tribes combined was found to be 22.5 per cent with the Scheduled Castes accounting for 17 per cent and the Scheduled Tribes accounting for the remaining 7.5 per cent. Following the report many Scheduled Castes and Scheduled Tribes increasingly referred to themselves as *Dalit*, a Marathi language terminology used by B. R. Ambedkar which literally means "oppressed".

The Scheduled Castes and Scheduled Tribes are provided for in many of India's educational programmes. Special reservations are also provided for the Scheduled Castes and Scheduled Tribes in India, eg. a reservation of 15 per cent in *Kendriya Vidyalaya* for Scheduled Castes and another reservation of 7.5 per cent in *Kendriya Vidyalaya* for Scheduled Tribes. Similar reservations are held by the Scheduled Castes and Scheduled Tribes in many schemes and educational facilities in India. The remote and far-flung regions of North East India are provided for under the Non-Lapsable Central pool of Resources (NLCPR) since 1998-1999. The NLCPR aims to provide funds for infrastructure development in these remote areas.

The government objective for the *Sarva Shiksha Abhiyan* (SSA), started in 2001, is to provide education to children between 6–14 years by 2010. The programme focuses specially on girls and children with challenged social or financial backgrounds. The SSA also aims to provide practical infrastructure and relevant source material in form of free textbooks to children in remote areas. The SSA also aims at widening computer education in rural areas. SSA is currently working with Agastya International Foundation-an educational NGO-to augment its efforts in making science curriculum current and exciting. However, some objectives of the SSA, eg. enrollment of all children under the

scheme in schools by 2005 remain unfulfilled. Education Guarantee Scheme and Alternative and Innovative Education are components of the SSA.

Women from remote, underdeveloped areas or from weaker social groups in Andra Pradesh, Assam, Bihar, Jharkhand, Karnataka, Kerala, Gujarat, Uttar Pradesh, and Uttarakhand, fall under the *Mahila Samakhya Scheme*, initiated in 1989. Apart from provisions for education this programme also aims to raise awareness by holding meetings and seminars at rural levels. The government allowed 340 million rupees during 2007–08 to carry out this scheme over 83 districts including more than 21, 000 villages.

Currently there are 68 *Bal Bhavans* and 10 *Bal Kendra* affiliated to the *National Bal Bhavan*. The scheme involves educational and social activities and recognising children with a marked talent for a particular educational stream. A number of programmes and activities are held under this scheme, which also involves cultural exchanges and participation in several international forums.

India's minorities, especially the ones considered 'educationally backward' by the government, are provided for in the 1992 amendment of the Indian National Policy on Education (NPE). The government initiated the Scheme of Area Intensive Programme for Educationally Backward Minorities and Scheme of Financial Assistance or Modernisation of Madarsa Education as part of its revised Programme of Action (1992). Both these schemes were started nationwide by 1994. In 2004 the Indian parliament allowed an act which enabled minority education establishments to seek university affiliations if they passed the required norms.

Women's Education

Women have much lower literacy rate. Compared to boys, far fewer girls are enrolled in the schools, and many of them drop out. According to a 1998 report by U.S., Department of Commerce, the chief barrier to female education in India are inadequate school facilities (such as sanitary facilities), shortage of female teachers and gender bias in curriculum (majority of the female characters being depicted as weak and helpless)

The number of literate women among the female population of India was between 2-6 per cent from the British Raj onwards to the formation of the Republic of India in 1947. Concerted efforts led to improvement from 15.3 per cent in 1961 to 28.5 per cent in 1981. By 2001 literacy for women had exceeded 50 per cent of the overall female population, though these statistics were still very low compared to world standards and even male literacy within India. Recently the Indian government has launched Saakshar Bharat Mission for Female Literacy. This mission aims to bring down female illiteracy by half of its present level.

Sita Anantha Raman outlines the progress of women's education in India:

Sita Anantha Raman also maintains that while the educated Indian women workforce maintains professionalism, the men outnumber them in most fields and, in some cases, receive higher income for the same positions.

THE CHALLENGE OF EXPANDING TERTIARY EDUCATION

Despite the rapid growth of tertiary enrolments in most developing and transition countries over the past decades, the enrolment gap in relation to OECD economies has not decreased. In fact, quite the opposite has occurred. In 1980, the tertiary enrolment rate in the US was 55 percent compared to an average of 5 percent for developing countries as a whole. In 1995, the numbers were 81 and 9 percent respectively. Equally worrisome is the low level of development of postgraduate education in many parts of the world. In the Latin American and Caribbean region, for example, students enrolled in postgraduate programmes represent, on average, only 2.4 percent of overall tertiary enrolment, compared to 12.6 percent in the United States.

In many countries, the fiscal constraints experienced in recent years have undermined their financial capacity to undertake further expansion of the public tertiary education system while maintaining satisfactory levels of quality. Moreover, the problem of insufficient—sometimes even declining— funding is often compounded by the inefficient use of available resources. Examples of such inefficiencies include low student-staff ratios, underutilized facilities and duplicative programme offerings (even in situations of overcrowding), high dropout and repetition rates and a large share of the budget devoted to overheads and non-educational expenditures, such as subsidized student housing, food, transport and other services. These management inefficiencies drain scarce resources away from the fundamental objectives of increasing access, quality and relevance. Despite the rapid growth of tertiary enrolments in most developing and transition countries over the past decades, the enrolment gap in relation to OECD economies has not decreased.

In fact, quite the opposite has occurred. Despite more rapid demographic growth in the developing world than in OECD countries, transition rates from secondary to tertiary education have been higher in the latter countries as a result of several factors, including significant increases in secondary school completion rates, altered perceptions among students towards educational attainment as a means to achieve higher incomes; and the perceived need for highly skilled labour in a rapidly changing global economy. In the developing world, the most rapid growth of secondary enrolment rates has been observed in East Asia (from 47 to 66 per cent between 1990 and 1997), followed by Latin America (from 51 to 62 per cent) and the Middle East (from 52 to 57 per cent).

Another reason why tertiary enrolment levels are still relatively low in many parts of the developing world is the lack of institutional differentiation to accommodate the growing demand. In Latin America for instance, while countries like Cuba (79.1 per cent), Peru (42,9 per cent), Brazil (37.7 per cent) or Chile (35.3 per cent) have a significant non-university tertiary sector, others like Panama, Guatemala, Honduras, Nicaragua or El Salvador have a very small non-university sector (less than 5 per cent).

This is a major concern not only because non-university institutions can absorb a significant share of the demand for tertiary education but also because they are in general more capable to respond rapidly to changing labour market needs, as illustrated by the positive contribution of the two-year Technological Institutes in Mexico. Within the context of the growing enrolment gap, an equally worrisome issue is the slow rhythm of expansion of postgraduate education in many parts of the world. In the Latin American and Caribbean region, for example, students enrolled in postgraduate programmes represented on average only 2.4 percent of overall tertiary enrolment in 1997, compared to 12.6 percent in the United States. Given the fact that more than two-thirds of all postgraduate students are concentrated in only two countries, Brazil and Mexico, the proportion is significantly smaller in most countries of the region. This circumstance represents a serious constraint to building up the elements of national innovation systems so essential for increasing national productivity: research capacities, university-trained researchers and professionals, graduates with advanced technical and managerial skills and more dynamic university-industry linkages.

In many countries, the fiscal constraints experienced in recent years have undermined their financial capacity to undertake further expansion of the public tertiary education system while retaining satisfactory levels of quality. In the past 10 to 15 years, expenditure for tertiary education as a percentage of the total public education budget have experienced a significant reduction in several countries. In Ecuador, Mexico and Peru, per student expenditures fell by 30, 20 and 30 percent respectively between 1980 and 1990. In several countries, public sector enrolments have not grown or, where they have continued to grow, it has happened with reduced resources. The financial constraints have become even more acute in times of economic and financial crises, leading sometimes to significant decreases in tertiary enrolment.

In East Asia, for example, typical responses to the 1997-98 fiscal crisis in Korea, Thailand or Indonesia have been for low income students to drop out without completing their studies and for middle and high income students to shift from private to public tertiary institutions. Similar patterns have been observed in a few South American countries, notably Bolivia and Colombia. In the latter case, the National Association of Universities has calculated that the private universities of the country had lost close to 20 percent of their students since 1999. Moreover, the problem of insufficient — sometimes even declining— funding is often compounded by the inefficient use of available resources. These management inefficiencies drain scarce resources away from the fundamental objectives of increasing access, quality and relevance. Examples of such inefficiencies include underutilized facilities, duplicative programme offerings, low student-staff ratios, high dropout and repetition rates, uneconomical procurement procedures and a large share of the budget devoted to non-educational expenditures. Even though many public tertiary institutions are overburdened by students, their facilities are often-underused. In accordance

with civil service regulations or agreements with trade unions, many university facilities are closed evenings and weekends. Low student-staff ratios and high repetition and dropout rates also drive up the cost per graduate. In China and Brazil, for example, student-teacher ratios are very low in the public universities: between 5:1 and 9:1 in Chinese universities, 9:1 in the Brazilian federal universities, compared to a range of 15:1 to 20:1 in European universities. In many parts of the world, the high repetition and dropout rates are among the most important sources of low internal efficiency in public universities, especially in countries with open access. Low internal efficiency is especially prevalent in countries with open access, as is the case with most francophone African countries and a few Latin American countries (Argentina, Uruguay, Dominican Republic, Guatemala). This problem is sometimes compounded by the longer than usual duration of first degrees. In Bolivia, for instance, where the length of the first degree is five years, students actually spend 9 years to graduate on average. In Guatemala, the public universities spend 22 student-years to produce a graduate of a 6-year undergraduate programme.

STRATEGIC FRAMEWORK FOR FUTURE WORLD BANK SUPPORT

Investment in tertiary education is an important pillar of development strategies emphasizing the construction of democratic knowledge economies and societies. In this context, the World Bank can play a supportive role along the following three dimensions:

- Policy dialogue on tertiary education reforms.
- Knowledge sharing and lending to help implement reforms.
- Enabling framework for international public goods crucial for tertiary education development.

Supporting Reforms through Knowledge Sharing and Project Funding

The World Bank can contribute to the actual implementation of reforms through knowledge sharing and lending. It will give priority to supporting tertiary education reform programmes which can bring about positive developments and innovations in the following areas:

- Increase in institutional diversification to expand coverage and establish a life-long learning framework with multiple points of entry (including construction of accessible pathways from secondary to tertiary education, articulation mechanisms across tertiary education segments and capacity building for distance learning).
- Strengthening of science and technology research and development capacity (possibly in selected areas linked to a country's priorities for the development of comparative advantages).
- Improvement of the relevance and quality of tertiary education.
- Promotion of greater equity mechanisms intended to create and expand access and opportunities for disadvantaged students.

- Establishment of sustainable financing systems to encourage responsiveness and flexibility.
- Installation of management information systems for improved accountability, administration and governance.
- Enhancement and expansion of information technology and communications capacity to reduce the digital divide.

This menu of priority areas does not apply equally to all countries at all times. The relative emphasis and mix of interventions appropriate for any given country is very much linked to its specific circumstances at both the macroeconomic and the tertiary education levels. Income level, size of the country, existence of a post-conflict situation are all important factors to be taken into consideration. In setting priorities for the appropriate mix of lending and non-lending services in any given country, the Bank should be guided by the following criteria:

- Need to change, reflecting the gravity of issues faced by the country's tertiary education system and the degree to which change is urgently needed. For example, in countries with rapid growth of low quality private institutions, setting up an accreditation/evaluation system would be of high priority.
- Willingness to reform, based on the government's commitment to implement reforms and its ability to mobilize the major stakeholders in support of the reform agenda. Having already undergone a consensus-building exercise and formulated a national vision on the future of tertiary education would be a clear sign of reform readiness. In establishing and applying this evaluation matrix, it is important to distinguish between first and second generation reforms.

First generation reforms, which aim at addressing core problems of tertiary education systems (financing, efficiency, equity, quality assurance), are the first steps in moving from one way of doing things to a more appropriate approach. These include for instance changing from open-ended admission to selective access, introducing cost-sharing through fees and reduced subsidies in institutions which did not charge anything, establishing accreditation and evaluation in unregulated systems, transforming scholarships into student loans, creating non-university institutions alongside traditional universities, adopting the academic credit system, starting to rely on formula funding, *etc.*

Second generation reforms are undertaken by countries which have already dealt with their basic problems but need to do some fine-tuning to take first generation reforms one step further or to correct some unintended effects. Examples would be extending the eligibility of a student loan programme to all tertiary education institutions in a diversified system, introducing flexible mechanisms of articulation and credit transfer among institutions, or establishing competency-based evaluation mechanisms for on-line courses. In Chile, for instance, a framework for life-long learning is being established and financing mechanisms like student loans applicable to all types of tertiary education

institutions are under consideration. In Korea, the government recently launched a seven-year US$ 1.2 billion investment plan for tertiary education institutions.

The plan, known as Brain Korea 21, is essentially a new incentive financing scheme which offers institutions matching funds awarded on a competitive and selective basis to induce them to excel in cutting-edge research training areas such as biotechnology and information technology. Similarly Brazil, Mexico, Chile and Venezuela provide incentives through highly competitive mechanisms to world-class research training via the Millennium Science Initiative supported by the Bank.

Operational Implications: Based on the lessons of recent experience about the relative effectiveness of different forms and types of support, the following operating principles will be applied to guide Bank intervention in client countries.

Bank support should be:

- Appropriate to the specificity of a country's circumstances;
- Predicated on strategic planning at national and institutional levels;
- Focused on promoting autonomy and accountability;
- Sequenced in agreement with the time requirements of the capacity building objectives; and
- Sensitive to the political dimensions of tertiary education reform.

Encouraging and Facilitating Policy Dialogue

In many countries, the relationship between the government and the university sector and/or between public and private tertiary education institutions, is tense at best when not outright conflictive. Attempts at tertiary education reform are usually fraught with controversy. Proposals which are likely to affect established practices and vested interests are always met with fierce resistance and opposition by those groups most concerned by the intended redistribution of power and wealth. Under the right circumstances, the Bank may play a catalyst role in encouraging and facilitating the policy dialogue on tertiary education reforms. In the first place, the Bank can be a bridge builder by bringing to the same table various stakeholders who would not normally converse and work together. In the second place, the Bank can contribute information about a great variety of national and institutional experiences which can aliment the debate in a given country and offer objective elements on the range and content of policy options worth considering. This can lead to the formulation of a long term vision for the country's tertiary education system as a whole and of strategic plans at the level of individual institutions.

Social assessments can also be used for the purpose of identifying stakeholders concerns and addressing issues that are potentially controversial. The first social assessments for tertiary education projects were carried out in Indonesia, Jordan, Chile and China. In Indonesia, the analysis caused the universities in the outer islands to include young academics recently returned from graduate studies overseas in the self evaluation taskforces which had before been composed of established faculty members. All stakeholders were methodically consulted—

students, parents, faculty members, employers and community leaders and the assessment results proved invaluable in ascertaining their respective aspirations and suggesting acceptable reform instruments with respect to issues of access, gender, internal perceptions and employer expectations. In Chile, a thorough stakeholder analysis led to changes in the government's communication strategy during project preparation, including the inclusion of students in the government's project preparation team.

QUALITY ASSURANCE IN TERTIARY EDUCATION

The World Bank has supported the formation of national quality assurance systems in a number of countries. This is a necessary instrument in diversifying tertiary education systems. In some instances the Bank has helped establish specific accreditation mechanisms as required, but the general strategy has been to move towards comprehensive systems that cover the entire tertiary education landscape and that are consistent with international developments in standard setting, evaluation and accreditation. In Indonesia, the World Bank supported the introduction of accreditation mechanisms in a project to improve teacher training standards in public institutions.

It started with a pilot programme to define a set of standards by which all teacher training institutions would be evaluated and establish a baseline for institutional development. Small planning grants were made available to the five institutions which participated in the pilot to enable them to do a self study which was then externally evaluated and validated. The pilot study proved useful in generating acceptance for accreditation as a mechanism to improve the quality and relevance of teacher training. In the same country, the Bank has supported the establishment of the National Accreditation Board for Higher Education.

Sometimes, two sets of policy interventions can be mutually reinforcing in a synergetic way. In Indonesia, Romania, Argentina and Chile for instance, only programmes which are evaluated by the respective national quality assurance system are eligible to compete for innovation and quality enhancement grants. Brazil has a long standing tradition and positive experience from supporting its graduate programmes in this way (CAPES). Many tertiary education projects supported by the World Bank have also facilitated the introduction or development of management information systems (MIS) at the national and institutional levels, on the assumption that neither the state nor individual institutions can formulate and implement reforms without effective monitoring and management tools.

In Argentina, for instance, the Bank-financed project helped put in place a network of Intra- and Internet linking all the public universities among each other and with the outside world. Software for all dimensions of academic management were also developed and put together into an integrated MIS which provides information at the level of each individual institution as well as consolidates this information into a programme run by the national tertiary education authorities for monitoring and planning purposes. Even though many

universities were circumspect at first, they ended up embracing these innovations with enthusiasm because it helped them invest in modern information and communication technologies and provided them with useful management tools.

Some consequences of globalization and the growth of borderless education are turning into important issues which affect tertiary education in all countries but are often beyond the control of any one national government. Among those challenges of particular concern to countries seeking to build up their advanced human capital capacity are the new forms of brain drain, the absence of a proper international accreditation and qualifications framework, the lack of clear rules for the protection of intellectual property rights of distance education programmes and issues of access to information and communication technologies, including the Internet. There may be a need for global institutions such as the World Bank to get involved directly with its partners in the international community in efforts to create a discussion platform and promote an enabling framework for global public goods which are crucial for the future of tertiary education in the developing world.

With respect to the brain drain dimensions, at the very least, donor agencies should not contribute themselves to this trend through their cooperation programmes and projects, notably those supporting professor and student exchanges between tertiary education institutions in developing and transition nations and universities in industrialized countries. Clear rules of conduct should be defined and adhered to among donor agencies in order to facilitate the return of professionals trained overseas with external funding. These rules could even be extended to all public agencies in the donor country following the example of the British National Health Service that has adopted ethical guidelines excluding recruitment from any country likely to suffer a negative effect on its own health care services.

Another possible approach could be to promote increased reliance on "sandwich" or joint degrees, whereby scholars from developing countries who study in a foreign institution receive their degree from their home institution within the framework of established academic partnerships. A third dimension, suggested by the experience of the German Academic Exchange Programme (DAAD) which includes, as part of its scholarship funding, resources to purchase the minimum working equipment and material needed by returning scholars as well as travel funds to allow them regular visits to their professors to update their skills and knowledge. The World Bank could work with client countries to ensure that capacity building activities include measures to create a favourable work environment for national researchers and specialists. Finally, considering the high opportunity cost born by countries losing their advanced human capital to the brain drain and the corresponding net gain for the recipient economies, the possibility and feasibility of introducing a special tax on qualified immigrants from developing countries could be envisaged.

The rapid development of virtual providers of tertiary education programmes on a global scale, the increasing mobility of professionals across national borders

and the absence of quality assurance infrastructure and capacity in many developing countries make it important to establish an international quality assurance framework that can establish minimum common standards to guide countries and individuals. This is already happening in some regions of the world. For instance, the tertiary education policies of transition countries in Eastern Europe are very much influenced by the international coordination efforts to promote mobility, employability and competitiveness that are taking place in Europe as a result of the 1999 Bologna declaration, the 2001 Prague Declaration and the 2001 Salamanca convention.

In South America, the Ministers of Education of the Mercosur countries (Brazil, Argentina, Uruguay, Paraguay, Chile and Bolivia) have defined a minimum accreditation framework to facilitate the circulation of professionals from all member countries in their common labour market. In addition to the support provided through accreditation projects in individual countries, the World Bank could contribute towards the goal of establishing an international qualifications framework through consultations with partners in the donor community and specialized professional associations, as well as through grants from the Development Grant Facility. Two sets of complementary initiatives could be envisaged. First, the Bank could provide technical and financial assistance to groups of countries intent on setting up regional quality assurance systems.

The six Spanish-speaking countries of Central America, for example, are in the process of constructing a regional accreditation system instead of having each country creating its own quality assurance mechanism. Second, the Bank could support global quality assurance initiatives on a thematic basis, such as the current efforts of the World Federation for Medical Association towards the establishment of "International Standards in Medical Education." In the past few years, the World Trade Organization (WTO) has spearheaded international efforts to reduce national trade barriers. The inclusion in these negotiations of an increasing number of goods and services is now raising fears in the academic community, especially in developing countries, that WTO rules for tradable goods and services might extend progressively to tertiary education services. The specter of invasion by virtual and other non-traditional providers is leading some governments to take very conservative stands against foreign providers.

In this context, the World Bank could work at both international and national levels to help define rules of conduct and appropriate safeguards that would permit to protect students from low quality offerings and fraudulent providers without constituting rigid entry barriers. The following principles could serve to guide governments, licensing bodies and tertiary education institutions: (i) minimum infrastructure, facilities and staffing requirements, (ii) appropriate, transparent and accurate information on policy, mission statement, study programmes and feedback mechanisms of foreign providers, including channels for complaints and appeals, (iii) capacity building partnerships between foreign providers and local institutions and (iv) comparable academic quality and

standards including the full recognition, in home country, of degrees and qualifications delivered by foreign providers in a developing country.

A related issue faced by tertiary education institutions in developing countries is that of intellectual property rights for on-line programmes and courses and for access to digital libraries and digital information. The current debate involves two diametrically opposite views. At one end of the spectrum, many universities in industrialized countries favour enforcing strictly commercial rules of protection of the intellectual ownership of digital courses and materials, either on behalf of the university itself or of its professors as intellectual authors. At the other extreme are the partisans of a public good approach who, following MIT's recent initiative to offer all its course materials free of charge on its web site, advocate flow-cost access to digital courses, textbooks and journals for tertiary education institutions and scholars in the poorest countries.

The World Bank could play a brokering role to help create dissemination partnerships among publishing companies, universities in advanced nations and tertiary education institutions in developing countries along the model of the recently announced agreement among six leading publishers of medical journals to give free access to their scientific journals to more than 600 institutions in the poorest sixty countries of the world and low cost access to an additional 30 low income countries. Many developing countries, especially low income nations and small states, have limited resources to build up their information and communication technologies infrastructure.

They also lack the economic and political leverage to negotiate favourable access and price conditions. As part of its commitment to contribute to decreasing the digital divide between advanced and developing countries, the World Bank can work with specialized international agencies such as the International Telecommunications Union (ITU) and INTELSTAT and large communication technology companies to support the efforts of the poorest nations of the planet. In the same way as the Bank was instrumental in negotiating, on behalf of the small countries of the Caribbean, the introduction of a special tax paid by the big cruise companies to finance waste management programmes in that region, the Bank could intervene on behalf of low income and small states to help them get preferential treatment from telecommunications firms.

3

Technology in Education

Educational technology is the study and ethical practice of facilitating learning and improving performance by creating, using and managing appropriate technological processes and resources." The term educational technology is often associated with, and encompasses, instructional theory and learning theory. While instructional technology is "the theory and practice of design, development, utilization, management, and evaluation of processes and resources for learning," according to the Association for Educational Communications and Technology Definitions and Terminology Committee, educational technology includes other systems used in the process of developing human capability.

Educational technology includes, but is not limited to, software, hardware, as well as Internet applications, such as wiki's and blogs, and activities. But there is still debate on what these terms mean.

Technology of education is most simply and comfortably defined as an array of tools that might prove helpful in advancing student learning and may be measured in how and why individuals behave. Educational Technology relies on a broad definition of the word "technology." Technology can refer to material objects of use to humanity, such as machines or hardware, but it can also encompass broader themes, including systems, methods of organization, and techniques. Some modern tools include but are not limited to overhead projectors, laptop computers, and calculators. Newer tools such as "smartphones" and games are beginning to draw serious attention for their learning potential. Media psychology is the field of study that applies theories in human behaviour to educational technology. Consider the *Handbook of Human Performance*

Technology. The word technology for the sister fields of Educational and Human Performance Technology means "applied science." In other words, any valid and reliable process or procedure that is derived from basic research using the "scientific method" is considered a "technology." Educational or Human Performance Technology may be based purely on algorithmic or heuristic processes, but neither necessarily implies physical technology. The word technology comes from the Greek "techne" which means craft or art. Another word, "technique," with the same origin, also may be used when considering the field Educational Technology. So Educational Technology may be extended to include the techniques of the educator.

A classic example of an Educational Psychology text is Bloom's 1956 book, *Taxonomy of Educational Objectives*. Bloom's Taxonomy is helpful when designing learning activities to keep in mind what is expected of—and what are the learning goals for—learners. However, Bloom's work does not explicitly deal with educational technology *per se* and is more concerned with pedagogical strategies.

An Educational Technologist is someone who transforms basic educational and psychological research into an evidence-based applied science of learning or instruction. Educational Technologists typically have a graduate degree in a field related to educational psychology, educational media, experimental psychology, cognitive psychology or, more purely, in the fields of Educational, Instructional or Human Performance Technology or Instructional Systems Design. But few of those listed below as theorists would ever use the term "educational technologist" as a term to describe themselves, preferring terms such as "educator."

The transformation of educational technology from a cottage industry to a profession is discussed by Shurville, Browne, and Whitaker.

HISTORY

Educational technology in a way could be traced back to the emergence of very early tools, *e.g.*, paintings on cave walls. But usually its history starts with educational film or Sidney Pressey's mechanical teaching machines in the 1920s. The first large scale usage of new technologies can be traced to US WWII training of soldiers through training films and other mediated materials. Today, presentation-based technology, based on the idea that people can learn through aural and visual reception, exists in many forms, *e.g.*, streaming audio and video, or PowerPoint presentations with voice-over. Another interesting invention of the 1940s was hypertext, *i.e.*, V. Bush's memex.

The 1950s led to two major, still popular designs. Skinners work led to "programmed instruction" focusing on the formulation of behavioural objectives, breaking instructional content into small units and rewarding correct responses early and often. Advocating a mastery approach to learning based on his taxonomy of intellectual behaviours, Bloom endorsed instructional techniques that varied both instruction and time according to learner requirements.

Models based on these designs were usually referred to as computer-based training", Computer-aided instruction or computer-assisted instruction in the 1970s through the 1990s. In a more simplified form they correspond to today's "e-contents" that often form the core of "e-learning" set-ups, sometimes also referred to as web-based training or e-instruction. The course designer divides learning contents into smaller chunks of text augmented with graphics and multimedia presentation. Frequent Multiple Choice questions with immediate feedback are added for self-assessment and guidance. Such e-contents can rely on standards defined by IMS, ADL/SCORM and IEEE.

The 1980s and 1990s produced a variety of schools that can be put under the umbrella of the label Computer-based learning. Frequently based on constructivist and cognitivist learning theories, these environments focused on teaching both abstract and domain-specific problem solving. Preferred technologies were micro-worlds, simulations and hypertext.

Digitized communication and networking in education started in the mid 80s and became popular by the mid-90's, in particular through the World-Wide Web, e-Mail and Forums. There is a difference between two major forms of online learning. The earlier type, based on either Computer Based Training or Computer-based learning, focused on the interaction between the student and computer drills plus tutorials on one hand or micro-worlds and simulations on the other.

Both can be delivered today over the WWW. Today, the prevailing paradigm in the regular school system is Computer-mediated communication, where the primary form of interaction is between students and instructors, mediated by the computer. CBT/CBL usually means individualized learning, while CMC involves teacher/tutor facilitation and requires scenarization of flexible learning activities. In addition, modern ICT provides education with tools for sustaining learning communities and associated knowledge management tasks. It also provides tools for student and curriculum management.

In addition to classroom enhancement, learning technologies also play a major role in full-time distance teaching. While most quality offers still rely on paper, videos and occasional CBT/CBL materials, there is increased use of e-tutoring through forums, instant messaging, video-conferencing, *etc.* Courses addressed to smaller groups frequently use blended or hybrid designs that mix presence courses with distance activities and use various pedagogical styles.

The 2000s emergence of multiple mobile and ubiquitous technologies gave a new impulse to situated learning theories favouring learning-in-context scenarios. Some literature uses the concept of integrated learning to describe blended learning scenarios that integrate both school and authentic settings.

OBJECTIVES OF TECHNOLOGY EDUCATION

Technology education allows learners to explore a variety of activities related to many areas of human endeavor. Learners can develop problem solving strategies and work habits that will be useful in almost any career and or

occupation. Learners should develop a greater appreciation for the work of craft workers and the skill required of that work.

Within the scope of this educational programme is the effort to develop "soft skills" within the learner, as well as an opportunity for the learner to see how systems work together and the chance to put much of the academic class instruction to work in an applied way.

Many of the skills that are to be taught in technology education classes are identified in the 1991 SCANS Secretary's Commission on Achieving Necessary Skills Report. Many of these soft skills may be better developed in a lab, problem solving setting than in a typical academic classroom.

TECHNOLOGY EDUCATION VERSUS TECHNICAL EDUCATION

The goal of technology education is different from technical education. Technical education teaches specific skills directed towards a specific type of work, for example, carpentry, automotive maintenance, welding, or computer network administration, whereas technology education has a more generalized approach. Technology education is often confused with the term educational technology, which denotes technological devices and methods of delivering or developing educational information.

TECHNOLOGY IN EDUCATION

Many people warn of the possible harmful effects of using technology in the classroom. Will children lose their ability to relate to other human beings? Will they become dependent on technology to learn? Will they find inappropriate materials? The same was probably said with the invention of the printing press, radio, and television. All of these can be used inappropriately, but all of them have given humanity unbounded access to information which can be turned into knowledge.

Appropriately used— interactively and with guidance— they have become tools for the development of higher order thinking skills. Inappropriately used in the classroom, technology can be used to perpetuate old models of teaching and learning. Students can be "plugged into computers" to do drill and practice that is not so different from workbooks. Teachers can use multimedia technology to give more colourful, stimulating lectures. Both of these have their place, but such use does not begin to tap the power of these new tools. In this area, you will find descriptions of how computers can be used to stimulate and develop writing skills, collaborate with peers in foreign countries, do authentic kinds of research that is valuable to the adult world, and do complex kinds of problem solving that would otherwise be impossible.

NATURE OF INFORMATION TECHNOLOGY

It seems obvious that a world with information technology is somehow different from a world without information technology. But what is the difference? Is it a difference of order (faster, closer, clearer, etc) or is it a

difference of kind? How can we make sense of these questions? Does technology shape society or society shape technology, or both shape each other? What is the nature of this shaping? Is it in practices, in ways of thinking, or is it more fundamental? The answers to these questions will obviously influence the judgments we make about the social and ethical implications of information technology when we consider the policy and practical concerns of using information technology in a particular domain (such as commerce, education or government).

The answer to these questions are also grounded to a large extent in one's particular, implicit or explicit, ontology of information technology itself; what is the nature—way of being—of information technology as such? Obviously many different ontological positions are possible and have emerged. Nevertheless, it may be useful for the purposes of this entry to discern at least three contrasting and prevailing views.

INFORMATION TECHNOLOGY AS AN ARTIFACT OR TOOL

The most common view of information technology is that it is an artifact or tool simply available for humans to achieve their objective and outcomes. Some of these tools might be useful and others not. When users take up a tool or artifact (word processor, mobile phone, etc) it will tend to have an impact on the way they do things.

For example if I write with a word processor I would tend to have different writing practices than I would with pen and paper. According to this view, we need to understand the impact that information technology has on society as it is taken up and used in everyday practices. For example, how will communication with mobile phones change our social interaction and social relationships? In asking such a question this view does not primarily concern itself with the development of the technology—why and how did it come about in the first instance.

It mostly assumes that the particular technology—mobile phones in this case—operates in a more or less uniform manner in different social settings. In other words, it assumes that a particular technology has certain determinate effects on, or in, the context of its use.

This way of conceptualizing information technology leads to questions such as "what is the impact of the internet on education" or "what is the impact of CCTV on privacy". This view of technology is often criticized for a greater or lesser degree of technological determinism. Technological determinism is the view that technology more or less causes certain ways of doing or ways of organizing to come about. For example, a technological determinist may argue that the Internet's open and non-hierarchical architecture can more or less *cause* a society that uses it to become more open and less hierarchical. The work of Postman (1993) is an example of this type of critical evaluation of the impact of technology on society.

INFORMATION TECHNOLOGY AS SOCIALLY CONSTRUCTED ARTIFACTS AND ACTORS

Many scholars argue that the 'impact view' above of information technology does not give an adequate account of the relationship between information technology and society. Firstly, it does not take into account that the technology does not simply appear but is the outcome of a complex and socially situated development and design process. In this development and design process many alternative options become excluded in favour of the technology that is now available—obviously with important implications. In other words there are many cultural, political and economic forces that shape the particular options suggested as well as the way the selected options become designed and implemented. It is not only technology that 'impacts' on society; technology itself is already the outcome of complex and subtle social processes—in other words it is socially constructed. Moreover, they argue that when we look at the actual uses of particular technologies we discover that users use them in many diverse and often unexpected ways—leading to many and diverse unintended consequences. Both in its design and in its actual use there is an ongoing reciprocal relationship in which society and technology co-construct each other; they act through and upon each other. It is therefore very difficult to make general statements about the 'impact' of a technology. One can, at most, speak of some general trends for which many exceptions will invariably exist. For the proponents of the this constructivist view it is important to understand, through detailed descriptive accounts, the particular ways in which technologies emerge and become embedded in particular social practices. Examples of such studies can be found in the work of Bijker (1995), Law (1991) and Latour (1991).

INFORMATION TECHNOLOGY AS AN ONGOING HORIZON OF MEANING AND ACTION

For the phenomenologist the 'impact view' of technology as well as the constructivist view of the technology/society relationships is valid but not adequate (Heidegger 1977, Borgmann 1985, Winograd and Flores 1987, Ihde 1990, Dreyfus 1992, 2001). They argue that these accounts of technology, and the technology/society relationship, posit technology and society as if speaking about the one does not immediately and already draw upon the other for its ongoing sense or meaning. For the phenomenologist society and technology co-constitute each other; they are each other's ongoing condition or possibility for being what they are. For them technology is not just the artifact. Rather, the artifact already emerges from a prior 'technological' attitude towards the world (Heidegger 1977). For example, as the already technologically oriented human beings that we are, we will tend to conceive communication as a problem requiring a technological solution. Thus, technology is already the outcome of a technological way of looking and relating ourselves to the world. Once in place technology allows the world to 'show up' in particular ways (Introna and

Ilharco 2003). For example you are a different person to me with a mobile phone than without one. With a mobile phone you become disclosed, or show up, as 'contactable', 'within reach' as it were. It is this way of thinking about information technology, as a horizon of meaning and action that we want to elaborate further before considering how these various ways of conceptualizing technology shape our views on the social and ethical implications of information technology.

EDUCATORS DISINTERESTED OR PHOBIC ABOUT COMPUTERS AND COMMUNICATIONS

Thus far, most educators who use technology to implement the alternative types of pedagogy and curriculum are "pioneers": people who see continuous change and growth as an integral part of their profession and who are willing to swim against the tide of conventional operating procedures—often at considerable personal cost.

However, to achieve large-scale shifts in standard educational practices, many more teachers must alter their pedagogical approaches; and schools' management, institutional structure, and relationship to the community must change in fundamental ways.

This requires that "settlers" must be convinced to make the leap to a different mode of professional activity—with the understanding that, once they have mastered these new approaches, their daily work will be sustainable without extraordinary exertion. How can a critical mass of educators in a district be induced simultaneously to make such a shift?

Studies of innovation in other types of institutions indicate that successful change is always bottom-up, middle-out, and top-down. The driver for bottom-up innovation in a district is the children. Typically, students are joyful and committed when they are given the opportunity to learn by doing, to engage in collaborative construction of knowledge, and to experience mentoring relationships.

That these types of instruction are accomplished via educational technology will excite some kids, while others will be indifferent—but all will appreciate the opportunity to move beyond learning by listening. Educators can draw enormous strength and purpose from watching the eager response of their students to classroom situations that use alternative forms of pedagogy. Often, teachers have shifted from pioneers to settlers because they were worn down by the unceasing grind of motivating students to master uninteresting, fragmented topics; and administrators have undergone a similar loss of enthusiasm by being inundated with paperwork rather than serving as instructional coordinators. The professional commitment that kids' enthusiasm can re-inspire is a powerful driver of bottom-up change.

The source of middle-out change is a district's pioneers. Many teachers entered the profession because they love students of a certain age and want to help them grow—or love their subject matter and want to share its beauty and

richness. Often, these teachers feel alienated because the straightjacket of traditional instruction and school organization walls them away from meaningful relationships with their students and their subject. Similarly, many administrators want to serve as leaders and facilitators, but are forced by conventional managerial practices into being bureaucrats and bosses. Middle-out change is empowered when educators who have given up hope of achieving their professional dreams see pioneer colleagues using technology to succeed in those goals—and realise that, if everyone made a similar commitment, no one would have to make continuous personal sacrifices to achieve this vision.

The lever for top-down innovation is the community served by the district. Educators want respect—yet teaching has fallen from a revered professions to a much lower status.

The relationship between educators and their community is seldom seen as a partnership; instead, teachers and administrators often feel isolated, forced to perform a difficult task with inadequate resources. Parents, the business sector, and taxpayers bitterly debate the purpose of schools and sometimes attempt to micro-manage their operation.

In contrast, when homes, classrooms, workplaces and community settings are linked via new media to achieve distributed learning, much more positive interactions emerge between schools and society. Educators can move from isolation to collaboration with the community, from a position of low esteem to an respected role in orchestrating children's learning across a spectrum of settings. This shift in status is a powerful driver for innovation.

To activate these bottom-up, middle-out, and top-down forces for improvement, educators must take the lead in developing a shared vision for systemic reform, distributed learning, and sophisticated utilization of technology. Making such a commitment to large-scale educational innovation is not only the right thing to do, but is increasingly essential to educators' professional integrity. In many ways, physicians working in health maintenance organizations face challenges similar to teachers and administrators working in today's schools.

These doctors are responsible for the well-being of their patients, but work within administrative structures that restrict their decision making capabilities, that are focused on saving money at least as much as on combating illness, and that do not provide the latest technology or much time and resources for professional development. Yet we expect those physicians to do whatever it takes—fight the system for what the patient needs, spend personal time mastering the latest medical advances and technologies—to help those whom they serve. To do otherwise would be malpractice, a betrayal of trust, a breach of ethics as a professional.

Given advances in information technology that are reshaping the knowledge students need and the ways educators can help them learn, we need to accept a professional obligation—despite current institutional constraints—to do whatever it takes in changing traditional instructional practices so that a generation of children is truly prepared for the 21st century.

PERSPECTIVES AND MEANING OF EDUCATION TECHNOLOGY

HISTORY

One comprehensive history of the field is Saettler's *The evolution of American educational technology*. Another worthy title is Larry Cuban's *Oversold and Underused-Computers in the Classroom.*

For several decades, vendors of equipment such as laptop computers and interactive white boards have been claiming that their technologies would transform classrooms and learning in many positive ways, but there has been little evidence provided to substantiate these claims. To some extent, the history of educational technology has been marked by a succession of innovations that arrive with much fanfare but often fade into the background once fully tested, as Cuban argues in the above title.

THEORIES AND PRACTICES

Three main theoretical schools or philosophical frameworks have been present in the educational technology literature. These are Behaviourism, Cognitivism and Constructivism. Each of these schools of thought are still present in today's literature but have evolved as the Psychology literature has evolved.

BEHAVIOURISM

This theoretical framework was developed in the early 20th century with the animal learning experiments of Ivan Pavlov, Edward Thorndike, Edward C. Tolman, Clark L. Hull, B.F. Skinner and many others. Many psychologists used these theories to describe and experiment with human learning. While still very useful this philosophy of learning has lost favour with many educators.

SKINNER'S CONTRIBUTIONS

B.F. Skinner wrote extensively on improvements of teaching based on his functional analysis of Verbal Behaviour, and wrote "The Technology of Teaching", an attempt to dispel the myths underlying contemporary education, as well as promote his system he called programmed instruction. Ogden Lindsley also developed the Celeration learning system similarly based on behaviour analysis but quite different from Keller's and Skinner's models.

COGNITIVISM

Cognitive science has changed how educators view learning. Since the very early beginning of the Cognitive Revolution of the 1960s and 1970s, learning theory has undergone a great deal of change. Much of the empirical framework of Behaviourism was retained even though a new paradigm had begun. Cognitive theories look beyond behaviour to explain brain-based learning. Cognitivists consider how human memory works to promote learning.

After memory theories like the Atkinson-Shiffrin memory model and Baddeley's Working memory model were established as a theoretical framework in Cognitive Psychology, new cognitive frameworks of learning began to emerge during the 1970s, 80s, and 90s. It is important to note that Computer Science and Information Technology have had a major influence on Cognitive Science theory. The Cognitive concepts of working memory (formerly known as short term memory) and long term memory have been facilitated by research and technology from the field of Computer Science. Another major influence on the field of Cognitive Science is Noam Chomsky. Today researchers are concentrating on topics like Cognitive load and Information Processing Theory.

CONSTRUCTIVISM

Constructivism is a learning theory or educational philosophy that many educators began to consider in the 1990s. One of the primary tenets of this philosophy is that learners construct their own meaning from new information, as they interact with reality or others with different perspectives.

Constructivist learning environments require students to utilize their prior knowledge and experiences to formulate new, related, and/or adaptive concepts in learning.

Under this framework the role of the teacher becomes that of a facilitator, providing guidance so that learners can construct their own knowledge. Constructivist educators must make sure that the prior learning experiences are appropriate and related to the concepts being taught. Jonassen (1997) suggests "well-structured" learning environments are useful for novice learners and that "ill-structured" environments are only useful for more advanced learners. Educators utilizing technology when teaching with a constructivist perspective should choose technologies that reinforce prior learning perhaps in a problem-solving environment.

CONNECTIVISM

Connectivism is "a learning theory for the digital age," and has been developed by George Siemens and Stephen Downes based on their analysis of the limitations of behaviourism, cognitivism and constructivism to explain the effect technology has had on how we live, how we communicate, and how we learn. Donald G. Perrin, Executive Editor of the International Journal of Instructional Technology and Distance Learning says the theory "combines relevant elements of many learning theories, social structures, and technology to create a powerful theoretical construct for learning in the digital age."

IMPACT ON INSTRUCTIONAL TECHNOLOGISTS

As part of the federal Individuals with Disabilities Act IDEA amendments of 1997 and 1999, statements now require assistive technology devices and services

to be considered on an individualized basis and become a part of the individual education plan (IEP) if the child needs them to benefit from his educational programme. The individualized education programme (IEP) is a written statement for a child with a disability that is developed, reviewed, and revised at the child's school.

The IEPs occur each year for every child with a disability and they are developed by members of the IEP team including parents, teachers, special education teachers, administration and others. Section 508 of the Rehabilitation Act Amendments of 1998 is the most extensive new law with wide ranging effects.

This ruling requires that all US federal agencies make their information technology accessible to their employees and customers with disabilities. The law gives federal employees and members of the public the right to sue if the government agency does not provide comparable access to the information and data available to people without disabilities. Section 508 applies to Web sites that are produced for government agencies. All state agencies that receive federal funds under the Assistive Technology Act of 1998 are also required to comply with Section 508 requirements. Schools seeking to comply with legal requirements regarding students with disabilities need faculty with knowledge of assistive technology applications. Based on NCATE accreditation requirements, it would be reasonable for a school administrator or other official to expect that an educational or instructional technology graduate from an NCATE accredited programme would be able to effectively contribute to a student's IEP team. These expectations would include that such a graduate be able to make effective judgments and recommendations concerning assistive technology and universal access.

ASSISTIVE TECHNOLOGY COURSE DEVELOPMENT

With the rapidly aging population of the United States, there is also a growing need for assistive technology and universal design. To receive federal funding organizations must be IDEA and Section 508 compliant. There exists a need to provide instruction on assistive technologies and methodology to make technology products such as computer programmes and web pages handicapped accessible. Instructional and educational technology specialists require more extensive experience and education concerning assistive technology than they currently receive. Instructional/Educational Technology graduate programmes should devote a course to the presentation of the basic concepts and applications of assistive technology. This course could be offered as a requirement in the current university master's instructional technology programme and as an elective in its master's of education or exceptional education programmes. The NCATE and ISTE standards state that for initial certification, a teacher should "demonstrate awareness of resources for adaptive assistive devices for students with special needs." These standards would be well met by such a course. The technologies and strategies presented in a course concerning the application of

assistive technology would also address many of the other NCATE guidelines associated with speciality programmes such as educational computing and technology leadership.

An assistive technology course could be designed as an introductory or survey course in the application of technology as assistive and adaptive devices, software and strategies. This course could present strategies for students who are physically or mentally impaired, and may be in a mainstreamed situation. The purpose of the course material would be to teach about the use of technologies to overcome handicaps and improve functionality. Course topics could include: basics of assistive technology; legal/ethical issues associated with assistive technology; assistive technology and the individual education plan (IEP); levels of assistive technology; technology adaptations; Windows and Macintosh built-in accessibility tools; text-to-speech and speech-to-text; universal design and the internet; English as a second language, and physical and learning disabilities. An additional facet of such a course should also be designing web-based information to be universally accessible, covering such topics as making web pages more accessible and designing multimedia to overcome user handicaps. The assessments and activities of the course should include hands-on experiences with assistive technologies. Activities should be designed to include visitations to schools or labs to see assistive technology being used, the application and use of text-to-speech and speech-to-text programmes, experiences with adaptive switches and toys, and even experimentation with environmental control hardware and software.

During discussions and interviews with inservice teachers, counsellors, physical therapists, parents, and assistive technology organizations, a need for training and education in the area of assistive technologies was identified. Through continuing discussions, some basic areas of need in assistive technology education were identified. Visitations were conducted at the Assistive Technology Educational Network (ATEN), Florida Diagnostic Learning Resources (FDLRS) and Florida Instructional Materials (FIMSE) labs.

The goal of the visitations was to learn about the state of the art and the programmes being offered, and to understand the components of the AT community. Additional research continued through conducting a literature survey in the field, observing at schools and labs, and studying current Exceptional Student Education (ESE) and Instructional Technology (IT) programmes offered at universities.

In order to begin to fill the need that was perceived, a course outline was developed and components were taught at daylong hands-on workshops designed to introduce instructional technologists and teachers to assistive technology. From these preliminary discussions with professionals in the assistive technology community, it was found that an assistive technology course would be appreciated and that course delivery through distance learning would be preferred. Many of the potential students expressing interest in such a course were unable to travel to a university. As an educational technology programme course, it would have an added benefit as a recertification course for ESE professionals and general education teachers. After an initial course outline was developed,

members of parent support organizations such as the Statewide Advocacy Network on Disabilities (STAND), university professionals in special education, assistive technology state organizations such as Florida Diagnostic and Learning Resources System (FDLRS) and Assistive Technology Education Network (ATEN), future students in exceptional education, and other instructional technology professionals were asked to provide feedback on the course design, goals, topics and assessments. All were extremely pleased with the idea of the material becoming available for instructional technologists, exceptional student education (ESE) and general education educators. In its current form, the AT course "Technologies for Special Populations" is designed as an introductory course in the application of technology as assistive and adaptive devices in education. The course itself should model effective design practices. For example web pages will be designed for universal access and course materials and multimedia will be developed to be handicapped accessible. Because of its online delivery, the course serves as a model of information presented through an assistive medium.

COURSE LEARNING STRATEGIES

The Technologies for Special Populations course stresses hands-on experiences with various assistive technology approaches and devices. One of the main course goals is designing methods for a student to have actual experiences with the technology going beyond readings and looking at images about the technology. Students are expected to purchase, train, and use voice input systems, install and use an environmental control system, purchase and use a voice repeater, and use speaking software and hardware devices. Student interactions with assistive technologies fall into five areas. Students interact in an online forum, they have field experiences, and they complete technology projects, in addition to using standard materials such as tests and papers.

One of the strategies used in the Technologies for Special Populations course is the forum. Students participate for themselves and also analyze what other students have done and provide feedback to their classmates' thoughts. Forum topics include case studies that students use in experimenting with, suggesting and explaining assistive technologies. Further forum topics encourage students to discuss and evaluate the impact that the assistive technologies have on them while they use various devices and programmes such as environmental control, voice input, and text-to-speech.

Students will be required to observe the use of assistive technology as part of their field experiences. Students are asked to observe a student who uses assistive technology devices, or investigate and visit an assistive technology demonstration lab. Using an assistive technology device checklist and observation form, students would observe assistive technologies being used and then contribute in an online exchange concerning their observations. Additional experiences include assistive hearing, assistive audio, voice control, DVD applications, and environmental control.

4

Approaching the Learning of Statistics

In this aspect of Study Two I am concerned both with the constitution of the Approaches to Learning Statistics Questionnaire for gauging deep and surface approaches to learning Statistics and the characterisation of students' learning by means of scores on its Deep and Surface Scales.

THE APPROACHES TO LEARNING STATISTICS QUESTIONNAIRE

The ALSQ consists of 18 of the 28 items completed by students in the survey Initially the 28 items were based on the Study Process Questionnaire and the Approaches To Learning Mathematics Questionnaire. Consistent with these questionnaires, even numbered items are considered to indicate deep approaches to learning Statistics, while odd numbered items are considered to indicate surface approaches to learning Statistics. The following data analyses was done using the software package SPSS.

Scale Reliabilities

Initial analysis was carried out on all 28 items completed by the students. I used the Cronbach alpha coefficient, a common indicator of the internal consistency, on the two scales consisting of odd and even numbered items respectively. The alpha coefficient indicates the extent to which the items in the scale "agree" with each other. That is, it is a way of assessing the scale's reliability, indicating the extent to which the items are consistent. The inter-item correlation on the deep approach items indicated that internal consistency would be increased by omitting items

6 and 12. A number of students wrote on the survey that they did not understand item 12, so perhaps this is the reason for its poor correlation with other deep approach items. Item 6 consists of two ideas and I believe that some students agreed with one or other portion of the item, rather than with the item as a whole.

I defined the Deep Scale for the ALSQ as the scale consisting of the following 12 items: 2, 4, 8, 10, 14, 16, 18, 20, 22, 24, 26, 28. This scale yielded á=0.86, indicating a high level of internal consistency.

This coefficient compares favourably with internal consistency results obtained by Biggs with samples of university students on the Deep Approaches Scale of the Study Process Questionnaire, which ranged from 0.65 to 0.81. It is also similar to the two values of Cronbach's alpha obtained on two successive occasions, about six months apart, on the Deep Approaches Scale of the Approaches to Learning Mathematics Questionnaire. Scales indicating surface approaches to learning have consistently resulted in lower reliability than those denoting deep approaches in previous research.

My reliability analyses indicated that the maximum internal consistency on a scale indicating surface approaches to learning Statistics would be achieved by using a Surface Scale consisting of the following six items: 5, 7, 13, 21, 23, 27. For this scale, á=0.70. This value is in the middle of the range of Cronbach alpha coefficients cited by Biggs on the Study Process Questionnaire, but lower than the alpha coefficients (0.78, 0.77) obtained on our two trials of the Approaches to Learning Mathematics Questionnaire. By restricting the number of Surface Scale items to six, I narrowed the construct of a surface approach. Students who gained high scores on the six items were evidently concerned with uncritically memorising details, as presented by their teachers, in order to satisfy the demands of assessments.

From this point on, I will refer to this construct as a surface approach to learning Statistics as it is similar, though not identical, to the construct developed by Biggs and is consistent with the interpretation of a surface approach to learning given by Marton, Watkins and Tang. My Approaches to Learning Statistics Questionnaire is the questionnaire consisting of two scales - the Deep Scale, consisting of the twelve items and the Surface Scale consisting of six items, as described.

STATISTICS FOR DEEP AND SURFACE SCALES

Each item on the ALSQ has a possible score of between 1 and 5 with high scores indicating that the student usually adopted the approach to learning Statistics suggested by that item. On average, students had considerably lower scores on the Deep Scale than on the Surface Scale. Table shows that the average item score for the Deep Scale was a low 2.19, compared to an average Surface Scale item score of 3.43. This substantial difference between the item means for the Deep Scale and Surface Scale is statistically significant.

Table. Descriptive Statistics on Deep and Surface Scales

	Mean	Median	Standard Deviation	Cases
Deep Scale	2.19	2.08	0.68	267
Surface Scale	3.43	3.5	0.81	270

Moreover, for the vast majority of the surveyed students (214 students, 77 per cent) the average item score on the Surface Scale exceeded the average item score on the Deep Scale. This indicates that over three quarters of the students adopted surface approaches to learning Statistics more frequently than deep approaches.

Females had statistically significant, higher scores, on average, than males on the Surface Scale of the ALSQ ($t=2.98$, $p<0.01$). The item mean for females on this scale was 3.53 (standard deviation 0.76, N=196) while males attained an average of 3.18 (standard deviation 0.90, N=73).

On further analysis I found that while the mean item scores were higher for females than for males on all six items of the Surface Scale, these gender differences were greatest (at least 0.5) and statistically significant ($p<0.01$) on items 5, 13 and 27.

This suggests more concern about assessment and a greater reliance on the authoritative view of Statistics by females than males. Average scores on the Deep Scale do not differ significantly for males and females. However, on two items of the Deep Scale: item 20 and item 22, differences are statistically significant ($p<0.05$). On these two items males scored on average somewhat higher than females (0.3 or less) although the means for both sexes were low on these two items - at most 1.8. Both of these items are concerned with the level of interest in Statistics.

In general, the lowest average item scores were on the five items loading positively and substantially on Factor 1 of the Deep Scale - Find Interesting. The mean item score on these five items was only 1.67 (with standard deviation 0.69, N=269). A few students even indicated on their survey papers that had a rating lower than 1 ("only rarely") been available for these items they would have chosen it.

The most frequent reason cited by students, in response to the first open-ended survey question, was they found Statistics uninteresting. Mean item scores on the other two subscales of the Deep Scale were also low: 2.44 and 2.67 respectively.

On average, the students expressing reluctance to learn Statistics ("NO" students) scored substantially lower on the Deep Scale and substantially higher on the Surface scale of the ALSQ than those reporting that they would have chosen to study Statistics ("YES" students). Table shows the different means and other descriptive statistics for "YES" and "NO" students. The differences between the means on the Deep and Surface Scales are statistically significant.

Deep Scale

Table. Descriptive Statistics for"Yes" and"no" Students on and Surface Scales

	Mean	Standard Deviation	Cases
"Yes" Students	2.56	0.64	69
"No" Students	2.06	0.64	194

SURFACE SCALE

	Mean	Standard Deviation	Cases
"Yes" Students	2.89	0.68	71
"No" Students	3.64	0.77	195

Given that most of the participants of Study Two viewed the task of learning Statistics as imposed–a demand to be met which lacked personal relevance–it is to be expected that they would tend to adopt surface approaches to learning it. Most did not want to learn Statistics. It is therefore unlikely that they would strive to make the learning personally meaningful, interesting, or cohesive–the hallmarks of the deep approach. The statistics for the Deep and Surface Scales are in the predicted direction. This supports the validity of the Deep and Surface Scales in the sense of substantiating the sense and recognisability of these results. What each scale purports to denote is credible. The structure of the ALSQ fits with a qualitative analysis of students' approaches to learning Statistics.

DIFFERENCES IN STUDENT'S EXPERIENCES OF LEARNING STATISTICS

Cluster analysis is a way of identifying subgroups of students, rather than scores, on the basis of similarity of scores on given variables. Clusters are determined so that the members of each cluster are closer to each other than to members outside the cluster with respect to the chosen clustering variables. In factor analysis, variables that are negatively correlated may load on the same factor as those that are positively related. Cluster analysis, however, can be done so that variables that are negatively related appear in different groups. I have used these two research devices in a complementary fashion to illustrate different aspects of the inter-relationships. I will discuss cluster analysis, as a way of interpreting data.

There are many techniques for cluster analysis. Initially, I used the "average linkage" between-groups method to determine the number of clusters that would be appropriate. This method is a hierarchical clustering technique that successively groups individuals into clusters by minimising "distances" between clusters. The distance between clusters is defined to be the average of the squared Euclidean distances between all pairs of cases in which one member of the pair

is from each of the clusters. This process starts with as many clusters as there are cases and halts when all cases have been agglomerated into only one cluster. Hence the problem of the investigator is to decide at which stage the agglomeration process is to stop.

The primary consideration is to avoid grouping together cases which do not belong in the same cluster. It is generally suggested that large changes between fusions are useful indicators of the appropriate number of clusters as these combine groups which are distant or disparate. The steps in the hierarchical clustering solution produced by the SPSS Cluster procedure can be represented visually using "dendograms". The distances at which the clusters are combined. Inspection of these dendograms led me to decide that a two cluster solution would be appropriate to describe the data, as it was at the final stage, where two clusters were combined into one, that the distances between the clusters being combined appeared largest.

I then used the SPSS procedure Quick Cluster to produce the clusters. In this procedure the number of clusters is specified by the user. The cluster centres are estimated from the data. Each case is assigned to the cluster for which the distance between the case and the centre of the cluster (centroid) is smallest. The procedure is an agglomerative hierarchical method which is efficient in producing a solution once the number of clusters has been determined. The final cluster centroids describe the profile for each cluster. For an explanation of dendograms and examples.

My aim was to complement the factor analysis by investigating more closely the relationship between students' performances on assessments and their approaches to learning Statistics. Hence, clustering was done on the variables for approaches to learning and performance (Class1, Exam1, Class2, Exam2). These six variables were first standardised (to Z-scores, with overall mean=0, standard deviation=1) to ensure that all the data units being clustered were of similar magnitudes. This is necessary for the procedures used. Data on the clustering variables was available for 211 cases. The No Meaning students are included in the analysis as Concept was not one of the clustering variables.

Two Cluster Solution

The profiles for the two clusters are shown in Table. The majority of the 211 students fell into the first cluster. This cluster was characterised by relatively poor performances, on average, on all four assessment tasks, below average scores on the Deep Scale of the ALSQ and higher than average Surface scores. In contrast, the profile of the second group was of students who, on average, performed well on the assessments, particularly on the class tests. Although still low, the average Deep Scale score for this cluster was above the overall mean for the 211 students (Z=0.39). The average Surface Scale score for this group of students was well below the overall mean (Z=–0.49). Z-scores greater than 0 (that is, above the overall average) are shown in bold font in the table. Statistical tests of significance are not valid, as cluster analysis guarantees "significant" differences between the pairs of means on all the clustering variables.

In summary, the two clusters differed dramatically on performance in tests and examinations and on approaches to learning Statistics. The mean final Statistics mark for the Cluster 1 students was 44.5 per cent, while for those in Cluster 2 it was 75.1 per cent, a substantial difference in achievement.

Table. Standardised Means (Z-scores) for Cluster 1 and Cluster 2 on Assessments and Approaches to Learning Statistics

	Cluster 1 Low Achievers; Above Average On Surface (N=114)		**Cluster 2 High Achievers; Above Average On Deep**	
Class1	–0.65	(45%)	**0.87**	(81%)
Exam1	–0.62	(41%)	**0.81**	(69%)
Class2	–0.59	(51%)	**0.72**	(79%)
Exam2	–0.60	(41%)	**0.73**	(71%)
Deep Scale	–0.30		**0.39**	
Surface Scale	**0.37**		–0.49	

Raw means are shown in brackets. In the case of the Deep and Surface Scales, these are average item means for that scale. I then ascertained whether the two clusters differentiated between students with different conceptions of Statistics as well as other features of learning. This assisted me in understanding how performance and approach related to other aspects of students' learning, such as willingness to study Statistics.

It also enabled me to further check the construct validity of the variable Concept, that is, the conception categories for Statistics. Table below, shows the percentages I investigated in the two clusters for all the categorical variables. By reading across the rows of Table it can be seen what percentage of the students fell into each cluster for the various categories of description.

Table. Percentages of Category Groups in two Clusters

		Clusters (N=211)		
		1 (Low Achievers; Surface) (%)	2 (High Achievers; Deep) (%)	N
Choice	"Yes"	30.5	**69.5**	59
	"No"	**64**	36	149
Conception	No Meaning	**75**	25	8
	Process	**64**	36	53
	Mastery	**57.5**	42.5	73
	Tool	42	**58**	59
	Thinking	40	**60**	5
	Less Than Year 12	**65**	35	17

Prior Level	Mathematics In Society	**83**	17	12
of	2 Unit Mathematics	**72.5**	27.5	69
Mathematics	3 Unit Mathematics	45	**55**	47
	4 Unit Mathematics	14	**86**	7
	University	36	**64**	53
Prior	Mathematics			
	No	**59**	41	144
Statistics Age				
*Yes	43	**57**	61	
	19 Year Olds	**61**	39	105
	21 To 24	29	**71**	31
Gender	25 And Over	**51.5**	48.5	33
	Males	40	**60**	53
	Females	**59**	41	157

Percentages are out of N. The majority in each category is shown in bold. General Statistical Methods for Arts students or non-university statistics courses.

It is interesting to see which categories of students fell mainly into Cluster 2, the cluster characterised by a high average for achievement on assessments, a relatively high average on the Deep Scale and a relatively low average on the Surface Scale. As Table shows, nearly 70 per cent of those reporting that they were studying Statistics willingly were in cluster two. Most of the students who conceived of Statistics as applicable to the real world were in this cluster.

The majority of those who had studied mathematics at one of the two highest levels for their Higher School Certificate, or had taken university mathematics, were in the second cluster as were most of those who had taken a previous course in statistics. The only age group for which the majority was in Cluster 2 was the group aged 21 to 24 years old, suggesting that these students benefited from their maturity and life experiences but were not so far away from school mathematics that they were disadvantaged. In contrast to females, males fell primarily into cluster two.

In summary, the cluster analysis presents evidence of relationships among students' conceptions of Statistics, their approaches to learning it and the outcomes of their learning. The dichotomy supports the validity of the construct: conceptions of Statistics, as denoted by Concept. The shift in majority from cluster 1 to cluster 2 coincides with the shift from Mastery to Tool, that is, from categories associated with conceptions of Statistics as performance in assessments to those which present Statistics as applicable to the real world. The two cluster solution also highlights categories associated with successful learners of Statistics.

Four Cluster Solution

The two cluster solution presented has the advantages of simplicity and parsimony. However, the in-depth investigation on the learning of mature

students, carried out in Study One, shows that marks on assessment tasks and approaches to learning Statistics are not related in a simple fashion.

Further, a careful examination of the qualitative data for Study Two, namely interview data and students' responses to the open-ended questions in the survey, suggested to me that four clusters would present a more recognisable reality than two. That is, although the data suggested that there were high achievers who expressed educationally desirable approaches to learning Statistics, there also appeared to be those who succeeded in assessment tasks in spite of approaches characterised by a lack of interest or personal meaning and a fragmentation of the material. Also, not all of those who performed poorly in examinations and tests expressed predominantly surface approaches to learning Statistics.

I identified four different profiles of achievement and approaches by clustering on the same variables as for the two cluster solution using the SPSS procedure Quick Cluster. In this case I specified the number of clusters as four. As expected, differences between the means for the four groups were considerable on all the clustering variables. Table summarises this data for the four clusters. They show the standardised means (Z-scores) for final marks in Statistics and approaches to learning it for the four clusters.

Table Standardised and Raw Means of Final Statistics Marks and Scores on Deep and Surface Scales for Four Clusters.

Means	Cluster 1	Cluster 2	Cluster 3	Cluster 4
	High Achievers;	High Achievers;	Low Achievers;	Low Achievers;
	Above Average	Above Average	Above Average	Above Average
	On Deep Scale	On Surface Scale	On Deep Scale	On Surface Scale
Mark	**1.03** (77%)	**0.62** (70%)	–0.37	(52%) –1.1 (37%)
Deep	**0.98**	–0.41	**0.16**	–0.4
Surface	–1.06	**0.38**	–0.50	**0.7**

Cluster means which are above the overall average (Z>0) are indicated in bold font. Raw means are shown in brackets.

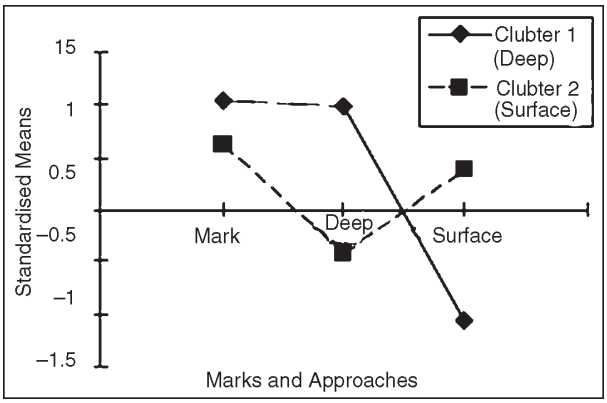


Fig. Marks and Approaches for The two Higher Achieving

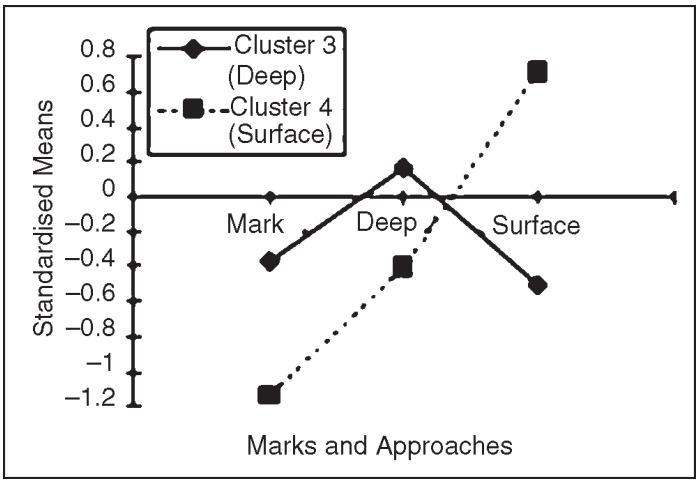


Fig. Marks and Approaches for The two Lower Achieving Cluster Groups

Most of the males were in the higher achieving clusters, Clusters 2 and 1 (38 per cent and 30 per cent, respectively) while females fell mainly into Clusters 4 and 2 (32.5 per cent, 26 per cent) which have a "high surface" profile. The four groups had, on average, different attainments in, and approaches to, learning Statistics. Like the two cluster solution presented earlier, the four clusters also differed in their profiles with respect to other variables, such as Concept. These profiles are discussed below and the key points summarised in Tables at the end of this discussion.

The profiles of the first two clusters are particularly interesting. While both these groups of students achieved high marks in Statistics, attaining averages of 77% and 70 per cent respectively, the orientations to learning Statistics characterised by these two clusters were very different.

Cluster 1 is characterised by students who, on average, achieved very high marks on assessments. This is the only cluster for which the average item score on the Deep Scale exceeded that on the Surface Scale. In addition, unlike the other three clusters, most of the students in this group (53 per cent, 25 students) expressed willingness to learn Statistics. The most common, or modal, category for their conceptions was TOOL (49 per cent of them, 23 students) with a further 6 per cent (3 students) in the Critical Thinking category.

The students classified in Cluster 2, as a group, evidently had a facility for learning Statistics, while lacking the motivation or awareness of the subject of their colleagues in Cluster 1. Most (75 per cent, 46) of them reported reluctance to study Statistics. Their modal category of conception was Processes (36 per cent, 22 students), followed by Mastery (30 per cent, 18 students). These distributions in Choice and Concept are consistent with the predominantly surface approach profile of this cluster. In all, they suggest a focus on reproducing aspects of the material for purposes of assessment, rather than a view of Statistics as personally meaningful, relevant and cohesive knowledge.

Clusters three and four performed, on average, at a much lower level than the first two groups, attaining averages of 52 per cent and 37 per cent respectively. Hence, the Cluster 3 students, on average, were less successful in the academic context than the Cluster 2 group, although their approaches to learning Statistics appear to be more educationally desirable than those in Cluster 2. The modal category of conception for this group was TOOL (40 per cent, 17). A further 5 per cent (2) of these students were classified in the Critical Thinking category. However, well under half (40 per cent, 17 students) reported that they would have studied Statistics given a choice.

The profile of the Cluster 4 students is particularly dismal with 93 per cent of the students in this cluster reporting that they would not have studied Statistics given a choice. On average, these students adopted predominantly surface approaches to their learning. Their average item score on the Surface Scale of the Approaches to Learning Statistics Questionnaire was very high: 4.01. The conceptions of 82 per cent of them were classified in the first three categories. Most of them (53 per cent, 32 students) perceived Statistics as being about Mastery of statistical ideas and skills in order to meet assessment demands. Although 13 per cent (8 students) expressed conceptions of Statistics that were classified as Tool, none of these students reported an intention to actually use this tool.

The following table summarises the percentages in each cluster for the variables Choice (YES or No) and Concept. The percentages shown are out of the number of cases (N) that is, column percentages and are rounded to whole numbers. Missing data accounts for differences from 100 per cent. I have indicated the modes in each cluster for Concept and Choice in bold font.

Table. Percentages of Choice and Concept Categories in Four Clusters

		Cluster 1 High Marks; Deep N=47	**Cluster 2 High Surface Marks; N=61**	**Cluster 3 Low Marks; Deep N=43**	**Cluster 4 Low Marks; Surface N=60**
Choice	"Yes"	**53%**	**23%**	40%	50%
	"No"	45%	75%	**61%**	**93%**
Concept	No Meaning	9% 28%	8%	5%	2%
	Processes	**49% (17%)**	**36%**	26% 23%	27%
	Mastery	6%	30% 18%	**40%** (7%)	**53%**
	*Tool	6%	(3%)	5%	13% (0)
	Critical Thinking				

The percentages shown in brackets are for a personally useful Tool.

The four clusters also distinguished between the prior levels of mathematics undertaken by the surveyed students. Table highlights the shifts by showing the

modal cluster for each level of the variable Prior Maths. As can be seen from this table, the modal cluster for the groups of students with lower backgrounds in mathematics was Cluster 4.

The modal cluster shifted to Cluster 2 for the students who had studied 3 Unit Mathematics for their Higher School Certificate. Encouragingly, for those students who had taken the highest level of mathematics at school or had studied mathematics at university, the most common cluster was Cluster 1.

Table. Modal Cluster for Prior Levels of Mathematics

Prior Maths	Cluster 1 High Marks; Deep	Cluster 2 High Marks; Surface	Cluster 3 Low Marks; Deep	Cluster 4 Low Marks; Surface	N
Less Than Year 12				**47%**	
Maths In Society				**58%**	12
2 Unit Maths				**44%**	69
3 Unit Maths				**38%**	47
4 Unit Maths				**57%**	7
University Maths				**40%**	53

The percentages shown are out of the number of cases that is, row percentages. Percentages are rounded to whole numbers. Finally, there was a relation between the two cluster solution and the four cluster solution. The 97 students who were in the Cluster 2 for the two cluster solution were almost equally divided into Clusters 1 and 2 in the four cluster solution. Six per cent were in Cluster 3 and none in Cluster 4.

In summary, the four cluster solution highlights differences in orientations to learning Statistics and performance on assessments with greater depth and complexity than the two cluster solution. The directions and shifts of the categorical variables such as Choice and Concept are in the predicted directions, thus enhancing the validity of the analysis.

Vignettes From Student Interviews

The profiles are of clusters, not individuals. Hence any individual learner may fit some but not necessarily all the characteristics of the cluster group into which she or he has been placed.

The following excerpts are from interviews with students from each of the four clusters. These illuminate the profiles described, yet present features of learning Statistics which are highly individual. Tilly and Ben were high achieving students.

Tilly achieved a final mark of 87.5 per cent on Statistics, while Ben attained 95 per cent. The written responses of these two students to the survey Question 2a indicated that they adopted similar strategies of reading lecture notes and doing examples. The intentions of the students were, however, quite different.

Tilly wrote for Question 2b that she was trying to achieve:

- A stable and satisfactory understanding of statistics and the ability to handle data confidently.

Ben, however, wrote that his intention was to:

- Get the right answers!

Their comments in the interviews substantiate these different approaches.

- When I'm trying to summarise my lecture notes I try to integrate all the information I have on each set topic. And understand that and therefore know how to apply it. What we're doing this semester is a lot more real statistics, how you really apply statistics to what you're doing.... Obviously on a higher level there's probably a lot more to it. I'll learn that in third year. But so far I think that's how psychologists would go about testing and researching these things. I'm interested in that sort of area–like experimental research, that sort of stuff. I want to learn it because I'm going to need the knowledge. And also
  - In a lot of ways it overcomes my frustration I had with maths last year. I looked at some of the work I thought last year was so difficult—and it's just basic to what we're doing now. So it's a triumph, almost, to have overcome that big barrier to statistics that I had built up last year.

Tilly evidently saw Statistics as a tool and one, furthermore that she intended to put to good use. Ben's focus on the other hand, was mainly on statistical processes–getting the right answer.

- It doesn't so much interest me, but it's easy going. You don't have to muck 'round doing reports and stuff. You can just learn it. And it's sort of half relevant to what I'm doing, and that's fine. It doesn't bother me a great deal if I don't understand, say, the exact theory behind different distributions and stuff like that. It doesn't really concern me a great deal. As long as I understand the basics and be able to just get through the questions.

Tessa despite taking the highest level of mathematics at school, achieved only 47.5 per cent for Statistics. This was mainly due to her poor performance in the multiple choice examination in first semester, where she attained only 20 per cent.

Her survey responses to the third open-ended question led me to categorise her conception of Statistics in the highest category–Critical Thinking. This was consistent with a relatively high score on the Deep scale of the ALSQ. In her survey she indicated a high level of engagement with learning Statistics and insight into statistical knowledge.

She wrote the following responses to the open-ended questions:

- Yes. Statistics is required for almost every field of study and is a very useful form of comprehensive analysis.
- From GSM notes, text books & lectures & tutorials. Statistics is structured mathematics and should. be considered as a friend to psychology.. I should enjoy learning its various methods to my advantage.

- A comprehensive understanding of everything for use in the Ψ department.
- Stats is about methodology which is used as a comprehensive form of analysis to interpret and test theories and correlations psychologists create. Substantiated method.

Tessa's interview responses confirmed that she was an exceptional student.

- It's very interesting–how statistics moulds itself into psychology. Right now I think psychology's trying to become a science because of the emphasis that people put on science being the main area of Abbreviation for psychology knowledge. The way statistics moulds itself into psychology kinda gives psychology a basis. Like–how would you put it–like raw facts that can be analysed scientifically, I guess. In 4 Unit Maths there wasn't that much application to things in real life. A lot of it was just formula based, following the formulas, plugging in the numbers. With statistics, however, you've got an aim. In a real life situation, in society–statistics is like a tool to analyse whatever happens when you do experiments.... It's helped me think more logically. Like inductive reasoning. It's a lot like inductive reasoning. You come up with a hypothesis and you have to follow through in order to get an answer.... I guess that's one of the main reasons I chose psychology and philosophy. I wanted a broader view of life. Science and maths in high school was more a regurgitation of theories. I just thought that in psychology and philosophy I might be able to contribute some new ideas.

Hence, despite her low mark in Statistics, Tessa seemed to have orientations to learning Statistics that most educators would find highly commendable. Her survey and interview responses indicated that she was exceptionally motivated to learn Statistics.

She evidently conceived of Statistics as a way of understanding experimental data and interpreting information from the world around her in a scientific way. Tessa's approach to learning Statistics seemed to have all the hallmarks of the deep approach–interest, personal engagement and relational understanding.

Narelle was classified into Cluster 4. She had studied General Statistical Methods in her first year of university, but attained 50 per cent for Statistics, well below the 58 per cent average for the surveyed students. Her score on the Surface Scale of the ALSQ was almost one standard deviation the mean while on the deep scale she was well below average.

Her survey responses to Question 2a and 2b were brief, indicating a lack of interest in the topic:

- Did exercises
- To pass the year.

In her interview, Narelle indicated a preoccupation with mechanical procedures, verifying our classification of her conception category as Processes. Her lack of relational understanding meant that her previous knowledge of statistics was confusing to her, rather than helpful.

An excerpt follows.

- I usually come home and have a look at what we've done in class and hope I understand it. If I don't then I'll perhaps contact my tutor, which I've only done once, or just keep doing the exercises until I work out what's going on. I want to be able to do it automatically rather than it be such a difficult process. I'd find what I need out of the problem, the mean and whatever, and then work through that–rather than read so much into the problem; what's going on with the rats or whatever. I'm trying to get the relevant information from the problem without being distracted by so much to think about, like the whole experiment. I found it just a little bit confusing because sometimes we have been doing the same sort of problem but doing it round a different way. I found that a little bit confusing. I mean, if I'm confused then those that didn't do GSM last year must be really confused. I don't enjoy it particularly. I enjoy it once I get an answer out I guess, but the working out is just not my thing. I'm not a very mathematical person.

Narelle, endured, rather than experienced, learning Statistics.

LEARNING STATISTICS AND APPROACHES TO LEARNING IT

How students' reported conceptions of Statistics relate to their approaches to learning it. The variable, Concept, denoting students' reported conceptions of Statistics, has five nominal values, according to the category in which the student's response was classified, namely No Meaning, Processes, Mastery, Tool and Critical Thinking. Students' approaches to learning Statistics are indicated by the scores obtained on the Deep and Surface Scales of the ALSQ.

The line graphs show the strong relationship between the categories of conception, obtained from the phenomenographic analysis of students' responses to the open-ended questions on the survey, and their scores on the two scales of the ALSQ.

I obtained for the deep approach versus conception category as follows:

- I calculated each student's average item score on the Deep Scale of the ALSQ.
- I then calculated the means of these for each group of students whose conceptions fell into the same category. Hence five means for deep approaches were obtained, one for each of the conception groups: No Meaning, Processes, Mastery, Tools and Critical Thinking.
- The five averages were plotted and the points joined in order to guide the eye.

I applied a similar procedure to obtain the line graph for the surface approach versus conception category. The means and standard deviations for each of the

five conception groups on the Deep and Surface Scales, as well as the number of cases for each group. In addition to showing the trends evident in the line graphs, these tables indicate the variability of the groups.

In particular, the standard deviation for the small NO Meaning group is comparatively high on the Surface Scale of the ALSQ as are the standard deviations for the Critical Thinking group on the Deep and Surface Scales. These groups are far from homogeneous. A discussion of the patterns follows these graphs and tables.

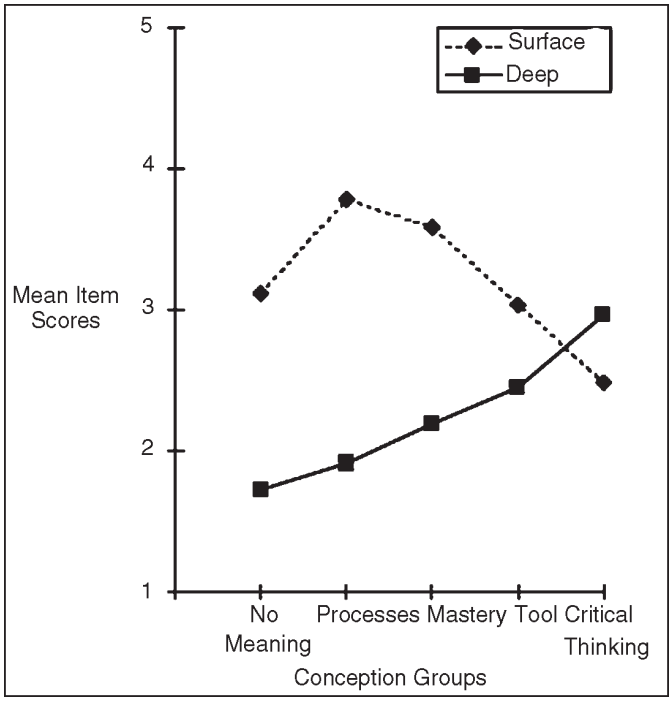


Fig. Mean Item Score for Each Conception Group on Surface and Deep Scales of The Alsq

Note: The Y-axis shows the range for Deep and Surface Scale items, from "1", indicating that the student adopts this approach "only rarely", to "5", indicating that the approach is adopted "almost always".

Table. Summary of Deep Scale Results for Conception Groups

Category	Item Mean for Group	Standard Deviation	Cases
No Meaning	1.69	0.39	11
Processes	1.91	0.62	66
Mastery	2.19	0.60	88
Tool	2.45	0.64	69
Critical Thinking	2.96	0.61	7
Overall	2.19	0.66	N=241

Table. Summary of Surface Scale Results for Conception Groups

Category	Item Mean for Group	Standard Deviation	Cases
No Meaning	3.11	0.98	10
Processes	3.79	0.66	66
Mastery	3.61	0.75	90
Tool	3.03	0.73	71
Critical Thinking	2.48	0.87	7
Overall	3.44	0.81	N=244

These trends were similar for males and females. Firstly, the graphs and tables portray the overwhelming preference for surface approaches to learning Statistics expressed by the participants in Study Two. Only for the small group categorised in Critical Thinking was the average score on the Deep Scale of the ALSQ higher than the average score on the Surface Scale. This is strong evidence for the effect of context as asserted by Vygotsky and the activity theorists. Students' approaches to learning Statistics depend on their interpretations of the learning task within its setting. In this case surface approaches to learning Statistics evidently fitted with these interpretations.

Secondly, reveals both a striking pattern and an apparent anomaly. The pattern relates to the evident strong relationship between the categories of conception and scores on the two scales of the ALSQ. On average, scores on the Deep Scale increase with each increase in conception category while scores on the Surface Scale decrease with each increase in conception category with the exception of the No Meaning group. The anomaly refers to the unexpectedly low average score of the No Meaning group on the Surface Scale.

I expected that the students who conceived of Statistics as a meaningless and pointless subject would tend to adopt surface approaches to learning it. This proved not to be the case. These students did, however, have low scores on the Deep Scale, as would be expected. The unexpected result led me to examine more closely the No Meaning students' responses to the open-ended survey question asking them about their strategies for and intentions in learning Statistics. These students were heterogeneous, having levels of prior mathematics ranging from 2 Unit Mathematics to university mathematics. However, my inspection revealed a common thread in their responses–these students evidently had a low level of engagement with the task. Their responses indicated that they were not attempting to memorise or rote learn the material presented, worrying about their performances on assessment tasks, attempting to conform to their teachers' ideas–they simply made the minimal effort required.

As one of these students wrote, in response to Question 2a, her strategy was to:

- Ignore it until I have to study for a test.

This finding adds further weight to my hypothesis that the No Meaning category does not form part of the hierarchical structure of the conception

categories. Hence, I suspended this group in some of the further analyses (correlations, factor analysis) carried out on the relationships among conceptions of Statistics and other variables, where hierarchical ordering of the selected variables is required.

For the four category groups: Processes, Mastery, Tool and Critical Thinking, an increase in students' awareness about Statistics was accompanied by a change in their reported approaches to learning it. The means on the Deep Scale increased linearly over the four conception categories. The means on the Surface Scale decreased linearly over the four 2 categories. As investigated, the ways that students conceive of the subject Statistics and their intentions and ways of learning it are strongly related. This supports Leont'ev's (1978, 1981) premise that mental reflections, goals and actions are inseparable. Students' conceptions of Statistics are expressed in actions taken, purposefully, to learn it and, conversely, their actions shape their conceptions of the subject.

PROBLEM-BASED LEARNING IN UNDERGRADUATE SCIENCE IN 21ST CENTURY PEDAGOGIES

In the PBL process, student learning is motivated using a problem, puzzle, or complex scenario presented in the same context, as it would be encountered in real life. Information needed to investigate the problem is not initially provided. Instead, when first presented with the problem, students organize their ideas and previous knowledge related to it, and attempt to define its broad nature. As they brainstorm initial hypotheses, the students find that they need to consult additional resources to fill in conceptual holes. They identify this needed information by posing questions that help to define why the information is needed – how it relates to the problem resolution.

They then assign one another the responsibility for researching this information before the next class meeting, and discuss the best sources (textbook, library, Internet, *etc.*) for finding the needed information. When students reconvene, they teach one another the results of their research on the questions posed in the previous session, ideally integrating their new knowledge and skills into the context of the problem.

The students continue to define new areas of needed learning (digging progressively deeper into the underlying content and assumptions) as they work through the problem, which typically enfolds in several stages through progressive disclosure. The PBL cycle thus provides a means for students to learn concepts in the context of their applications, and a forum for them to hone their ability to think critically, to find and process new information, to communicate effectively, and to become influential members of productive teams.

Problem-based learning as a formal cycle of teaching and learning strategies had its origins in medical schools to facilitate learning basic science concepts

in the context of clinical cases. In this earliest PBL setting, students typically worked in groups of eight to ten, guided by a faculty facilitator. The role of the group facilitator was to guide, probe, and support students' initiatives, not lecture, direct, or provide easy answers. The group facilitator thus monitored both the quality of the information and intellectual arguments, and the quality of group interactions.

This original medical school model is generally not applicable to typical undergraduate courses for many reasons, including class size. However, the collection of strategies encompassed by PBL can be reassembled in ways that are appropriate for undergraduate courses and students without necessarily comprising its essential nature.

For example, a faculty instructor can facilitate many classroom groups by serving as a "roving" facilitator who walks around the classroom to listen to groups' conversations and pose and answer questions when appropriate. The PBL instructor in this setting may choose to use PBL problems that are structured so that they allow for natural interventions for instructor guidance, whole class discussions, or short clarifying lectures to move students beyond conceptual impasses.

The whole class discussions can be tailored to provide feedback on group progress, to challenge students to examine the underlying assumptions of evidence they apply to the problem's resolution, and to guide students by encouraging them to consult good quality resources. Faculty using this model in essence supply in a structured and planned way the guidance supplied in a more extemporaneous and informal way in the original PBL model.

Another model for monitoring multiple PBL groups in an undergraduate classroom is the use of peer group facilitators – undergraduates who have completed a PBL course who return to work alongside the faculty instructor as guides for one or more groups. In both instances, the roving facilitator or peer facilitator model, PBL in the undergraduate setting is accompanied by use of typical cooperative or collaborative learning structures – that is, drafting by students of group contracts or guidelines, rotation of student roles as group members (for example, recorder, reporter, discussion leader, accuracy coach), and peer evaluation of performance as group members.

The relative absence of PBL instructional materials can be a major roadblock to implementation of PBL in a particular discipline. To incorporate the goals of PBL, these problems must be complex enough to engage students attention for several class periods or longer, and challenge students to think critically as they explore new content areas.

Fortunately, collections of these materials are becoming more readily available – for example, the PBL Clearinghouse (accessible through a secure server at https://www.mis4.udel.edu/Pbl and the LifeLines OnLine collection offered through BioQUEST available at http://bioquest.org/lifelines; see PBL resources list).

My first encounter with PBL was introduced to the University of Delaware (UD) in 1992 in a workshop to prepare faculty to teach in a new Medical Scholars

Programme. A group of science faculty immediately recognized the potential of the method and adapted it to their introductory science courses.

PBL at UD has since expanded to impact thousands of students and 30 per cent of the faculty in an array of undergraduate disciplines. My role in this effort was to co-design and implement a number of PBL courses in the biological sciences (both small and large enrollment), and to develop (with Harold White in UD's Department of Chemistry and Biochemistry) a formal in-service support programme for undergraduate peer facilitators.

Materials and methods from one of these courses (introductory biology for science majors) are published in Thinking Towards Solutions, and summarized in a recent essay on PBL in biology in Cell Biology Education.

With a group of PBL faculty on my campus, I also co-founded and continue to co-lead a faculty institute that was created in 1997 with support from NSF's Institution-Wide Reform programme. The institute uses a "faculty mentoring faculty" model to help faculty to develop a new PBL course or adapt an existing to incorporate active, group learning strategies such as PBL.

ADAPTING PROBLEM-BASED LEARNING TO THE UNDERGRADUATE SETTING

It was clear from the outset that the original medical school model for PBL would have to be adapted to fit the greater class sizes (and therefore greater demands on faculty resources) and less intellectually mature population of learners in the undergraduate setting. Fortunately, PBL is not a single strategy, but rather a collection of strategies that can be assembled in many combinations, and thus lends itself to adaptation without necessarily comprising its essential nature. Nevertheless, the challenges of implementing PBL in an undergraduate setting are numerous. In the process of implementing PBL, we discovered many different strategies that could be used in the design of PBL activities for courses with different enrollment sizes, learning objectives, and populations of learners, and by faculty with varying perspectives and time constraints.

FURNISHINGS OF EXISTING CLASSROOMS

The layout and furnishings of existing classrooms were an obvious barrier to implementation of the group learning aspect of PBL. Fortunately, the University of Delaware's administration responded readily to faculty requests by funding and expediting the renovation of PBL classrooms as the need for them grew. These classroom renovations were designed to maximize blackboard space (for student groups to use when reporting on their research), include tables for group work, and provide cabinets for storing resource materials between classes.

Nevertheless, some faculty using PBL in large classes (in which enrollment exceeds 80 students, the capacity of the University's largest PBL classroom) are still challenged by classrooms where seating is fixed. When possible, instructors in these classrooms typically under enroll them, so that the seating plan can include vacant rows for greater ease of instructor access to student

groups. Groups are typically four students, so a group can use two adjacent rows by having the two students in the front row turn in their seats to more immediately face the remaining two students in the row behind.

MONITORING MULTIPLE GROUPS

The most daunting challenge is how faculty can facilitate the PBL efforts of many classroom groups simultaneously. In the earliest model of PBL, an expert facilitator guides the group process by observing, asking questions, and intervening when appropriate. The facilitator should also prompt the group to dig deeper into content, ensure that students make connections and tie information together, keep students on track during discussions, help to locate resources, examine evidence that can be used to support conclusions, ensure that all students are involved in the process, model the process of giving and receiving feedback, and help the group learn to plot its own course. Clearly, few undergraduate classes are small enough that the instructor can be a dedicated facilitator of a single small group of students in this intensive fashion. A roving facilitator strategy that is used in many cooperative learning settings, also works well in PBL classrooms.

In this model the instructor walks around the classroom to observe groups in action, looking for signs of engagement with the problem and for the participation of all students in their group discussions. Typically the roving instructor poses questions that encourage students to dig more deeply into essential content or get beyond a conceptual impasse.

The instructor looks for signs of behaviours that seem counterproductive to good group function and may enter into discussions when appropriate. This roving facilitator strategy works well with use of PBL problems constructed to provide natural break points that allow for insertion of instructor-led discussions.

Whole class discussions at key intervals in the problem-solving process allow the instructor to provide feedback and model the process of evaluating resources, and analyzing and summarizing information. Groups that are moving at a substantially slower pace can benefit from hearing about their peers' progress, but care must be taken to prevent groups from intentionally piggy-backing on other groups' efforts. During these whole-class discussions, the instructor can provide, in a more structured and formal way, some of what the classic PBL facilitator contributes when s/he guides a single group.

Assigned roles of responsibility and drafting of group guidelines - strategies commonly used in cooperative learning classrooms - can also work effectively in PBL.

The roles of discussion leader, reporter (for group products and class discussions), recorder, and accuracy coach (the "skeptic") may rotate among group members on a regular schedule or with each new problem.

Students draft ground rules for effective group work that typically include their notions of acceptable attendance and preparedness, and include penalties for non-adherence that escalate with each incident.

Peer facilitators can also be effective as group monitors. Students who have successfully taken the course return to guide student groups as either a roving or dedicated facilitator, working alongside the faculty instructor. The use of peer facilitators has proved to be an excellent model for enhancing the effectiveness of classroom groups and is a model that can be extended to active learning activities other than problem-based learning.

THE LARGE CLASS

Instructors in large classes enlist the help of undergraduate and graduate TAs to have more individuals to monitor groups. They use carefully staged problems that allow the instructor to intervene at roughly 15-20 min intervals to help guide progress through the problem. The instructors typically choose to implement closely defined group monitoring strategies such as rotating roles and ground rules. They ask students to record their roles each week or with each problem to verify that roles have actually rotated among group members. Students are asked to make suggestions for policies and penalties for group guidelines. Group evaluations often include ratings of each other's contributions to assignments and products. Highly streamlined versions of the written and verbal feedback strategies used in smaller-class PBL can be used.

PBL instructors of large enrollment classes also intersperse other classroom activities between the PBL problems. While the PBL problem often serves as the central focus, lectures, discussions, and short active learning activities associated with the problem are used to help students build conceptual frameworks. In "Who Owns the Geritol Solution Problem" a concept mapping exercise has been used. Between the first two stages of the problem, students are given map titles such as "the light-independent reactions of photosynthesis," "the carbon cycle," "the Geritol solution," or "the flow of energy through the biosphere" for which the student group constructs a concept map. This helps provide timely feedback to both students and instructor about whether the major concepts evoked by the problem have been understood and synthesized.

PBL strategies for large classes include use of hybrid models in which, for example, four to six shorter problems are presented for each major content unit. In these hybrid models, traditional methods such as lecture are used to support the PBL instruction but do not supplant it. In these large classes, the problems typically are used in association with clearly defined final products, such as whole class debates, position papers, and mock town meetings, trials or congressional hearings, that are naturally embedded in the context of the problem. For instructors reluctant to commit to using PBL throughout the entire semester, even use of a single problem of several weeks duration can be effective 001). For instructors of large enrollment course who find the management of large numbers of classroom groups to be a daunting prospect, discussion, recitation or laboratory sections of the same course can provide the requisite smaller class setting for PBL. If using this strategy in conjunction with teaching assistants, however, care must be taken to select individuals who support the underlying goals and assumptions of PBL.

Is there a limit to how large a PBL class can be? Shipman and Duch compared selected outcomes in a class of 120 with those in a class of 240 students. Their preliminary findings suggest that PBL can work in the class with the larger enrollment. Students reported that problem solving and group work enhanced their learning and helped prepare them for their working lives, and these perceptions were backed up by independent assessments of classroom performance. From the perspective of both students and faculty, the PBL experience was better in the smaller class. Students reported more positive attitudes towards, and greater interest in learning the subject in the smaller class, and instructors were not enamored of the sheer magnitude of the management task involved in monitoring up to 60 classroom groups.

SOURCES OF PROBLEMS

A major roadblock when PBL was first implemented in undergraduate courses, particularly in the introductory basic sciences, was the absence of suitable problems. To meet the goals of PBL instruction, problems must be able to engage active, cooperative learning activities within student groups for up to a week or more. End-of-chapter textbook problems in general do not have the contextual richness, nor do they require the analytical, synthetic, and evaluative thinking needed for PBL. Consequently, a major hurdle for adapting PBL was the necessity to write problems appropriate to the instructional goals.

While that hurdle ensured that only fully committed instructors became involved, it undoubtedly discouraged others from trying. Fortunately, this barrier is being lowered as more and more faculty drawn to PBL turn their creative energies towards writing and disseminating college level course materials.

PBL problem writers seek inspiration from a variety of sources, including fictionalized composites of events in a typical person's, articles from the popular press about inventions and discoveries, science and society interactions, or landmark experiments. Harold White (Chemistry and Biochemistry) and David Sheppard at the University use problems that provide explicit models for research. In White's Introduction to Biochemistry course, students build their understanding of what biochemists have learned about a specific area of biochemistry by reading a carefully selected series of articles that document the history of the seminal experiments in the field.

In a course required for biology majors, Sheppard asks students to mine research quality nucleic acid and protein databases to resolve problems about important recent experiments in the field of genetics. In both cases, students discuss and resolve these research-oriented problems, using the PBL process.

Although the sources of problems and the contexts for their classroom use may vary, PBL problems have common features. To be appropriate, problems should engage students' interest and motivate learning, require students to develop a line of reasoning that is backed up by evidence, be complex enough to motivate participation of a group of students rather than just a single individual, be open-ended enough at the outset to allow participation by all students,

incorporate the learning objectives of the course and allow for many legitimate resolutions or many paths to a single resolution. These objectives are embedded in the problem, rather than posed separately or otherwise set apart by the instructor.

The problem "Riverside's Dilemma" exemplifies the PBL process as it unfolds in a chemistry context. The problem, written for use in general chemistry courses, presents students with the dilemma faced by a town council that must decide on allowable limits for wastes flowing into a local river. In working on this problem, students encounter concepts related to the chemistry of weak and strong acids and bases, neutralization reactions and related equilibrium calculations. In Stage One, Riverside's town council receives a proposal from a multinational chemical corporation to buy several old, closed-down factories. The dilemma is that the old design of the factories would result in discharge of wastes directly into the river; retrofitting of the factories seems prohibitively expensive.

Placed in the role of a consultant to the town council, students must consider the impact of these potential waste streams on the health of the river. Knowing the amount and types of emissions projected for each plant (various strong and weak acids and bases), students first determine what level of dilution of each waste stream would be needed to bring it to an acceptable pH value, then decide whether dilution would be an effective strategy. In doing so they are introduced to the concept of neutralization and its relationship to dissociation equilibria. In order to refresh students' memories of high school chemistry, the students complete a quiz - first as individuals, then within groups – on the basic terminology of acid-base chemistry.

In Stage Two of "Riverside's Dilemma," students realize that dilution is not a reasonable solution because of the very large volume of water that would be required. They now must consider the feasibility of neutralizing the waste streams to achieve an acceptable pH range in the river. Students, still in the role of consultants, must consider the use of two relatively inexpensive neutralization agents and determine how much of the appropriate agent would be required. This stage helps students discover the relationship between neutralization and dissociation reactions. Stage Two, like Stage One, typically requires one 50 minute class period for resolution.

The third and final stage of the problem is the most difficult and open-ended. Students must pull together the material dealt with in the previous two stages to consider a more complex situation. The town council has requested that the student consultants evaluate the feasibility of combining some of the waste streams before they are released into the river in order to bring pH into the acceptable range. In addressing this problem, students must recognize which combinations of waste streams are possible, decide which combinations constitute neutralization processes and of these, which have equilibrium constants appropriate for the desired results.

A stage-by-stage account of how a biological problem is worked out in the PBL classroom can be found in the University of Delaware's PBL Clearinghouse.

This problem centers around John Martin's novel scheme to cure global warming by seeding the oceans with iron to harness the latent primary productivity of marine phytoplankton. The problem motivates students to research major concepts related to cellular energy transformations, biogeochemical cycles, global climate change, and marine ecosystems.

"A Bad Day for Sandy Dayton" provides an example for how an introductory physics problem is constructed and unfolds in the classroom. This problem is designed to help non-science majors understand forces, motion and mechanical energy by reconstructing a rear-end auto collision that occurs outside their classroom. In doing so, they explore the relationship between speed and stopping distance, reaction time and stopping distance, and the design and safety features of seatbelts and airbags.

Several collections of problems are available on-line for instructors who would like to use PBL. These include the University, which contains problems and teaching notes for the sciences and disciplines outside the sciences, Life Lines On-Line, a collection of introductory life sciences problems produced through collaboration between Southeast Missouri University and the BioQUEST Curriculum Consortium, The National Center for Case Study Teaching in Science's which includes some appropriate for PBL instruction, a library of several hundred engineering case studies available through Carleton University, and a set of pharmacology problems.

Books containing PBL problems are not common but a collection of problems with teaching notes for general biology is available. There are also several books of case studies that provide ideas that can be adapted for use in PBL. To be appropriate for PBL, the cases would need to be written in a format of progressive disclosure, and the content would have to be reduced so that students would be motivated to do independent research.

EVALUATION OF PBL OUTCOMES

Faculty who use PBL instruction want students to develop skills such as the ability to find and analyze information from a multitude of sources, to engage effectively in self-directed study, to communicate well using diverse media, and to work productively with a group of peers.

These goals are often assessed by comparing student performance several times during the semester at such tasks as giving oral presentations, writing reports, or answering exam questions. Classroom observations and peer evaluations of group performance are helpful in assessment of students' contributions to their groups.

Development of other skills, such as the ability to reason critically and creatively, and to make reasoned decisions in unfamiliar situations is not so easily assessed. Documentation of student attainment is hampered by the lack of instruments that have the sensitivity needed to detect changes in critical thinking (as defined by particular instructors or within particular disciplines) over the course of a semester.

PBL was not designed as a way to enhance content understanding alone, but the constructivist nature of its approach often invokes concerns about whether students are learning essential course content. Specific experiences with PBL, and meta-analyses of outcomes from PBL curricula in the medical school context have shown that content learning in PBL matches that in a traditional curriculum. Additional outcomes in PBL include greater retention of knowledge and greater satisfaction with the educational experience.

Data comparing traditional and PBL classrooms can be more difficult to obtain in the undergraduate setting, with its diversity of majors and tracks. Disciplines such as physics, for which national standardized surveys of content learning outcomes exist, provide some comparison data. Williams reports gains in force concepts inventory scores for a PBL course in introductory physics that are nearly twice those found in courses using traditional methods. This is a common finding for physics courses that use active learning methods. However, care must be taken in interpreting the outcomes of these scores or those from any multiple choice pre- and post-test, since they capture only one of the goals (content understanding) of PBL instruction, and because students in PBL courses typically do not encounter multiple choice tests in the course of the semester.

Conversely, it would be inappropriate to evaluate students in lecture-based courses with instruments that assessed PBL's additional goals if these students had had little opportunity to practice these skills in their courses.

The University of Delaware recently completed a broad study on PBL outcomes that was funded by the Pew Charitable Trusts. This study sought to document the instructional use and impact of PBL at the undergraduate level. Preliminary results include the following findings. Exposure to PBL positively and significantly affected the number and/or quality of student-faculty interactions, as well as the number of diversity-related experiences in which students participated. Exposure to PBL was defined as a composite measure of the number of PBL courses completed by each students and students' report of participation in several activities that collectively represent PBL activities in the classroom.

Discussion with students in structured focus groups offered further insights into students' learning experiences. Students indicated that the collaborative nature of PBL increased their level of comfort and inclusion in the class. In addition, students believed that their learning was enhanced because PBL increased their ability to consider, evaluate, and respect different points of view.

They discovered that there might be more than one good answer to a problem or an issue developed within the context of a problem. The PBL setting helped students to apply theory to real world issues, made course content more interesting, and helped them to learn course content more thoroughly. Students also believed that their communication and interpersonal skills had improved as a result of participation in PBL courses.

While gains in critical thinking skills were measured (using standard instruments such as the Watson-Glaser Critical Thinking Appraisal), it was not

possible to conclude that these changes were due to PBL, or to other aspects of the students' undergraduate experiences. Interpretation of findings was hampered by the presence of many different models of PBL instruction, including hybrid ones, on the University of Delaware campus, as well as the lack of a uniquely PBL curriculum track that could be compared to a more traditional track, as is the case in some medical school settings.

INSTITUTIONAL COSTS OF PBL TEACHING

Transformation of courses to incorporate PBL strategies has some costs associated with it. Classrooms were refurbished to include new furnishings and seating arrangements more conducive to group work; while these classrooms are ideal, they are not essential to successful use of PBL strategies. Other costs resulted from the additional demands placed on faculty time in the early phases of PBL adoption, when materials and activities were planned and developed. Outside consultants were brought in the first few years of PBL adoption on our campus to assist with faculty development in PBL instruction, but in later years, University of Delaware faculty took over this role. These faculty members formed a PBL institute supported in part by an extramural grant to provide the training and mentoring often needed by faculty attempting to redefine their teaching. Faculty incentives to attend the institute and to transform their courses were provided by a match of institutional funds to funds from extramural sources, including the National Science Foundation, the Pew Charitable Trusts, and for biomedical sciences faculty, the Howard Hughes Medical Institute.

These incentives have taken the form of professional development accounts through which faculty can purchase materials, hire technology assistants for aspects of course design, or attend and present at education-related conferences in their scholarly disciplines. At the University of Delaware, there was no reduction in class size with adoption of PBL, so no additional faculty time was needed to accommodate greater numbers of course sections; faculty adopted PBL strategies that would work within existing class sizes. Although having additional graduate teaching assistants (TA) to help facilitate student groups might have been ideal, reallocation of TAs towards PBL courses or creation of new TA lines did not occur.

It is important to point out that these costs are mitigated by the way in which PBL contributes to a unique definition of instructional productivity. That is, PBL allows more students and student hours to be engaged in educational activities that resemble those of a faculty-directed undergraduate research experience without the associated costs of such one-on-one faculty-student interactions. Nearly 100 per cent of the science and engineering faculty at the University of Delaware already take undergraduate collaborators into their research and they are now serving as many undergraduate researchers as they can handle. Therefore, this classroom-based mode of student engagement in the discovery process effectively scales up research-based learning so that all undergraduates can benefit.

5

Educational Measurement

Science builds on an evolving network of measurement. The realization that scientific measurement has special characteristics is not new. Campbell (1920) showed that what physical scientists mean by measurement requires an ordering system and the kind of additivity illustrated by physical concatenation. Campbell called this "fundamental measurement."

FUNDAMENTAL MEASUREMENT

Intelligence test scores with which arithmetic could be done were called for by Terman and Thorndike before 1920 and required for measurement by Thurstone in the 1920's. Thurstone (1925) constructed some rather good interval scales with his absolute scaling. With his Law of Comparative Judgement he did even better, producing results that are successful instances of fundamental measurement (Thurstone, 1927). Insistence on order and additivity recurs in Guilford's (1936) definition of measurement. The significant consequence of additivity is the maintenance of the unit of measurement and hence of the invariance of comparisons of measures across the scale.

During 1940's Guttman realized that a test score would be ambiguous in meaning unless the total score on the items was sufficient to reproduce the response pattern the score represented. This led Guttman (1950) to a criterion for judging whether data were good enough to build a scale. The data must demonstrate a joint order shared by items and persons.

During 1950's the Danish mathematician George Rasch found that he could not obtain an invariance of test item characteristics over variations in persons

unless he could represent the way persons and items interacted to produce a response in a special way. It is possible to represent a response by an exponential function in which the participation of person and item parameters could have a linear form (Rasch, 1960, p. 120).

Rasch also noted that invariance could be maintained only when data could be gathered so that they cooperated with a stochastic response model that produced a joint order of response probabilities-similar to the joint order Guttman called for (Rasch, 1960, p. 117). As he worked with this first "Rasch model," Rasch discovered that the best (*i.e.*, minimally sufficient) statistics from which to estimate person measures and item calibrations were none other than the good old unweighted sums of right answers for persons and for items—the familiar raw scores.

Then Luce and Tukey (1964) showed that an additivity just as good for measurement as that produced by physical concatenation could be obtained from responses produced by the interaction of two kinds of objects (*e.g.*, persons and items), if only this interaction were conducted so that its outcomes (*e.g.*, item response data) were dominated by a linear combination of the two kinds of quantities implied (*e.g.*, the differences between person measures and item calibrations). Luce and Tukey showed that if care taken; their "conjoint measurement" could produce results as fundamental as Campbell's fundamental measurement. For Luce and Tukey, the moral seems clear: *when no natural concatenation operation exists, one should try to discover a way to measure factors and responses* (*e.g.*, gather data) *such that the 'effects' of different factors are additive* (p. 4).

The realization that Rasch's models could do this when data were collected carefully enough followed (Brogden, 1979; Fischer, 1968; Keats, 1967). Perline, Wright and Wainer (1979) provide empirical demonstrations of the efficacy of the Rasch process in constructing fundamental measurement. When Andersen (1977) showed that the sufficient statistics that enable fundamental measurement depend on scoring successive categories with the equivalent of successive integers—exactly the ordered category scoring most widely used on intuitive grounds. The Rasch model for dichotomously scored items was extended to response formats with more than two ordered categories (Andrich, 1978; Wright & Masters, 1982).

COMMON PRACTICE

As this brief history shows, there has been enough successful theoretical and practical work on the nature and implementation of fundamental measurement to establish its necessity as a basic tool of science and its ready accessibility for educational researchers.

Fundamental measurement is obtained from educational data and, in an intuitive form, educational researchers have relied on it for a long time. In their use of unweighted raw scores and successive integer category weights, educational researchers have been practicing the scoring required for

fundamental measurement all along. Intuitive rather than explicit reliance, however, has meant that they have neither recognized nor enjoyed the benefits of fundamental measurement, and they have not built on its strengths to improve their understanding of education.

ILLITERATE THEORY

The ready accessibility of fundamental measurement and its necessity for the successful practice of science is still unknown to most educational researchers and misunderstood by most psychometricians. In spite of 60 years of literature explaining the strong reasons for and illustrating the successful application of fundamental measurement in educational research, few psychometricians and fewer educational researchers attempt to construct fundamental measures. Much of the recent psychometric writing and practice goes on as though no knowledge concerning the theory and practice of fundamental measurement existed. In books claiming to provide the latest ways to make measurements with educational data, one might expect that an exposition of fundamental measurement would not only appear but would dominate the discussion, the choice of methods advanced, and the choice of data analysed. In *Item Response Theory* by Hulin, Drasgow and Parsons and *Applications of Item Response Theory*, edited by Hambleton, there is no discussion of the nature, meaning, or practice of fundamental measurement.

Indeed, a theory of measurement is denied in Hulin, Drasgow and Parsons, who say "in social science we have no well-articulated meta-theory that specifies rules by which one can decide among competing (item response) theories on the basis of the propositions, assumptions, and conclusions of the theories" (p. vii). The possibility of a relevant theory is despaired of in Hambleton, where the "ineluctable conclusions" are "that no unidimensional item response model is likely to fit educational achievement data" and "the Rasch model is least likely to fit" (Traub, in Hambleton, p. 65).

Response models for one, two and three item parameters are described in detail but their motivation through Thurstone's (1925) assumption of a latent response process of continuous distribution is unnecessarily abstract and circuitous. It would be so much better education and science to begin with a binomial model for the observable dichotomy as Rasch does (1960, pp. 73). That simple story is not only easier to follow but leads directly to fundamental measurement.

These two books are rife with the usual misunderstandings. The three parameter model is called a logistic model although, when guessing varies, Birnbaum (1968, p. 432) says it is not. The estimation virtues extolled by Swaminathan in a chapter featuring the three parameter model apply, in fact, only to the one parameter Rasch model. Indeed "the likelihood function (of the three parameter model) may possess several maxima" and its value at infinite ability "may be larger than the maximum value found" when ability is finite (Swaminathan, in Hambleton, p. 30).

The Rasch model is referred to as a special case of the three parameter model when, in fact, its singular significance for measurement is that it is a unique (necessary and sufficient) deduction from the (fundamental) measurement requirements of joint order and additivity.

The sum of discrimination estimates for items answered correctly, which cannot be a statistic because a statistic must be a function of data and not of other statistics, leads a double life. It is called a sufficient statistic (Bejar, in Hambleton, p. 11; Swaminathan, in Hambleton, p. 30) and then shown to be insufficient because "it is not possible to extend the conditional (estimation) approach to the two parameter model" so that "sufficient statistics exist for ability... only in the Rasch model" (Swaminathan, in Hambleton, p. 36).

Bejar worries that "the logistic model ignores the difficulty of the items answered correctly in assigning a score" (Bejar, in Hambleton, p. 14).

But that is exactly what raw scores have always done. Although the score pays close attention to the difficulty of the test, it must be indifferent to items within the test which are answered correctly. This indifference is a necessary consequence of the local independence assumed by all of these models and required for any item banking.

The study of items which a person answers correctly, that is, the investigation of the joint order between observed person responses and estimated item difficulties, is the study of person fit. Does the person's data fit the measurement model? Is the person's performance valid? Indifference as items which within a test are answered correctly is necessary for building measurement. But the measure constructed from this indifference is only valid when the person's performance pattern is stochastically consistent (jointly ordered) with the items' difficulties.

IMPRACTICAL PRACTICE

Methods for using these response models are left to computer programmes, mostly LOGIST. Lord help the poor researcher who hasn't got his LOGIST working. There are no explicit procedures, detailed examples, or even estimation equations provided. Any reader hoping to learn how to apply IRT will have to look elsewhere (*e.g.*, Rasch, 1960; Spearritt, 1982; Wright & Stone, 1979). Even in Hambleton's Chapter, which is dedicated to LOGIST, there are no instructions for how to use it. The basic problems with the two and three parameter models are made plain by the steps taken to deal with their symptoms. During "estimation in the two and three parameter models... the item parameter estimates drift out of bounds" (Swaminathan, in Hambleton, p. 34). Thus "Range restrictions (must be) applied to all parameters except the item difficulties" to control "the problem of item discriminations going to infinity" (Wingersky, in Hambleton, pp. 47-48). Worse than that, "items with vastly different discrimination and difficulty parameters (can) nonetheless have virtually identical ICCs in the ability interval" where the data are present and the estimation work is to be done (Hulin, Drasgow, & Parsons, p. 100).

All this happens because an empirical binomial ICC expressed on a scale that must be defined by the same data contains only enough information to identify one item parameter. The only way more item parameters can be estimated is to assume that the particular persons participating in the item calibration are random examples of some "true" distribution of ability for which these items are always to be used.

The problem is aggravated by the way these models use ability in two ways at once. The first use is as a "difference" that specifies a distance between person's ability and item difficulty. This difference is essential for the construction of measurement. The second use is as a "factor" to multiply item discrimination so that a different unit can be specified for each item. This interacts with the first use to confound estimation and prevent the construction of joint order or additivity. The non-linear combination of item parameters prevents the algebraic separation of difficulty and discrimination and hence the derivation of sufficient statistics for estimating them. When ICC's are allowed to cross (*e.g.*, Hambleton, pp. 45, 46, 163), the manifest difficulty order items varies with ability. This prevents the construction of a variable that can be defined in any general way by the relative difficulties of its items.

The estimations of item discrimination and person's ability are based on a feedback between (i) summing the product of observed response and current ability estimates over persons and (ii) summing the product of observed response and current discrimination estimate over items (Birnbaum, 1968, pp. 421-422). This process cannot converge because the cumulative effect of the feedback between ability and discrimination pushes their estimates to infinity.

Yen describes the kind of trouble this can get one into: The biggest surprise occurring with the CTBS/U interlevel linking related to the type of scale produced.

As grade increased, the scale scores had decreasing standard deviations and corresponding decreasing standard errors of measurement and increasing item-discriminations.... This result was unexpected because the scaling procedure used with previous tests, Thurstone's absolute scaling, produced a scale with standard deviations increasing with grade. (Yen, in Hambleton, p. 139).

When the response model allows person's ability and item discrimination to interact, this is bound to happen. The decreasing standard deviations are not describing children on an interval scale. They are describing the consequences of an estimation procedure that cannot keep item discriminations from drifting towards infinity.

Parameterizing guessing is much prized in theory. But "attempts to estimate the guessing parameter... are not usually successful" (Hulin, Drasgow, & Parsons, p. 63). In one study, "40per cent of the guessing parameter estimates did not converge even with a sample size of 1593" (Ironson, in Hambleton, p. 160). Even when some estimates of guessing are obtained there are problems.

"If a test is easy for the group (from which guessing parameters are estimated) and then administered to a less able group, the guessing parameters (from the

more able group) may not be appropriate" (Wingersky, in Hambleton, p. 48). "When dealing with three parameter logistic ICCs, a non-zero guessing parameter precludes a convenient transformation to linearity" (Hulin, Drasgow, & Parsons, p. 173). None of this should be least surprising. The formulation of the model shows quite plainly that the explicit interaction between guessing and ability must make any guessing estimates inextricably sample dependent.

ANALYSIS OF FIT

The *analysis of fit* for persons and items is addressed at length—two chapters in Hulin, Drasgow, and Parsons, and three in Hambleton. The techniques mentioned and their sources are basically the same. But the substantive discussion and motivation are much better in Hulin, Drasgow, and Parsons. Unfortunately, as with estimation, the methods described are made to seem unnecessarily complex and arcane. If a reader want to apply one of them, he or she will have to go elsewhere to learn how.

Harnisch and Tatsuoka call for person fit statistics with standard normal distributions and no correlations with ability (in Hambleton, pp. 114, 117-119). That is a good idea when fine tuning fits statistics against data simulated to fit. But they use their ideal to evaluate fit statistics applied to real data. With real data one would prefer fit statistics that are skewed by misfit and correlate negatively with ability (to detect the guessing of low ability persons). But these criteria are the opposite of the ideal Harnisch and Tatsuoka use in their comparisons.

A special technique for fit analysis used by Yen (in Hambleton, p. 126) on CTB items seems attractive at first, but there is good reason to avoid it. Yen's first step is to smooth out the misfit occurring in her data by regressing observed scores of ability-groups on a monotonic function of their estimated expectations.

Then she makes her "fit" comparisons, not against the regression residuals into which the misfit has just been pushed, but against the regression predictions from which the misfit has just been removed. The reason her "fit statistic is more stable" is that she is no longer analysing misfit.

Ironson does a chapter (in Hambleton, Chapter) on item bias, but the topic is better dealt with in Chapter of Hulin, Drasgow, and Parsons. The trouble with both approaches to bias is that bias found for groups is never uniformly present among members of the group or uniformly absent among those not in the group. For the analysis of item bias to do individuals any good, say, by removing the bias from their measures, it will have to be done on the individual level of the much more useful person fit analyses described in other chapters.

TEST EQUATING

Test equating appears in Chapter of Hulin, Drasgow, and Parsons, only on the way to connecting tests written in more than one language, whereas Hambleton provides a chapter eleven on equating. Unfortunately item banking, the purpose of test equating, is not mentioned. Neither book describes the actual details of equating sufficiently well to enable the reader to learn how to do it.

There is a refreshing irony in the various accounts (Hulin, Drasgow, & Parsons, pp. 174, 202; Hambleton, pp. 45-46, 132, 178, 181-182; 191) of how equating with the three parameter model is actually accomplished. First of all, whereas the three parameter model is always claimed, the two parameter model is the one actually used to calibrate most of the tests to be equated. This happens because the guessing parameter proves too elusive to be kept as a variable. Then, when it comes to actually connecting two tests, even item discrimination is given up. The actual equating is based entirely on item difficulty, the one item parameter of the Rasch model. In other words, whenever these authors get involved in actually building a measuring system, even of only two tests, they are forced by what then happens to them to use the only response model that can build fundamental measurement, namely the one item parameter or the Rasch model.

The irony is particularly poignant in the Cook and Eignor attempt to distinguish between three methods for equating tests (in Hambleton, p. 191). Test equating must result in a single linear system of item calibrations and person measures (associated with test scores), which is invariant (but for scale and origin) over the data and also over the methods used to construct this system. Unless all three methods produce the same result, none of them has worked.

These authors claim concern with advancing science by providing better methods for building knowledge. But they offer models for item response data that systematically prevent the construction of fundamental measurement. The only way educational researchers can hope to build knowledge is to use models that insist on data from which fundamental measurement could be constructed—models that contain in their formulation the joint order and additivity conditions required for its construction.

For most of the models presented in these books to produce meaningful item estimates, it is necessary to assume that the persons who provided the data are random events sampled from the "right" simple standard distribution. Persons do not appear as individuals to be measured in these models. Person parameters are not even subscripted in most presentations.

The scale reported is routinely standardized to give the sample person a mean of zero and a variance of one as though those particular persons were a true random sample of exactly the population of persons with whom the scale would always be used—and as though any individual person subsequently measured could be usefully understood as no more than a random instance of the particular sample of other persons with whom the item analysis was done.

FIGMENTS OF DESPAIR

Despair is the latent message in books like this—despair of ever constructing any good mental or psychological variables. Review the testimony. The item response models described are said to be:

- *Hard to Understand*: The "procedures involved in estimation are complex and require sophistication on the part of the user" (Bejar, in Hambleton, p. 3).

- *Difficult to Use*: "The difficulty in applying these models is stressed" (Bejar, in Hambleton, p. 1). "Working with IRT is an arduous process" and "the largest hurdle is the estimation" (Bejar, in Hambleton, p. 3). Would-be users are warned that they "must be ready to pay... by investing substantial resources in parameter estimation and model monitoring" (Bejar, in Hambleton, p. 17).
- *Demanding of large samples of persons and many test items* (Bejar, in Hambleton, p. 3; Ironson, in Hambleton, p. 160; Wingersky, in Hambleton, p. 46). "Even long tests and large samples do not necessarily allow accurate estimation of the guessing parameter" (Hulin, Drasgow, & Parsons, p. 100).
- *Unreliable in Action*: Discrimination eludes capture because of its interaction with ability. Guessing can't be laid hold of, not only because it is persons and not items who are doing it, but because when plausible estimates do emerge they are sample dependent.
- *Hopeless in any Event*: "Multiple-choice items... are unlikely to be modelled very well by any unidimensional item response model" (Traub, in Hambleton, p. 57). "Speeded administration and an item format that permits guessing will each introduce another trait into the response process" (Traub, in Hambleton, p. 62). "Data from examinees who vary in their guessing propensities will result in systematically biased calibrations of items and systematically biased estimates of examine's ability" (Traub, in Hambleton, p. 63). As for using a model to select data good enough to make measurements with, "It will be a sad day indeed when our conception of measurable educational achievement narrows to the point where it coincides with the criterion of fit to a unidimensional item response model" (Traub, in Hambleton, p. 64).

If this were all we had to look forward to, the future of educational research would be dim indeed. Fortunately there is a readily accessible and remarkably hearty antidote to all the confusion and despair, even to Traub's terrors.

INVESTIGATIVE AND COMPARATIVE STUDIES

All fields of study appear to be marked by similar phases of growth. In the beginning, contributions to the field tend to be discrete and unsystematic, prompted by the curiosity of the observers and their inherent interest in the subject. There are no rules, just the special insights and motivations of single observers, whose accounts are descriptive and usually lack systematic reporting or an expressed framework of theory. As work in the field of study increases, reporting becomes more systematic and comprehensive, and the reporters are more self-conscious about the accuracy of their data and more concerned about the ways in which they arrive at conclusions from the data.

The curiosity of observers becomes more focused on the possibilities of practical applications of new knowledge, and they are inclined to be critical of their own

work and that of their colleagues. Particular types of studies emerge, marked by particular theoretical approaches to the subject, characteristic ways of observing and reporting, and broad agreement on what is or is not relevant. Practitioners in the field become aware of the precedents and of their intellectual ancestors, as well as the kinds of effort their contemporaries are directing at similar targets of study.

Comparative education has demonstrated all these characteristics during its development. The literature includes a wide array of subjects and approaches, symptomatic of the varied motives for studying foreign educational systems. It encompasses narrative description of single nations prompted by interest and curiosity, selective and structured observations motivated by the desire to apply lessons from abroad to the solution of educational problems at home, and encyclopedic codification of the "facts" about many countries.

Such work may be impressionistic and even normative, providing a wealth of information and insights about the nations studied. In addition, it often reveals much about the culturally determined predilections of its authors. Historical reviews of the literature in comparative education show clearly that systematic studies of foreign education increased dramatically as nations began to develop their own public school systems.

Interest in foreign educational practices has been stimulated by nationalism, the growth in international communications, and the aftermath of major wars. For some, the motive was to help develop improved education modeled on foreign practices; for others, foreign study, travel, and teaching were seen as means to ease tensions among nations and foster an international perspective. Two general questions have especially shaped the investigations of writers in comparative education: Why do educational thinking and practices differ among nations? What are the differences and similarities? The conceptual frame within which answers have been offered becomes evident from a brief review of the literature since the beginning of the 20th century, when Sadler emphasized the intimate and interactive relation of educational and historical facts.

EDUCATION SYSTEMS OF NATIONS

The education systems of nations differ because of different historical and cultural traditions, but they are similar because there are common elements in human societies. In addition, important events transcending national boundaries have influenced their affairs — the Protestant Revolution, Marxism, the Industrial Revolution, Imperialism (or the achievement of national independence from colonial control), for example. The interplay among such factors has occupied the attention of recent generations of comparative educators.

Comparative education writing has been influenced by several important perspectives. First, the realization that educational phenomena are part of the whole fabric of a nation's culture and history for the most part put an end to works that described and assessed schooling without reference to the larger cultural context of a country. Studies thereafter tended to set educational events against a historical background and to describe the genesis of different types of

schools, educational philosophies, and school systems as parts of a series of political and social events. Kandel and Ulich, especially, focused upon the links among history, national culture, political ideology, and schooling.

Much attention was given to "national identity" as the key to understanding a nation's special educational characteristics. It was conceivable that this approach would end opportunities for comparison because of the tendency to assume that each nation and its education were unique. This danger was averted, however, by the great attention given to common factors and common problems presumed to affect many countries. The major assumption that characterized comparative education work during the first half of the 20th century was that such study could illuminate the past growth and current dynamics of educational change in whole societies.

It was not until after World War II, however, that the predominantly historical, philosophical, and theoretical approaches were challenged by more pragmatic considerations. Over a hundred years before, many writers had been prompted to study the schools in foreign countries with a view to improving classroom practices and school system policies in their own lands. From about 1950, this motive prompted a renewal of interest in comparative studies, whether in highly developed nations seeking a way out of the disorder of the postwar period or in the less developed world which was confronting the problems of newly achieved independence. Educational reform and planning for national survival and growth were everywhere of paramount importance.

Economists in particular led the move towards regarding the education system of a nation as its means for investing in human potential, as its way of developing national resources. But the problems were not only economic. The survival of a nation depended as much on its success in dealing with political and social problems as on the most efficient allocation of human resources. For the first time, some nations began to regard reform in education as a possible means of achieving a sense of national unity among disparate sub-populations, a tool for ameliorating gross disparities in status and opportunity among social classes, and a mechanism for improving skills and the quality of life. To study the experience of nations other than one's own seemed pertinent. Comparative study of education and of those social, political, and economic dimensions of society closely bound up with education was encouraged by these developments.

The social sciences thus provided a leavening for the predominantly historical and philosophical approaches of earlier comparative educators. Contemporary work has developed a greater specificity of criteria and a sharper awareness of causal relations. The underlying general questions now tend to be restated more in the following form: What factors in the school system or in the social, political, or other structures of the society explain variability in pupil achievement, administrative structure, school financing, instructional methodology, and other educational phenomena? To the questions of what the similarities and differences in educational practices among nations are and what explains these similarities and differences, a third question was added: What are the outcomes of these similarities and differences?

Comparative educators have for a long time been especially concerned with the availability of data comparable across nations, with controlling the biases of observers and interpreters of data, and with integrating the data, concepts, and analytic techniques of several academic disciplines. The work of international organizations has greatly improved the availability and, to some extent, the comparability of data. Collaborative work involving persons from different nations and disciplines has strengthened the methodological grasp of the problems involved. Further, the use of statistical analysis, model theory, and systems analysis for the purposes of comparative education has been heralded by some practitioners as signifying the emergence of the field as a science.

As a result, a new kind of comparative education research has developed: the empirical cross-national study in which large amounts of data are gathered and analyzed and a variety of social science concepts and techniques are used to test hypotheses about the relations between educational variables and political, economic, and social characteristics. Interest in the methodological problems of cross-national comparison has been sharpened, and much attention has been given to the possibilities of using those strategies and tactics that had become commonplace in empirical social science research. International organizations, such as UNESCO and other specialized agencies of the United Nations, and the Organization for European Cooperation and Development (OECD) were able to collect educational and other social data systematically and on a vast scale.

These agencies have also made valuable contributions to educational planning and policy efforts, for example, the series of studies on educational finance and planning produced by the International Institute of Educational Planning (a UNESCO agency), and a set of OECD country studies in which national policy and plans have been critically reviewed and analyzed by international teams of educational experts. Technical assistance programmes have encouraged the exchange of skilled professionals among developed and underdeveloped countries, and this has made evident the commonality and immediacy of socio-educational problems in many lands.

Thus the burgeoning of data sources, increasing methodological sophistication, the meshing of social science expertise with education, and the presence of urgent problems requiring attention at the national policy level all have combined to give renewed power and variety to comparative studies in education.

An excellent example of this type of work is the massive survey undertaken by the International Association for the Evaluation of Educational Achievement (the IEA project). This project has been devoted to cross-national assessment of student achievement in selected school subjects and attempts to explain variance in such achievement. The first project was a study of mathematics achievement in 12 countries. The most recent phases of work covered science, reading comprehension, and literature in 21 countries, and reports on three more school subjects — civic knowledge and English and French as foreign languages — are in preparation.

In addition, overall studies of the six-subject surveys are in preparation. Data were not, of course, collected merely on achievement. A vast amount of information in standardized form was obtained on student home background, school practices, teacher characteristics, and the nations' school systems, as well as selected social and economic data. Results were compared at three levels: among students, among schools, and among nations. The main statistical technique used for explaining variance in achievement was multiple regression. The problems inherent in the new wave of empirical cross-national research are somewhat different from those of the earlier generation of comparativists, but they are not altogether unique.

The latter could be faulted on grounds of personal or cultural subjectivity and bias, or because their global perspective was too theoretical, or because their descriptive detail was merely interesting or idiosyncratic and not generalizable. The more empirical studies, even when thoughtfully planned and rigourously executed, are subject to such familiar methodological criticisms as representativeness, the accuracy of data, and the appropriateness of analytical design. Two general concerns, however, are more important. First, the findings should have some relevance to decision making in education (whether at the national policy level or in the school or classroom in particular pedagogical terms), and second, the subtleties of human interaction in the teaching-learning process should not be neglected by undue emphasis upon easily quantifiable and more generally conventional dimensions of education.

In reference to the large-scale survey approach of IEA, the technical problems, while important, are not insuperable. If there is enough time, experience, cooperation among experts, and money, it is possible to reduce weaknesses in sampling, data collection, analysis, and inference to reasonable levels. Relating comparative studies of this kind to policy is a more difficult task, however. At one level is the widespread problem of how to communicate scholarly research findings to practitioners; at another is the selection of problems for investigation and the search for relations among factors that bear upon important professional policy issues. The IEA study does achieve this to some extent.

For example, it illuminates the arguments over selective versus comprehensive schooling, documents and refines knowledge about the relation of sex of pupil to achievement in different subjects, and highlights the variable relations between school and home factors in accounting for pupil achievement in different subjects, at different ages, and in different countries. Furthermore, first attempts were made, notably in the literature and civic knowledge studies, to gather and to compare data on non-cognitive variables.

The trend of the past decade towards empirical, quantitative, large-scale research has not been without its severe critics. In drawing upon the quantitative techniques of economists, psychologists, sociologists, anthropologists, and political scientists, researchers run the risk of becoming distracted from those topics that are more central to educational studies: curriculum, teaching methodology, and classroom and school organization. Furthermore, it is argued,

enthusiasts for empirical methodology may ignore its limitations as an investigative strategy. Critics also tend to stress the inappropriateness of applying models of investigation drawn from the physical sciences to the humane arts, such as education.

However, few practitioners are unaware of the differences in orientation inherent in the body of comparative studies: theoretical and practical, descriptive and analytical, objective and melioristic, philosophical-historical, and empirical. Each orientation has made and continues to make its own particular contribution to the understanding of data and educational problems.

But few researchers today will deny the complementary nature of the approaches that characterize the predominantly historical studies of the 1930s and 1940s and the empirical studies of the past decade. Schooling is a mass enterprise. As such there is value in analyzing its correlates and outcomes, using techniques of mass data collection and analysis. Because education is an international enterprise that is not limited to any particular time and place, it is therefore properly studied cross-culturally. This is not to reject the view of education as a small-scale individual process in which techniques of micro-observation, analysis of small-group behaviour, and observations of classroom interaction and culture are desirable. The two approaches should properly be regarded as complementary and, as they are developed, they should contribute to better understanding of the educational process at all levels.

In education and in the social sciences at large, approaches range from attempts to identify the regularities of human behaviour in social settings to emphasis upon the special, even unique qualities of the phenomena studied. Comparative education is no exception. At one end of the scale lies a group of works intended to test particular hypotheses systematically: quantified data are statistically analyzed and inferences and predictions are made, with conventional caveats. At the other end are studies of a different nature: colourful, intuitive, eclectic, impressionistic, ranging widely. over history, philosophy, and education, spiced with social comment. Whether a particular piece of work has value is not so much a matter of where it stands on this particular range of alternative approaches as it is of how well the work has been done in its own terms.

Comparative education studies include a valuable scholarly tradition in the more humanistically oriented direction, and in recent years the field has been enriched by a growing array of works built upon empirical social science models. Viewed as complementary modes of study, both can contribute substantially to knowledge in comparative education.

6

Teaching Methods and Technological Tools

Problem Based Learning, Project-based Learning, and Enquiry-based learning are active learning educational technologies used to facilitate learning. Technology which includes physical and process applied science can be incorporated into project, problem, enquiry-based learning as they all have a similar educational philosophy. All three are student centered, ideally involving real-world scenarios in which students are actively engaged in critical thinking activities. The process that students are encouraged to employ is considered to be a technology. Classic examples of technologies used by teachers and Educational Technologists include Bloom's Taxonomy and Instructional Design.

THEORISTS

This is an area where new thinkers are coming to the forefront everyday. Many of the ideas spread from theorists, researchers, and experts through their blogs. Extensive lists of educational bloggers by area of interest are available at Steve Hargadon's "Support Bloggers" site or at the "movingforward" wiki started by Scott McLeod. Many of these blogs are recognized by their peers each year through the edublogger awards. Web 2.0 technologies have led to a huge increase in the amount of information available on this topic and the number of educators formally and informally discussing it.

Most listed below have been around for more than a decade, however, and few new thinkers are listed here:

- Alan November
- Seymour Papert

- Will Richardson
- John Sweller
- Don Krug
- Alex Jones
- George Siemens
- David Wiley
- David Wilson
- Bernard Luskin

BENEFITS

Educational technology is intended to improve education over what it would be without technology.

Some of the claimed benefits are listed below:

- Easy-to-access course materials. Instructors can post the course material or important information on a course website, which means students can study at a time and location they prefer and can obtain the study material very quickly
- Student motivation. Computer-based instruction can give instant feedback to students and explain correct answers. Moreover, a computer is patient and non-judgmental, which can give the student motivation to continue learning. James Kulik, who studies the effectiveness of computers used for instruction, students usually learn more in less time when receiving computer-based instruction and they like classes more and develop more positive attitudes towards computers in computer-based classes. The American educator, Cassandra B. Whyte, researched and reported about the importance of locus of control and successful academic performance and by the late 1980s, she wrote of how important computer usage and information technology would become in the higher education experience of the future.
- Wide participation. Learning material can be used for long distance learning and are accessible to a wider audience
- Improved student writing. It is convenient for students to edit their written work on word processors, which can, in turn, improve the quality of their writing. The students are better at critiquing and editing written work that is exchanged over a computer network with students they know
- Subjects made easier to learn. Many different types of educational software are designed and developed to help children or teenagers to learn specific subjects. Examples include pre-school software, computer simulators, and graphics software
- A structure that is more amenable to measurement and improvement of outcomes. With proper structuring it can become easier to monitor and maintain student work while also quickly gauging modifications to the instruction necessary to enhance student learning.

- Differentiated Instruction. Educational technology provides the means to focus on active student participation and to present differentiated questioning strategies. It broadens individualized instruction and promotes the development of personalized learning plans. Students are encouraged to use multimedia components and to incorporate the knowledge they gained in creative ways.

MULTI-DIMENSIONAL APPROACH TO INTEGRATED EDUCATION

In order to cope with this highly complex situation we need to examine the possibilities for expanding the concept and the scope of integrated education and try to make it multi-dimensional. Let us look at some examples.

INTEGRATION OF SCHOOL SUBJECTS

Combining and integrating school subjects will reduce the number of subjects to be included in the curriculum. For example one subject to be called "Our Environment" could be a combination of physical, political and economic geography together with environmental protection and some aspects of civics.

Another subject to be called "Our Health" could be a combination of biology, hygiene, nutrition, and some aspect of moral education.

INTEGRATION OF AT-SCHOOL EDUCATION WITH SELF-LEARNING AT HOME

Modern educational technology is gradually making it possible to learn some subjects through various "distance education" or self-learning methods. This approach if properly devised and coordinated will reduce the burden of classroom teaching.

INTEGRATING SCHOOL- EDUCATION WITH OUT-OF-SCHOOL EDUCATION

There are various opportunities at home and in the community that can provide certain knowledge and skills that are normally offered by the school. Public libraries, summer jobs, assuming some responsibilities at home, services offered by public agencies and institutions can be enlisted to cover a portion of what is now expected to be included in school curriculum. This will be a kind of integration of formal and non-formal education.

INTEGRATING EDUCATION AND LEARNING WITH RECREATIONAL AND ENTERTAINMENT ACTIVITIES

Computer games, television programmes, educational video and DVD, some sport activities, summer camps, boy and girl scouts, are but a few examples of how education can be integrated with recreation and entertainment.

INTEGRATING ACADEMIC AND VOCATIONAL EDUCATION

Dividing the education process into two separate streams by separating ordinary secondary level education from vocational training, has resulted in bringing up two groups of individuals, thus dividing the society into tow camps: in one camp are those who utilize their hands namely their physical abilities and in the other camp are those who are expected to use their heads namely intellectual capabilities. The former group is not expected to use their intellectual potential and the latter are deprived of the many advantages of manual dexterity and practical skills. To give the benefits of both types of education to each individual and bringing up more balanced and capable members of society demands integrating the two streams of education into one which provides every person with the opportunity to actualize all his or her potential aptitudes and capabilities.

In short, implementation of such changes in the delivery system of education contributes to creating a learning and educational environment at home and in the community that will help the students to be surrounded, at all times, by opportunities that will be encouraging, supplementing and strengthening their educational development.

DIFFERING GLOBAL PERSPECTIVES

Perhaps the most important step in understanding and incorporating global education in classrooms and communities is to understand and relate to the themes of global awareness as presented by experts in the field. Hanvey (1976), one of the first scholarly experts to give a comprehensive definition of the concept "global awareness", proposes five dimensions that prepare students to achieve global awareness. These include perspective consciousness, state-of-the-planet awareness, cross-cultural awareness, knowledge of global dynamics, and awareness of human choices. Haavenson, Savukova, and Mason (1998/99) conducted their research on United States and Russian perspectives on teacher education reform and global education and found that these dimensions form the first level known as attitude formation upon which global education can be implemented. The second level is the development of cognition skills and the third level is an integrated view of the world. An explanation of Hanvey's five dimensions, paraphrased by Kirkwood (2001) and Haavenson et al. (1998/99) is provided below and an explanation of the other two levels will be identified.

PERSPECTIVE CONSCIOUSNESS

Perspective consciousness refers to an awareness of and appreciation for other images of the world and that a person's worldview is neither universally shared, nor necessarily right, yet may be profoundly different. It is the realization that an individual's worldview is both a matter of conscious opinions and ideas and more importantly to subconscious evaluations, conceptions and unexamined assumptions. Perspectives are shaped by ethnic, religious, differences in age,

sex, and social status, among many other factors. These differences, as stated by Haavenson et al. (1998/99), "have been one of the main causes of conflict and confrontation in the history of mankind" (p.38).

The authors go on to say that, "It is important to teach students to look upon a certain phenomenon or event from different perspectives so as to encourage respect and appreciation for beliefs, customs, and values different from their own" (p. 38). It is not only about racial and cultural differences, instead, a pluralistic view needs to be taken when looking at global perspectives.

STATE-OF-THE-PLANET AWARENESS

State-of-the-planet awareness requires comprehension of prevailing world conditions, developments, trends, and problems that are confronting the world community. It includes an in-depth understanding of global issues such as population growth, migrations, economic disparities, depletion of resources, and international conflicts, that require global learners to be aware of the world around them. Children need to be made aware that what affects the world affects them as well. In elementary school, students can be taught to make decisions about ways to prevent disaster by studying the consequences of environmental illiteracy.

CROSS-CULTURAL AWARENESS

This dimension includes the diversity of ideas and practices in human societies and how these ideas and practices are found in human societies around the world, including concepts of how others might view one's own society as perceived from other vantage points. According to Hanvey (1976), this dimension is the most difficult to attain most likely because it refers to the highest level of global cognition. The misconception about cross-cultural awareness is that people consider it no more than a set of stereotypes that do more harm than good as superficial knowledge engenders prejudice. An effective way to promote cross-cultural awareness, as explained by Haavenson et al. (1998/99), is by showing videos and then having discussions with students about these films to help them in separating stereotypical views from those that are more authentic.

KNOWLEDGE OF GLOBAL DYNAMICS

Knowledge of global dynamics refers to an understanding of the world as an interconnected system of complex traits and mechanisms and unanticipated consequences. A high level of sophistication on the part of the student is required because understanding these processes is difficult to achieve due to the unanticipated effects on the human condition. It includes a consciousness of global change and cannot be acquired through mass media. Haavenson et al. (1998/99) explain that "[s]tudents learn to identify subtle cause-effect relationships, anticipate side effects, model processes and make decisions about eliminating or altering undesirable consequences". Students may be asked to create webs of the factors influencing the issue, to suggest feasible solutions, and to foresee possible side effects of such actions.

AWARENESS OF HUMAN CHOICES

Hanvey (1976) challenges global thinkers to realize the problems of choice confronting individuals and nations as consciousness and knowledge of global systems expand. It is related to global dynamics in such a way that it focuses on making choices and develops a sense of responsibility for making decisions made which affect future generations. It also includes an awareness of the interconnectedness of individual, national, and international settings. It fosters a sense of responsible citizenship on the local and global levels. Students may be introduced to alternatives on thought and behaviour by looking at relationships and interactions between man and the world. Students are asked to account for their choices and are taught to be tolerant towards the view of others.

In their study with teachers, Haavenson et al. (1998/99) posits that the second level of global education implementation is cognition focused. This means that life demands both a thorough knowledge of a domain combined with a broad perspective of the world. This is similar to the 'interconnections' theme that Werner & Case (1997) identify and develop which explores both the international and inter-system linkages and conclude that we live in an interconnected world. Therefore students must be encouraged to see the different ways in which one situation is influenced by and influences others. Further exploration in the topic is explained by Haavenson et al (1998/99) that the brain often searches for common patterns and relationships and seeks to connect new knowledge with prior experiences that result in the fact that cognition operates in all concepts. The traditional approach of filling the minds with facts and information that students are simply asked to memorize and reproduce does nothing to promote global awareness and teachers must keep this in mind when working to plan curriculum. Instead, students need experience in critical thinking, in taking part in cross-cultural experiences, and to make decisions and substantiate them. In the study by Haavenson et al. (1998/99), students are taught to think for themselves and to be able to stand their ground. The authors advise that the atmosphere created by the teacher is very important.

The third level of global education implementation is an integrated view of the world as explained by Haavenson et. al. (1998/99). They state that "the third level aims to create a specific picture of the world where geographical, physical and linguistic features all fit into a complex pattern". This means that all discipline-focused world perspectives need to overlap due to the interdependence of facts, events, and phenomena. For instance, university interdisciplinary courses may be the most effective way to create a cross-disciplinary perspective.

GLOBAL AWARENESS ELEMENTS

Case (1993) identifies five key substantive elements that keep people informed of a range of global topics. The first element describes the universal values and cultural practices, and the second includes global interconnections, which refers to the study of the workings of the four major interactive global

systems: economic, political, ecological, and technological. The third presents worldwide concerns and conditions such as development and peace issues while the fourth forms the origins and past patterns of worldwide affairs such as global history and geography. The last presents alternative future directions in worldwide affairs. In addition to these substantive elements, he proposes perceptual elements that should be addressed which include open mindedness, resistance to stereotyping, anticipation of complexity, empathy, and nonchauvinism.

Kirkwood (2001) analyses Case's elements and explains that the substantive elements listed above "includes the objects of global education that incorporate the contemporary events, conditions and locations in the world that Hanvey (1976) addresses within the state of the planet awareness dimension". The perceptual elements focus on the development of world mindedness and empathy and resistance to prejudicial thinking as well as stereotyping and cross-cultural knowledge. These elements are similar to the Hanvey dimensions of perspective consciousness and cross-cultural awareness.

Case (1993) and Hanvey (1976) provide similar definitions for global awareness even though the terminology they use is different. Merry M. Merryfield, one of the leading scholars in the field of global education, combines the definitions of other scholars and provides us with a current framework in this field today. Kirkwood (2001) lists Merryfield's eight elements which include: human beliefs and values global systems, global issues and problems, cross-cultural understanding, awareness of human choices, global history, acquisition of indigenous knowledge, and development of analytical, evaluative, and participatory skills. Kirkwood (2001) concludes that "Merryfield's work contributes significantly in reducing, if not eliminating, the definitional ambiguities that still linger in the field".

NEW UNDERSTANDINGS OF GLOBAL AWARENESS

Kirkwood (2001) presents another dimension to the definition of global education that Lamy (1987) identifies as the acquisition of knowledge transmitted by indigenous people. He concludes that a global education must include knowledge about the contributions of native people who are representing the views of their world. In his words, "The teaching of historical and contemporary events must be balanced by listening to indigenous voices".

To provide further elaboration in regards to listening to Indigenous voices, Battiste and Henderson (2000) think of globalization as a new threatening transformation that is emerging. In the introduction of their book, they state that, "Globalization with its cognitive and linguistic imperialism is the modern force that is taking our heritages, knowledge, and creativity". For Indigenous people it is not just physical survival that concerns them; it is an issue of "maintaining Indigenous worldviews, languages, and environments". It is ironic that the world looks to Indigenous people for help in order to solve the world's crisis that its worldview has created. Battiste and Henderson (2000) state that

"in view of the history of relations between the colonizers and the colonized, this is an extraordinarily bold request." The work of David Selby and Graham Pike has brought new understandings in regards to ecological awareness or 'state of the planet awareness' as outlined by Hanvey. They have been influential in the global education field in the 1980's and are known nation wide. Influenced mainly by Richardson and Hanvey in the 1970's they have worked mainly with secondary schools. Hicks (2003) in his review of global education discusses the work of Selby & Pike. He explains that in 1988 they further developed the conceptual map of the field and highlighted what they called 'the four dimensions of globality'. These dimensions make up the core elements of global education. The first is 'issues dimension', which embraces five major problem areas and solutions to them: inequality/equality, injustice/justice, conflict/peace; environmental damage/care; alienation/participation. The second is 'spatial dimension' which emphasizes exploration of the local-global connections that exist in relation to these issues, including the nature of both interdependency and dependency. The third is 'temporal dimension' that emphasizes exploration of the interconnections that exist between past, present, and future in relation to such issues and in particular scenarios of preferred futures. The fourth is the 'process dimension' that emphasizes a participatory and experiential pedagogy which explores differing value perspectives and leads to politically aware local-global citizenship. Selby and Pike then relate this to both individual subjects in the curriculum and whole-school case studies.

Hicks (2003) further explains that each of these four elements needs to be present before one can claim to be involved in global education. Both Selby & Pike have written extensively on the importance of ecological thinking in global education and this is evident within the four-dimensional model that they propose for global education. It needs to be stressed that the environmental health of the world is just as important as taking care of all humanity and that the two must work together simultaneously. For example when explaining the 'spatial dimension' Selby (1998) writes that "this dimension also concerns the cycles and systems of nature and the relationships between human society and the environment".

Hicks (2003) explains that the 'temporal dimension' is a futures perspective that "looks at how global issues affect and are affected by interrelationships between past, present and future". He goes on to say that "this works to help young people think more critically and creatively about the future, especially in relation to creating more just and sustainable futures".

MAKING SENSE OF VARYING DEFINITIONS AND PERSPECTIVES

Given the comprehensive and overlapping definitions of global awareness and perspectives as they relate to global education it is evident that similar views are presented. The following reaction will analyze the strengths and weaknesses of the interpretations and what implications they have for global

education. Hanvey's (1976) description of global awareness doesn't include the relationship between perspective consciousness and the power one holds either locally or globally. According to Merryfield and Subedi (2001), this makes a significant difference.

They explain this relationship by explaining how "the development of perspective consciousness differs considerably depending upon the degree to which students perceive that people like themselves are on the margins or in the center of their society".

DuBois (1989) wrote of double consciousness as a coping response to racism: "Black children grew up conscious not only of their own culture learned from family and community but also the white culture that designated them an inferior race, a problem to be solved". Merryfield (2000) states that "DuBois' conceptualization of double consciousness helps to explain why people who are placed on the margins develop the ability to perceive multiple realities by looking at events and issues both through the perspectives of people in the mainstream and people on the margins".

Conflicts and misunderstandings that occur in present day classrooms seem to be an indication of this. If this double consciousness does not develop in white people due to their race-based dominant position and if the majority of teacher educators are middle class, white, more male than female; then it is difficult to expect that the ways of looking at the world in classrooms is looked at from multiple perspectives.

Although these teachers might interact with people different from them it is always from a privileged position. Merryfield and Subedi (2001) also state that "white people because of their race-based dominant position did not develop double consciousness".

Battiste and Henderson (2000) further elaborate on this perspective consciousness by saying that Indigenous students, "experience what scholar W.E.B Du Bois, in The Souls of Black Folk referred to as "double consciousness". They give a further explanation of double consciousness that states, "Double consciousness occurs when the dominators reject the assertions of the colonized that they are human and insist on imposing the standards of the colonizers as universal and normal. According to the above explanation it can be concluded that a global perspective does not mean a universal perspective. A universalism exists even in colonialism and imperialism. This has privileged a few and alienated many.

Colonization makes a difference in the way that Indigenous people view and accept globalization. Battiste (2000) states, "Indigenous scholars are now struggling to define Indigenous humanity. First they need to understand the systems of thought that gave rise to this alienation, and then they need to create a shared language both sides can use to discuss education, science, epic storytelling of huge devastation, painful struggle and persistent survival". Indigenous students need to begin to analyze knowledge and information and to be given the opportunities to participate in meaningful dialogue and

interactions. Wilson (2000) in a paper on lessons from Ghana further articulates this double consciousness by saying, "we may be conscious of our own perspective but often are not aware how strongly our nationality, our culture, and our experiences inform that perspective". Hanvey (1976), Case (1993), Merryfield (2000), among others, mention this and I agree that it is the first step in developing multiple perspectives. It is not enough to look at the world through one's own sunglasses, we must experience the world by taking another's sunglasses and looking through them in order to begin making sense of other worldviews.

In the same article, Wilson (2000) explains that along with fellow teachers he gains knowledge about the world from similar kinds of sources that include, course work in school, media, interaction with people from different countries, and people met while traveling. In the same way we can conclude that cross-cultural experiences, watching videos, reading online newspapers from other countries and organizing these activities around the expectation of free expression of ideas, respect for differing viewpoints, active participation and a desire to communicate are all effective ways in creating a global perspective. Textbooks and media sources used in order to retrieve information are not enough; rather meaningful experiences and reading and discussing books written by writers of differing cultures are more effective.

Edward Said, a Christian Palestinian who grew up in Egypt, makes sense of how perspectives are informed by defining the meaning of the East versus the West. He came to the United States and studied literary theory and came to understand that Americans had more myths than theories about the Middle East. He forms another perspective in the topic of global education by writing about embedded cultural understandings. Said provides a good description of how scholarly misinformation of different cultures came about and captures this idea in the following quote paraphrased by Merryfield (2003):

In his seminal work "Orientalism" Said demonstrates how European explorers, intellectuals, missionaries, settlers, travel writers, and others created scholarly misinformation because they relied on their own cultural frames of reference to describe, catalog, and interpret the cultures of Arabs, Muslims, Asians, Africans, and others. Five hundred years of this orientalist scholarship served both political goals and cultural beliefs in that it clearly differentiated Europeans as superior to other peoples and affirmed the European right to rule and "civilize" Africans, Asians, Arabs, Indians, *etc.*

Merryfield (2003) explains that Said's writings are useful to teachers in that they help explain the thinking and pedagogy of exemplary global educators who challenge "colonialisms" in the social studies. His work helps teachers to understand and see how the legacy of imperialism shapes mainstream academic knowledge through its framework of "us" and "them". This refers to "us" (the white men who created the dominant power and represent its ideals and "them" (the others who are divided from "us" by their inferior cultures, poverty, politics, language or other differences)". Social studies curriculum must go beyond

European or American constructions of knowledge and must teach experiences, knowledge, and perspectives of diverse peoples in Africa, Asia, Latin America and the Middle East.

Selby (1998) states that global education is nothing less than the educational expression of an ecological holistic or systemic paradigm and as such has implications for the nature, purposes and processes of learning and for every aspect of the functioning of a school or other learning community. Selby and Pike's research has led us to believe that "global education is a holistic paradigm that encompasses the interconnectedness of communities, lands and peoples and interrelatedness of all social, cultural and natural phenomena".

TEACHING SOCIAL STUDIES

The first step towards global awareness is the shaping of attitudes. There needs to be a shift from the traditional Eurocentric way of looking at the curriculum and incorporating more emphasis on critical thinking and decision-making skills. White (2002) in his article states that "Students who can think for themselves and look critically at societal problems will find their classroom a more exciting and challenging place. Engaging students in learning through dialogue enhances their journey to knowledge and competency". Focusing, understanding, and applying the Hanvey (1976) dimensions combined with substantive and perceptual elements by Case (1983) are an important first step in creating a global awareness in children. However, understandings must include the concept of double consciousness and other ways of knowing that Du Bois and other indigenous writers such as Battiste, Henderson and Hampton discuss.

White (2002) discusses the need for social studies to change in reaction to the constantly changing world. The subject of social studies has been ethnocentric and intent on socialization, instead of focusing on critical citizenship. He proposes that the most effective way to engage children in learning about global perspectives is that they view the world as a planet-wide society and understand the interdependence of human beings". As mentioned previously, teachers will need to develop a global knowledge, to experience cross-cultural experiences and to expand their own perspectives of the world in order to help students achieve the goals.

In order to view the world as a planet wide society, it is important to look at the work of Selby and Pike. They stress a four-dimensional model of global education as core elements of the programme and relate this to both the individual subjects in the curriculum and whole-school case studies.

One aim that Pike and Selby (1988) have pointed out is 'health of the planet awareness'. The aim states that "students should acquire an awareness and understanding of the global condition and of global developments and trends and to develop a future orientation in their reflection upon the health of the planet". Children need to possess an awareness and understanding of the global condition and of global developments and trends in order to become active in making sound choices and effective decisions at a variety of levels.

A challenge exists for society to make a definite change in the area of global education. Hanvey's definition of global awareness does not reflect this and it appears to be politically neutral. The elements that he identifies send a clear message that the world needs to make changes in order for the planet as a whole to sustain itself.

However, it does not really challenge people to making a definite change. It builds a foundation needed for society to change attitudes but does not indicate any radical ways to accomplish this. In order to begin to make changes definite plans need to be implemented. Personal, family, classroom and community goals need to be followed through and at the same time it is important to keep in mind that plans need to fit into one's world-view. The Eurocentric way of action planning does not work for everyone and perhaps there are alternative ways that communities can work together to achieve certain goals. Hanvey's intent was only to create awareness and to instigate people into thinking in a global way and is a great start.

The media is another area where caution needs to be exercised in the area of forming perspectives. Media influences Eurocentric thinking and instead of depending on what the media presents students can find other sources of information and compare understandings. Other ways to get information is through cross-cultural interactions, presentations by guest speakers, videos, biographies and documentaries and books among many others. Students need to be informed with current up to date material and then given a chance to reflect on this information. Therefore, it is important that there be a reflection component to encourage further critical and decision making thinking. Reflection could take place by using discussion boards, e-mails, chats, journals, and any other way that students feel comfortable. These are all effective ways to gain substantive knowledge about the world and its systems and perceptual understandings especially if a follow up with constructive feedback is given. This method validates children's knowledge and strengthens their confidence in decision-making and in self-esteem. Teachers must educate themselves first in local, global and national knowledge systems of the world and continue learning about global issues together with their students.

Merryfield (2001) describes the need for decolonizing of the mind to take place in order to incorporate global education. This makes sense because how can one truly understand the notion of globalization if one's culture is looked down upon and seen as inferior to mainstream culture and where European history remains at the center of world history. The Kenyan playwright and scholar Ngugi wa Thiong refers to the concept of decolonizing the mind in the article written by Merryfield and Subedi (2001). As Thiong states, "a colonial mentality deeply permeates many Kenyan's thinking today because it is not only embedded but unexamined". Aboriginal people in Canada experience this phenomenon as well. They are not conscious of how oppressors force their worldview on their lives and therefore other societal problems such as physical/mental abuse, poverty, and addictions that seem to be separated from colonization become apparent. A tremendous amount of energy that goes into dealing with societal

problems often neglects to look at imperial colonization. Merryfield (2001) explains this by stating, "later generations people may never realize that their ideas and choices are affected by colonialist or neo-colonialist perspectives".

In order to teach global perspectives as mentioned throughout this paper there is a need to decolonize the Social Studies. Merryfield (2001) discusses a strategy that teachers can use to accomplish this. Global educators use contrapuntal or opposing histories and literature to describe how they challenge the Eurocentric selection of historical events. To teach multiple perspectives or alternative histories instead of using a single universal history is crucial for students to critically examine and question their own historical understandings. This needs to be accomplished by current information and accurate content that includes the knowledge, voices and ideas of people from these regions. Haavenson et al. (1998/99) go on to explain that it is important to exercise the implementation in classrooms of the dimensions of perspective consciousness through the selection of updated, globally relevant content.

It is also important to keep in mind that it is not only the oppressed that form a colonized mind. Many young people are acculturated into thinking that white is superior. This gives a false sense of security that perpetuates across generations. McLaren (1995) explains that the white culture needs to be interrogated as well. He states, "unless we give white students a sense of their own identity as an emergent identity - we naturalize whiteness as a cultural marker against which otherness is defined".

Smith (2000) states that the self identity of Western civilization for the last three hundred years is a myth and that a "far more profound truth may be that there is no Self without Others, no Me without You. It is important to stay clear from the embedded ways of thinking of the "us" versus "them" attitude.

TECHNOLOGY IN THE CLASSROOM

There are various types of technologies currently used in traditional classrooms.

Among these are:

- *Computer in the classroom*: Having a computer in the classroom is an asset to any teacher. With a computer in the classroom, teachers are able to demonstrate a new lesson, present new material, show how to use new programmes, and show new websites.
- *Class website*: An easy way to display your student's work is to create a web page designed for your class. Once a web page is designed, teachers can post homework assignments, student work, famous quotes, trivia games, and so much more. In today's society, children know how to use the computer and navigate their way through a website, so why not give them one where they can be a published author. Just be careful as most districts maintain strong policies to manage official websites for a school or classroom. Also, most school districts provide teacher webpages that can easily be viewed through the school district's website.

- *Class blogs and wikis*: There are a variety of Web 2.0 tools that are currently being implemented in the classroom. Blogs allow for students to maintain a running dialogue, such as a journal,thoughts, ideas, and assignments that also provide for student comment and reflection. Wikis are more group focused to allow multiple members of the group to edit a single document and create a truly collaborative and carefully edited finished product.
- *Wireless classroom microphones*: Noisy classrooms are a daily occurrence, and with the help of microphones, students are able to hear their teachers more clearly. Children learn better when they hear the teacher clearly. The benefit for teachers is that they no longer lose their voices at the end of the day.
- *Mobile devices*: Mobile devices such as clickers or smartphone can be used to enhance the experience in the classroom by providing the possibility for professors to get feedback.
- *Interactive Whiteboards*: An interactive whiteboard that provides touch control of computer applications. These enhance the experience in the classroom by showing anything that can be on a computer screen. This not only aids in visual learning, but it is interactive so the students can draw, write, or manipulate images on the interactive whiteboard.
- *Online media*: Streamed video websites can be utilized to enhance a classroom lesson
- *Digital Games*: The field of educational games and serious games has been growing significantly over the last few years. The digital games are being provided as tools for the classroom and have a lot of positive feedback including higher motivation for students.

There are many other tools being utilized depending on the local school board and funds available.

These may include: digital cameras, video cameras, interactive whiteboard tools, document cameras, or LCD projectors:

- *Podcasts*: Podcasting is a relatively new invention that allows anybody to publish files to the Internet where individuals can subscribe and receive new files from people by a subscription. The primary benefit of podcasting for educators is quite simple. It enables teachers to reach students through a medium that is both "cool" and a part of their daily lives. For a technology that only requires a computer, microphone and internet connection, podcasting has the capacity of advancing a student's education beyond the classroom. When students listen to the podcasts of other students as well as their own, they can quickly demonstrate their capacities to identify and define "quality." This can be a great tool for learning and developing literacy inside and outside the classroom. Podcasting can help sharpen students' vocabulary, writing, editing, public speaking, and presentation skills. Students will also learn skills that will be valuable in the working world, such as communication, time management, and problem-solving.

Although podcasts are a new phenomenon in classrooms, especially on college campuses, studies have shown the differences in effectiveness between a live lecture versus podcast are minor in terms of the education of the student.

CHALLENGES FOR EDUCATIONAL TECHNOLOGY

Many exciting applications of information technology in schools validate that new technology-based models of teaching and learning have the power to dramatically improve educational outcomes. As a result, many people are asking how to scale-up the scattered, successful "islands of innovation" instructional technology has empowered into universal improvements in schooling enabled by major shifts in standard educational practices. Undertaking "systemic reform" requires policies and practices different than fostering pilot projects for small-scale educational improvement. Systemic reform involves moving from utilizing special, external resources to reconfiguring existing budgets in order to free up money for innovation. Without undercutting their power, change strategies effective when pioneered by leaders in educational innovation must be modified to be implemented by typical educators.

Technology-based innovations offer special challenges and opportunities in this scalingup process. I believe that systemic reform is not possible without utilizing the full power of high performance computing and communications to enhance the reshaping of schools. Yet the cost of technology, its rapid evolution, and the special knowledge and skills required of its users pose substantial barriers to effective utilization. One way to frame these issues is to pose six questions that school boards, taxpayers, educators, business groups, politicians, and parents are asking about implementing large-scale, technology-based educational innovations. After each question, I'll respond to the issues it raises. Collectively, these answers outline a strategy for scaling-up, leveraging the power of technology while minimizing its intrinsic challenges.

SCHOOLS AFFORD TO PURCHASE ENOUGH MULTIMEDIA

Giving all students continuous access to multimedia-capable, Internet-connected computers is currently quite fashionable. For politicians, the Internet in every classroom has become the modern equivalent of the promised "chicken in every pot." Communities urge everyone to provide volunteer support for NetDays that wire the schools. Information technology vendors are offering special programmes to encourage massive educational purchases. States are setting aside substantial amounts of money for building information infrastructures dedicated to instructional usage.

Yet, as an educational technologist, I am more dismayed than delighted. Some of my nervousness about this initiative comes from the "First Generation" thinking about information technology that underlies these visions. Multimedia-capable, Internet-connected computers are seen by many as magical devices, "silver bullets" to solve the problems of schools. Teachers and administrators

who use new media are assumed to be automatically more effective than those who do not. Classroom computers are envisioned as a technology comparable to fire: just by sitting near these devices, students get a benefit from them, as knowledge and skills radiate from the monitors into their minds.

Yet decades of experience with technological innovations based on First Generation thinking have demonstrated that this viewpoint is misguided. Classroom computers that are acquired as panaceas end up as doorstops. As discussed later, information technology is a costeffective investment only in the context of systemic reform. Unless other simultaneous innovations in pedagogy, curriculum, assessment, and school organization are coupled to the usage of instructional technology, the time and effort expended on implementing these devices produces few improvements in educational outcomes—and reinforces many educators' cynicism about fads based on magical machines.

I feel additional concern about attempts to supply every student with continuous access to high performance computing and communications because of the likely cost of this massive investment. Depending on the assumptions made about the technological capabilities involved, estimates of the financial resources needed for such an information infrastructure vary. Extrapolating the most detailed cost model to one multimedia-capable, Internet-connected computer for every two to three students yields a price tag of about ninety-four billion dollars of initial investment and twenty-eight billion dollars per year in ongoing costs, a financial commitment that would drain schools of all discretionary funding for at least a decade.

For several reasons, this is an impractical approach for improving education. First, putting this money into computers-and-cables is too large an investment in just one part of the infrastructure improvements that many schools desperately need. Buildings are falling apart, furnishings are dilapidated, playgrounds need repair, asbestos must be removed...otherwise, the machines themselves will cease to function as their context deteriorates.

Also, substantial funding is needed for other types of innovations required to make instructional hardware effective, such as standards-based curricular materials for the WorldWide Web and alternative kinds of pedagogy based on partnerships between teachers and tools. If most of the money goes into new media, little funding is available for the new messages and meanings that those devices could empower.

Second, without substantial and extended professional development in the innovative models of teaching and learning that instructional technology makes affordable and sustainable, many educators will not use these devices to their full potential.

"Second Generation" thinking in educational technology does not see computers as magic, but does make the mistake of focusing on automation as their fundamental purpose. Computers are envisioned as ways to empower "teaching by telling" and "learning by listening," serving as a fire hose to spray information from the Internet into learners' minds.

However, even without educational technology, classrooms are already drowning in data, and an overcrowded curriculum puts students and teachers on the brink of intellectual indigestion. Adding additional information, even when coated with multimedia bells-and-whistles, is likely to worsen rather than improve educational settings. Professional development needs are more complex than increasing educators' technical literacy. The issue is building teachers' knowledge and skills in alternative types of pedagogy and content, and such an increase in human capabilities requires substantial funding that will be unavailable if almost all resources are put into hardware.

Third, the continuing costs of maintaining and upgrading a massive infusion of school based technology would be prohibitive. High performance computing and communications requires high tech skills to keep operational and will become obsolete in five to seven years as information technology continues its rapid advance. Yet taxpayers now see computers as similar to blackboards: buy them once, and they are inexpensively in place for the lifetime of the school. School boards rapidly become restive at sizable yearly expenditures for technology maintenance and telecommunications usage—especially if, several months after installation, standardized test scores have not yet dramatically risen—and will become apoplectic if another $50B to replace obsolete equipment is required only a few years after an initial huge expenditure. For all these reasons, investing a huge sum in information infrastructures for schools is impractical and invites a later backlash against educational technology as yet another failed fad.

I would go farther, however, and argue that we should not make such an investment even if the "technology fairy" were to leave $100B under our virtual pillows, no strings attached. Kids continuously working on machines with teachers wandering around coaching the confused is the wrong model for the classroom of the future; I wince when I see those types of vendor commercials.

In that situation—just as in classrooms with no technology—too much instructional activity tends to center on presentation and motivation, building a foundation of ideas and skills as well as some context for why students should care. Yet this temporary interest and readiness to master curricular material rapidly fades when no time is left for reflection and application, as teachers and students move on to the next required topic in the overcrowded curriculum, desperately trying to meet all the standards and prepare for the test. Substantial research documents that helping students make sense out of something they have assimilated, but do not yet understand is crucial for inducing learning that is retained and generalized. Reflective discussion of shared experiences from multiple perspectives is essential in learners' converting information into knowledge, as well as in students mastering the collaborative creation of meaning and purpose.

Some of these interpretative and expressive activities are enhanced by educational devices, but many are best conducted via face-to-face interaction,

without the intervening filter and mask of computer-mediated communication.

What if instead much of the presentation and motivation that is foundational for learning occurred outside of classroom settings, via information technologies part of home and workplace and community contexts? Students would arrive at school already imbued with some background and motivation, ripe for guided enquiry, ready for interpretation and collaborative construction of knowledge. People are spending lots of money on devices purchased for entertainment and information services: televisions, videotape players, computers, Web TV, videogames. Many of these technologies are astonishingly powerful and inexpensive; for example, the Nintendo 64 machine available now for a couple hundred dollars is the equivalent of a several hundred thousand dollar graphics supercomputer a decade ago. What if these devices—many ubiquitous in rich and poor homes, urban and rural areas—were also utilized for educational purposes, even though not acquired for that reason? By off-loading from classroom settings some of the burden of presenting material and inducing motivation, learning activities that use the technology infrastructure outside of schools would reduce the amount of money needed for adequate levels of classroom-based technology. Such a strategy also enables teachers to focus on students' interpretation and expressive articulation without feeling obligated to use technology in every step of the process.

Such a model of "distributed learning" involves orchestrating educational activities among classrooms, workplaces, homes, and community settings. This pedagogical strategy models for students that learning is integral to all aspects of life—not just schooling— and that people adept at learning are fluent in using many types of information tools scattered throughout our everyday context. Such an educational approach also can build partnerships for learning between teachers and families; this is important because parental involvement is certainly one of the most powerful levers in increasing any student's educational performance.

In other words, unless "systemic reform" in education is conducted with one boundary of the system around the school and another boundary around the society, its affordability and sustainability are doubtful.

As a bridge across these boundaries, new media can play a vital role in facilitating this bi-level approach to large-scale educational innovation. For example, videogame players are the only interactive devices widely available in poor households and provide a sophisticated, but inexpensive computational platform for learning—if we develop better content than the mindless follies of Super Mario™ or the grim dystopias of Doom™. My research in virtual reality shows how multisensory, immersive virtual environments could leverage learning complex scientific concepts on computational platforms as commonplace as next decade's videogames.

Districts can leverage their scarce resources for innovation, as well as implement more effective educational models, by utilizing information devices outside of classrooms to create learning environments that complement computers and communications in schools. To instead saturate schools with

information technology is both very expensive and less educationally effective.

SCHOOLS AFFORD ENOUGH COMPUTERS AND TELECOMMUNICATIONS TO SUSTAIN NEW MODELS

Educational improvement based on distributed learning—utilizing information technologies external to school settings to enable increased interpretive and expressive activities in classrooms—does not mean that schools won't need substantial amounts of computers and communications. To empower project-based learning through guided enquiry, students must have access to sophisticated information devices in schools. Even if this is accomplished via notebook computers and wireless networks moved from class to class as required, with pupils also spending significant amounts of time learning without the aid of technology, districts must allocate more money to purchasing, maintaining, and upgrading computers and telecommunications than has been true historically.

Where will educators find the funds for equipment, software, technical staff, ongoing telecommunications services, professional development—the myriad of costs associated with a sophisticated information infrastructure? In the past, this money has come largely from special external sources: grants, community donations, bond initiatives. To be sustainable over the long run, however, resources for technology must come from reallocating existing budgets by reducing other types of expenditures. Of course, such shifts in financing are resisted by those groups whose resources are cut, and district admini-strators and school boards have been reluctant to take on the political challenges of changing how money is spent. An easy way to kill educational innovations is to declare that of course they will be implemented—as long as no existing activities must be curtailed to fund new approaches. Such an approach to institutional evolution is one reason why, if Rip Van Winkle awoke today, he would recognize almost nothing in modern society—except schools.

Educational organizations are unique, however, in demanding that technology implementation accomplished via add-on funding. Every other type of societal institution recognizes that the power of information devices stems in part from their ability to reconfigure employee roles and organizational functioning. These establishments use the power of technology to alter their standard practices, so that the cost of computers and communications is funded by improvements in effectiveness within the organization, by doing more with less. If educators were to adopt this model—reallocating existing resources to fund technology implementation—what types of expenditures would drop so that existing funds could cover the costs of computers and communications?

First, schools that have adopted the enquiry-based models of pedagogy find that outlays on textbooks and other types of standardized instructional materials decrease. While these materials are a smaller part of districts' budgets than salaries or physical plants, nonetheless they cost a significant amount of money. When students collect their own data, draw down information across the Internet, and interact with a larger pool of experts than teachers and textbooks, fewer

commercial presentational resources are required—especially if learners draw on topical data flowing through information sources outside of schools. Moreover, covering a few concepts in depth rather than surveying many ideas superficially reduces the amount of prepackaged information educators must purchase.

A second way to reconfigure existing financial resources is to reduce the staff involved in data entry operations. Educators are inundated with large amounts of recordkeeping functions, and one of the most debilitating aspects of this work is the continuous reentry of identical information on different forms. Businesses have saved substantial amounts of money by altering routine information processes so that data is only entered once, then automatically flows across the entire organization to each place in which it is needed. Were educators to adopt these already proven models for cost-efficient information management, the amount of time and staff required for data entry functions would decrease markedly, freeing funding for instruction-related uses of technology.

Third, and on a more fundamental level, teaching is more efficient and effective with new types of technology-based curriculum and pedagogy. At present, substantial re-teaching of knowledge and skills is required; presentational material flows into students' minds, is retained just long enough to perform on a test, and then is forgotten. Class sizes are typically between twenty-five and forty—somewhat too large for effective project-based learning, yet small given that lectures work as well for several hundred students as for several dozen.

The scheduling of class periods is too short, limiting teachers and students to fragmentary presentational and practice activities. Teachers all have comparable roles with similar pay structures—unlike other societal organizations, which have complementary staff roles with a mix of skill levels and salaries. Visions presented in the forthcoming 1998 ASCD Yearbook depict how altered configurations of human resources, instructional modalities, and organizational structures could result in greater effectiveness for comparable costs—even with the acquisition of substantial school-based technology. This case is also made at greater length in Hunter and Goldberg.

In the commercial sector, too often these types of institutional shifts result in layoffs. However, because of the coming wave of retirements among educators, districts have a window of opportunity to accomplish structural changes without major adverse impacts on employees. Over the next decade, large numbers of "baby-boom" educators will leave the profession, and a staged process of organizational restructuring could occur in parallel with those retirements. Coordinating technology expenditures as an integral part of that larger framework for institutional evolution is vital in districts' planning to afford computers and communications.

7

Importance of Research Methodology

STUDY IMPLICATIONS

The purpose of research is to inform action. Thus, your study should seek to contextualize its findings within the larger body of research. Research must always be of high quality in order to produce knowledge that is applicable outside of the research setting. Furthermore, the results of your study may have implications for policy and future project implementation.

One problem that often plagues progress in global health is the slow translation of research into practice. Often, a disconnect exists between those who conduct research and those who are positioned to implement the research findings. The underlying problem is that "the production of evidence is organized institutionally with highly centralized mechanisms, whereas the application of that science is highly decentralized.

This social distance prevails because scientists are more oriented to the international audiences of other scientists for which they publish than to the needs of practitioners, policy makers, or the local public."Thus, as researchers, it is imperative to take steps to overcome this barrier. Publishing your study may be one initial step to make your research known to the global community. Other proactive measures can be taken to encourage the uptake of evidence-based interventions. For example, you can present your research findings at various venues, such as Unite for Sight's Global Health and Innovation Conference. Furthermore, you can send the results of your study to local officials, policy-makers, and community leaders.

Goals of Research

There are relatively few published studies about eye care in developing countries, and Unite For Sight encourages all volunteers to consider developing a research study to contribute important knowledge to the eye care community on a global scale. Pursuing a research project will be a challenging and rewarding experience, and this opportunity enables you to pursue an in-depth original study about a topic of interest.

Well-conducted research is vital to the success of global heath endeavors. Not only does research form the foundation of programme development and policies all over the world, but it can also be translated into effective global health programmes. Research draws its power from the fact that it is empirical: rather than merely theorizing about what *might* be effective or what *could* work, researchers go out into the field and design studies that give policymakers hard data on which they can base their decisions. Furthermore, good research utilizes methodologies that can be replicated, produces results that are examinable by peers, and creates knowledge that can be applied to real-world situations. Researchers work as a team to enhance our knowledge of how to best address the world's problems.

The "Iterative" Process of Research

Ultimately, the key to a successful research project lies in iteration: the process of returning again and again to the research questions, methods, and data, which leads to new ideas, revisions and improvements. It is easy to think of research as a step-by-step "1, 2, 3" process, but it is important to be flexible and open to change.

Oftentimes, by discussing the research project with advisers and peers, one will find that new research questions need to be added, variables need to be omitted, and other changes need to be made. As a proposed study is examined and reexamined from different perspectives, it may begin to transform and take a different shape.

This is to be expected and is a component of a good research study. In addition, it is important to examine study methods and data from different viewpoints to ensure a comprehensive approach to the research question. In conclusion, there is no one formula for developing a successful study, but it is important to realize that the research process is cyclical and iterative.

RESEARCH METHODS V/S METHODOLOGY

METHOD

A method is simply the tool used to answer your research questions—how, in short, you will go about collecting your data. Examples of UX research methods include:

- Contextual inquiry
- Interview

- Usability study
- Survey
- Diary study
- Card sort

If you are choosing among these, you might say "what method should I use?" and settle on one or more methods to answer your research question.

Methodology

A methodology is the rationale for the research approach, and the lens through which the analysis occurs.

Said another way, a methodology describes the "general research strategy that outlines the way in which research is to be undertaken" (An Introduction to the Philosophy of Methodology, Howell 2013). The methodology should impact which method(s) for a research endeavor are selected in order to generate the compelling data.

Examples of methodologies, courtesy of Elin Bjorling, include:

- *Phenomenology:* Describes the "lived experience" of a particular phenomenon
- *Ethnography:* Explores the social world or culture, shared beliefs and behaviours
- *Participatory:* Views the participants as active researchers
- *Ethnomethodology:* Examines how people use dialogue and body language to construct a world view
- *Grounding theory*:* Assumes a blank slate and uses an inductive approach to develop a new theory

**Despite the fact that* grounding theory *has theory in its name, don't let that fool you—it is actually a methodology because it aims to generate theory from systematic application of research.*

If you wanted to know about the lived experiences purchasing food in the United States, for instance, you would be using the phenomenology methodology— and from there you could choose from different methods to collect that data.

For instance, you might perform a contextual inquiry and shop alongside participants; you might also interview a handful of participants and ask them to recount their most recent grocery shopping experience; you might equally choose to do a survey and ask the same questions to hundreds of participants.

Because the contextual inquiry gets the researcher much closer to the actual setting, the results may be considered stronger and more transferable in the future.

TYPES OF QUESTIONS AND CORRESPONDING METHODS

The insight that the method should follow from the question, now pretty obvious and something graduate students often hear from their professors, led

to another conundrum. There are so many research questions, how do we say anything cogent about method? The Committee reasoned that most scientific research questions were of three general types: (1) What's happening? (2) Is there a systematic (causal) effect? And (3) what is the causal mechanism or how does it work?

What's Happening?

The question—what's happening?—asks for a description. We could ask this question in a materials science laboratory or in a middle school. We could describe the years of experience possessed by elementary school teachers in the U.S., or the types of science instruction students receive, or the changes in students' mathematics achievement over 20 years. In these cases, statistical estimates of population parameters could be obtained from available national surveys. Or we could describe what school, home and community look like though the eyes of an inner-city youth using ethnographic methods. Or we could describe different approaches to the assessment of learning in colleges and universities that have been nominated as "exemplary" using case study methods.

Holland and Eisenhart's (1990) study exemplifies scientific research into what's happening, and, as we shall see, beyond. They were concerned about explanations for why so few women pursued careers in non-traditional majors such as science: women were not well prepared before coming to college, women were discriminated against in college, women did not want to compete with men for jobs.

They began their study by *describing, in depth over several years* through *ethnography*, the lives of 23 volunteer women from two small public residential colleges. Half of these women pursued traditional careers and the other half non-traditional careers. They were matched on grades, college majors, college activities, and college peers. Based on extensive, detailed data collected through participant observation and interviews, Holland and Eisenhart found that what best described these women's academic pursuits were, *contrary to popular conjecture*, how they viewed the value of schoolwork, what their reasons were for doing school work, and how they perceived financial and opportunity costs. Simply put, detailed description of the college lives of these women portrayed their career trajectories in a very different light than the distal conjectures based on statistical data.

Now if you are thinking, "well, this is just idiosyncratic description and is suspect," Holland and Eisenhart one up you. They took the next step, going beyond description and entering the arena of model testing by predicting with their descriptive models what each of the 23 women would do after college: continue in school, get a job in her field, get a job outside her field, get married, *etc.* At the end of 4 years and another 3 years, they followed these women up with telephone interviews. In *all* 23 cases, predictions based on their model of commitment to schoolwork were confirmed. In all cases, their model provided a better predictor than did data on precollege preparation (grades and courses taken), discrimination against women, or feelings about competing with men.

Is There a Systematic (Causal) Effect?

Questions about effects are, ultimately, questions about causal effects. Did *x* cause y? Perhaps the most widely known study of systematic effects in education is the Tennessee randomized trial on class size reduction. The question posed by the Tennessee legislature was whether reduced class size would improve students' achievement. To this end, within 79 schools across the state, a total of 11,600 students were randomly assigned to a regular class (22-26 students), a class with an aide (to decrease the student/adult ratio economically) or to reduced class size (under 13-17 students). Three findings stand out: (1) students in small classes outperformed students in the other classes, (2) minority students particularly benefited, and (3) the effect persisted when students returned to regular class sizes. Here the method, randomized trial, followed directly from the legislature's question and was feasible and ethical to implement. In such cases as this, randomized trials are the preferred method for ferreting out causal effects.

There are many cases, however, when randomization may not be feasible. Such cases include the effect of smoking on health and longevity, and the effects of hunger, alcohol use, drug use or child abuse on students' academic performance. For these research questions, other methods need to be used, are available, and include quasi-experiments (control and experimental groups without random assignment), correlational studies using large-scale probability-sampled data sets (that adjust for selectivity bias), and various time series designs. To be sure, as you move away from the randomization in some cases uncertainties increase; nevertheless, causal interpretations are possible and replication is important to increase confidence in the interpretations.

Application of Correlational

Loeb and Page's (2000) study of teacher salaries exemplifies the application of correlational (structural) modeling in a situation where random assignment is unlikely. They asked, "If teacher quality affects student achievement, why do studies that predict student outcomes from teacher wages produce weak results?" That is, shouldn't teachers' salaries reflect, at least to some degree, teacher quality after controlling for other things?

Loeb and Page tested two competing models. One was the usual production function model that links inputs (salary) to student outcomes (dropouts in this case) after controlling for relevant variables. The second model followed their reasoning that there are other things in the lives of teachers than salary that may have meaning, and there also may be local job markets that provide attractive alternatives to teaching in the area. So their second, competing model incorporated opportunity costs into the production function: non-pecuniary rewards and competition in the local job market. They replicated prior research with the usual production-function model, showing a weak effect of salaries on outcomes. However, once they adjusted this model for opportunity costs (non-pecuniary and job market incentives), they found that raising wages by 10 per cent reduced high school dropout rates by 3-4 per cent.

Three points about studying causal effects seem appropriate here. First, in dealing with causal assertions we are always trying to rule out all the possible counter hypotheses that we know of at the time. As a research programme moves along, new challenges (counter hypotheses) arise and get ruled out; in this way confidence increases in the causal interpretation. Oftentimes we don't know all the counter hypotheses; challenges arise with novel counter-interpretations, and research and debate continues as it has with the Tennessee study. This type of debate—hypothesis/counter-hypothesis—is the basis of science and should be looked upon positively and not as "backbiting" among scholars with different views when the issue is one of interpretation; it is backbiting when personal attacks are made.

A second point has to do with the role of description in causal studies—what's happening? When feasible, descriptive research should be used in causal studies to help us understand, as fully as possible, what "treatments" were actually implemented, and to reveal what possible causal mechanisms might be operating.

And the third point is that establishing a causal effect may be necessary when possible but not sufficient in policy and practice. The questions of mechanism and context inevitably should arise in order to design education policies or practices. We need to understand how interventions were articulated and implemented in diverse contexts with whom, under what conditions with what resources in order to design more than superficial education policy.

Causal Mechanism

The third type of research question focuses on the mechanism that creates a causal effect. For example, reducing class size seems to have a salutary effect according to the Tennessee study. But what was the mechanism that caused the effect and why did it persist even after students returned to regular class sizes Was the effect due to an increase in the number and personal nature of teacher-student contacts or to less off-task student behaviour or to the level of student engagement?

Empirical Studies of Mechanism

Empirical studies of mechanism, following on studies that have established causal effects are most common. Bryk, Lee and Holland (1993) sought to understand the causal mechanism(s) underlying the causal evidence that Catholic schools outperform public schools in the U.S., This longitudinal study used both qualitative (*e.g.*, case studies of effective Catholic schools) and quantitative data to address the mechanism question. Three potentially explanatory models were tested: (1) sector effects only (spiritual and private characteristics of Catholic schools), (2) compositional effects (kinds of students attending Catholic schools), and (3) school effects (school operations contributing to school context). A combination of models, characterizing "…the *coherence* of school life in Catholic schools … most clearly accounts for it relative success in this area".

Nevertheless, there is another way to approach the question of mechanism—namely, to build an artifact based on a causal theory and establish its causal effect. Studies such as "design experiments" or "design studies" posit a theory with a causal mechanism, and design educational artifacts (*e.g.*, a curriculum, a computer application) and iteratively test them out in complex real-world classroom environments, revising both artifact and theory along the way. Once evidence accumulates to suggest a causal mechanism, the onus, of course, is on design researchers to then establish generalizable causal effects.

TEACHER MANAGEMENT AND TRAINING

Historically, state governments have not had total control over the management of teachers in publicly-supported schools. While state schools were large in number, a sizeable number of schools were under the management of District Boards, Taluka Boards, Municipal Corporations, and so on. In the early 1970s, many state governments took direct control of these schools, creating a mega system of primary schooling. This arrangement began to undergo transformation in mid-1990s, following a Constitutional Amendment which called for the empowerment of panchayati raj bodies. Some, but not all, state governments took this opportunity to wholly or partially decentralize the teacher management system. Over time, the unevenness in service conditions of teachers within and between states has increased. In some states, local self-governing bodies are given a reasonably free hand to decide on teacher recruitment at a local level, and operate within a broad framework provided by the state government. While all teachers are technically state government employees, therefore, some teachers are employed by local village level bodies, whilst others are employed by block or district level bodies.

Another significant development during the 1990s, which has the potential to significantly influence the professional development of teachers, is the creation of a system of district and sub-district level teacher resource centres. Following the recommendations of the National Policy on Education 1986, every district in the country was provided with a district-specific teacher education institution, a District Institute of Education and Training. Subsequently, under the District Primary Education Programme, block resource centres and 10-15 cluster resource centres were established in each of the 42 districts which participated in the project. This is now being universalized to reach all districts of the country under SSA.

Availability and Deployment of Teachers

Mere availability of schools does not ensure children's participation. More than schools, motivated teachers play a vital role in ensuring that children attend school regularly and take an active part in the learning process. Though there has been an impressive increase in the number of school teachers in India in the 1990s, the imbalances in intra-state teacher allocation between districts and within districts, and between rural and urban areas continues to be a major

concern. An increase in the numbers of upper primary teachers has taken place in rural areas in Andhra Pradesh, Gujarat, Himachal Pradesh, Karnataka, Madhya Pradesh, and Rajasthan indicating improved completion of the lower primary cycle and increased transition to upper primary classes.

Indeed, in most of states, the percentage increase in upper primary teachers has been more than that of the primary teachers. This indicates a positive trend in transition rates from lower to upper primary classes. However, the allocation of teachers particularly between rural and urban areas continues to be a major concern in many states. In almost all states, trained teachers with higher qualifications are generally concentrated in urban areas.

Several factors, including political and bureaucratic interventions, non-availability of qualified local teachers in remote areas, lack of a proper database on the status of teachers at the school level and the low motivation of urban teachers to serve in rural areas, are obstacles to the rational deployment of teachers. Moreover, teacher transfer is highly influenced by political interventions, and more influential teachers are more likely to be found working in urban schools and schools located in habitations well connected with road or rail transport networks.

The lack of basic facilities also act as de-motivating factors, particularly for female teachers. These factors undoubtedly influence pupil-teacher ratios with the average PTR higher in rural areas than in urban areas.

RECRUITMENT OF PARA-TEACHERS

In the 1990s, a number of states began appointing para-teachers and other types of temporary teachers, mainly at the primary level. This was justified as necessary to meet the additional teacher requirements caused by increased enrolments. However, in some states parateachers have been appointed to fill vacancies left by regular full-time teachers. Available data shows that the relative share of primary and upper-primary parateachers is high in Gujarat, Andhra Pradesh, Himachal Pradesh, Uttar Pradesh, Uttarakhand and Maharashtra. The trend for appointing para-teachers continues in rural primary and upper primary schools in some states. In Gujarat, for example, more than 43% of rural primary teachers and 31% of rural upper primary teachers are temporary teachers. The Seventh All India Educational Survey, there were 103,270 para-teachers at the primary level, 104,894 para-teachers at the upper primary level, 37,950 at the secondary level, and 33,911 para-teachers at the higher secondary level in 2002. A study conducted under the District Primary Education Programme in 1998, covering 21 para-teacher schools over five states, highlighted that there were distinct advantages if the teacher was a local resident, including better community interactions, teacher punctuality and efficiency of schools.

While acknowledging some short-term advantages of the system, Pandey reiterates that, by accepting the scheme of para-teachers, the central government is encouraging states to evade their responsibility for building a strong cadre of qualified teachers. The trend has also challenged teachers' professional identities.

Para-teacher schemes may therefore serve the immediate purpose of universalisation of access to elementary education in far flung rural and hilly areas, but replacing regular teachers with para-teachers is in general detrimental to the quality of education and the effectiveness of schools, and should be avoided if possible. Large-scale non-formal education and alternative schools with para-teachers are often concerned with physical access over quality of provisions.

Meanwhile, investment in improving the capacity of teachers and organizing continuous educational resource support and pedagogic renewal has taken a backseat. The biggest problem with the concept and rationale behind these schemes, as pointed out by the PROBE Report, is that the Shiksha Karmi approach involves the potential risk of institutionalized dualism in the schooling system. This issue may not be so prevalent when Shiksha Karmis are posted as auxiliary teachers in mainstream schools and are spread evenly through the schooling system. However, Shiksha Karmis tend to be posted in areas where regular teachers are reluctant to go, such as tribal or backward villages. In these deprived areas, Shiksha Karmis tend to be used as low-cost substitutes for non-performing teachers, often managing the entire school on their own. This policy is often supported by regular teachers who, as a result, are more likely to be posted to more attractive areas. This results in deprived children being taught by poorly-qualified, low paid Shiksha Karmis, while those from more privileged families are more likely to be taught by a fully qualified teacher. It is worth noting, however, that regular school teachers are not necessarily better than contract teachers.

Problems relating to classroom processes, pedagogic techniques, classroom management and other constraints which effect para-teachers also impact on the quality of education offered by trained and qualifed teachers. Ramachandran et al in their study on teacher motivation, for example, found that trained teachers indicated several reasons for dissatisfaction, including high teacher-pupil ratios, inadequate infrastructure, erratic disbursement of salaries, being 'forced'to teach children of poor communities and specific social groups who are 'dirty', irregular attendance of children, and illiterate parents, all of which add to teacher workloads.

Tilak also points out that most state governments favour para-teacher schemes and EGS schools as they save huge resources and avoid teacher managerial problems. So while the central government can claim to have gone a long way towards fulfilling the constitutional directive on universal elementary education, little attention has been given to the likely effects on quality of education or the long term implications of the schemes.

As Govinda and Josephine argue:

- The schools, which are targeted for employment of contract teachers, are those where children from the poorer sections of the society study. Thus it would exacerbate inequity in the society by creating classes of government schools with different kinds of teachers for different classes of population.

To avoid such a situation, Govinda and Josephine suggest the creation and nurturing of a professional cadre of teachers governed by well-designed norms and standards as agreed by the state and teachers. A transparent process of recruitment, posting and promotion of teachers is therefore urgently required. Even though the appointment of contract teachers has come in for considerable questioning in public discourses, there are no systematic studies to analyse the long-term impact of such recruitments on the quality of schools on the one hand, and on the development of a professional cadre of teachers, on the other.

This suggests that a focus needs to be placed on the larger implications of policies related to the professional morale of teachers and the manner in which parateacher schemes may jeopardize efforts to build a strong cadre of professionally trained elementary school teachers and the impacts of provision of a second-grade alternative education for children from already deprived and disadvantaged sections of society.

INADEQUATE QUALITY AND RELEVANCE

While there are exceptions, the quality and relevance of research, teaching and learning has continued to decline in public tertiary education institutions. Many universities operate with overcrowded and deteriorating physical facilities, limited and obsolete library resources, insufficient equipment and instructional materials, outdated curricula, unqualified teaching staff, poorly prepared secondary students and an absence of academic rigour and systematic evaluation of performance. Similar conditions can be found in many of the new private universities and other tertiary institutions which have cropped up in many countries, especially in those countries which do not have any formal system of licensing or accreditation of new institutions. Even in the former socialist nations of Eastern Europe and Central Asia, drastic reductions in public funding are further jeopardizing the quality and susceptibility of existing programmes and even the survival of entire institutions. In many countries, the poor quality of teacher training programmes has detrimental effects on the quality of learning in primary and secondary education. Weak secondary education and scientific literacy, in turn, does not provide high school graduates with the necessary skills for successful tertiary level studies.

Most tertiary education institutions function at the periphery of the international scientific community, unable to participate in the knowledge production and adaptation necessary to confront the most important economic and social problems of the countries. While very few countries have exhaustive data to document the depth of the problem in a systematic manner, when information is available, the situation is quite alarming, as illustrated by figures from the Philippines. In 1995, a Task Force on Higher Education concluded, after reviewing information on critical education inputs and the results of professional examinations for the 1316 existing tertiary education institutions, that only nine universities and two colleges were comparable in quality to international institutions.

In both public and private institutions, the lack of full-time qualified teachers is an important factor of poor quality. In Latin America, for example, less than 6 percent of professors teaching in public universities have a doctoral degree and 26 percent a master's degree. More than sixty percent of teachers are part time in the public sector; in the private universities, the proportion is as high as 86 percent (Garcia Guadilla, 1998). Expansion and diversification of tertiary education systems has often led to internal brain drain because low paid professors in public institutions seek second and third jobs in extramural jobs such as teaching at private institutes and colleges which might offer better salaries.

While there is a global trend towards increased returns to tertiary education, specific market responses are not always favourable. For example when countries expand just any kind of tertiary education in order to meet the increasing social demand, there is a high risk of graduate unemployment. In many countries, the mismatch between the profile of graduates and labour market demands is mostly apparent among graduates in social sciences and humanities. Tertiary education institutions often lack adequate information for prospective students, parents and employers. Problems of quality and relevance are not exclusive to traditional universities. Even in countries which have diversified their tertiary education structure such as Indonesia, relevance can become a serious issue in the absence of close linkages between tertiary education institutions and the labour market.

Lack of access to the global knowledge pool and the international academic environment is increasingly becoming an issue. In many countries, poor command of foreign languages among staff and students complicates access to textbooks and the Internet, especially at the graduate level. Many countries which have experienced a doubling or tripling of tertiary enrolments in the last few decades, along with increased participation rates for young people, have also seen the negative effects of rapid expansion on quality. As a result, issues of quality assurance and quality enhancement have been a major focus of attention (El-Khawas et al., 1999). Despite differences in the size and stage of development of their tertiary education sectors, many governments have decided that traditional academic controls are inadequate to deal with today's challenges and that more explicit quality assurance systems are needed. There are wide differences among countries in their approaches to quality promotion. In some countries, governments have taken steps to strengthen quality by introducing new reporting requirements or other mechanisms of management control. Argentina, for example, has introduced quality assurance mechanisms that depend on an enhanced information and evaluation system and new rules for funding the universities.

Many countries have developed accreditation systems, while others have established evaluation committees or centres that carry out cycles of external review. In many countries, independent bodies have been established, often a single national agency but sometimes, as in the Netherlands, Mexico, or Romania, separate agencies that are responsible for different types of institutions, regions,

or purposes. Such variation in approach reflects political and cultural preferences within each country, differences in governmental leadership, as well as varying stages of development for the tertiary education sector. The scope of responsibilities given to quality assurance systems has also varied widely. Scotland and England, for example, have procedures to monitor teaching effectiveness, while Hong Kong is focusing on high-quality management processes. Some countries, Chile for one, have established systems to license new institutions and to certify education credentials.

Others have directed efforts towards rewarding research productivity, either of individual scholars (as in Mexico) or of entire academic departments (as in the United Kingdom). There is wide variation too in the extent to which quality assurance agencies have addressed issues related to student transfer and to study in other countries, as well as the issues related to the expansion of new modes of educational delivery, including video-based education, interactive transmission to remote sites or, most recently, Internet-based learning. The most comprehensive activity in relation to transnational credit transfer and student mobility is undertaken among European countries under the *Bologna Declaration*.

The further development of the ECTS (European Credit Transfer System) progresses rapidly and it is highly recommended that other nations which have build their tertiary education system on the European tradition are kept informed about the rapid development of such a large international marketplace for students. In many countries, the governance structure and management traditions of public tertiary institutions are characterized by rigidities and a total lack of flexibility which inhibits any type of reform or innovation. In the name of academic freedom, institutions (and their individual constituents, faculty, administrators and students) frequently operate with limited accountability for their use of public resources or for the quality of their outputs (*e.g.*, graduates, research). In grained institutional cultures, together with poor management practices and lack of accountability, explain some of the inefficiency dimensions identified earlier. The time-honoured committee approach to management in universities suffers from lengthy, sometimes politically-laden, consensus-based decision making. It often lacks the agility for effective interaction with a surrounding corporate culture.

The ownership of tertiary institutions have often shifted from clients, *e.g.*, society and students, to staff. The *reason d'eter* for some institutions has become providing employment and benefits for staff rather than being educational establishments geared towards the needs of the students. Such systems are rigorously guarded by cadres of academic leaders represented in academic councils who operate within a framework of institutional autonomy that is almost exclusively accountable to staff and academics. Academic leaders such as proctors, deans and heads of departments are not trained in management of large complex institutions.

In many public universities in Latin America and Eastern Europe, reform-oriented rectors stand little chances of getting elected because they are perceived

as a threat to established practices. When there is a change of proctor, the entire management team is changed with the ensuing loss of institutional continuity. Often the institutional support systems do not provide guidance in terms of monitoring and evaluation of the institutions' own performance. Few institutions have a governance structure allowing for the participation of representatives of employers and civil society. Universities in countries as diverse as Russia, Bangladesh and Bolivia have no Boards of Trustees that would constitute an explicit external accountability channel. Reliance on performance indicators as management and planning tools is not a common practice in most countries.

At the national level a stalemate often exists between academically powerful rectors conferences or councils and governments that continue to negotiate line item budgets seldom linked to institutional performance or national strategies, but generally reflect the needs of regional constituencies. This leads to a political rather than a professional system of management and governance. The consequence is a deficient governance system lacking flexibility and innovative capacity because programmes are developed to serve the needs of existing staff rather than the country's development goals and lack of programmatic accountability because academic autonomy is not paired with financial and legal responsibility.

The situation is often compounded by cumbersome administrative rules and bureaucratic procedures. In many countries, the Ministry of Education determines staffing policy, budgetary allocations, number of students admitted and universities have little say in relation to the number of positions, the level of salaries and promotions. In Brazil, the *Law of Isonomy* establishes uniform salaries for all federal jobs including those in the federal universities. Lengthy procedures at the level of the ministries of finance and education often cause delays in transfer of funds to tertiary education institutions. The purchase of laboratory equipment is also affected by such inefficiencies; by the time the equipment arrives it is often less up to date than originally intended and institutions end up receiving equipment supplies after the courses have taken place. In many countries and institutions, administrative procedures are also very rigid when it comes to making changes in academic structure, programmes or mode of operation.

In Uruguay, for instance, it is only when confronted in the mid-1990s with competition from new private universities that the venerable University of the Republic—which for 150 years had exercised a monopoly over higher education in the country—started a strategic planning process and considered establishing post- graduate programmes for the first time. Another example of institutional inflexibility occurred in Venezuela, where a dynamic private business administration institute called IESA had to wait several years to receive the official approval from the Council of Rectors for a new MBA programme designed and delivered jointly with the Harvard Business School.

Institutions for tertiary education are in general not designed to deal with civil constituents. Students can block entire systems from functioning over prolonged periods of time, as happened in 1999 in Mexico where UNAM the

largest university of the country was closed down for almost a year by a student strike in response to a proposed tuition fees increase from a few dollars to 140 dollars a year. Some countries have seen an alarming increase in campus violence which can be politically motivated (Colombia) or the result of criminal activities (Bangladesh) and which severely damages the functioning of the institutions. Cheating, which has also become widespread in many settings, is another dimension of inadequate operation. Finally, student democracy often works against the academic interests of the very students. Extended campaigning and election periods detract from teaching and learning and lead to inefficiencies rather than to better opportunities for the students.

THE PERSISTING INEQUALITIES

Even though rapid enrolment growth has increased access to tertiary education for traditionally less privileged groups, including women and students of rural origin, tertiary education, especially the university sector, generally remains elitist, with the majority of students coming from wealthier segments of society. In many countries, the substantial subsidies for non-educational expenditures represent regressive social spending in so much as the proportion of university students from high and middle income families is higher than their share in the overall population. This is as socially inequitable as it is inefficient. Finally, there is an equity backlash whenever fees are imposed only on certain groups, as happens in a number of former socialist economies where the students who do not pass the official entrance examination can still be admitted into regular university programmes, provided they are able to pay tuition fees. Although few countries and institutions collect data on the socio-economic origin of students in a systematic way, where statistics and household survey data are available the pattern is clear. In Latin America, for example, the proportion of students from the bottom third income group is only 6, 11 and 18 percent in Peru, Chile and Uruguay, respectively (García-Guadilla, 1998).

To a large extent, inequalities in access to tertiary education, including gender disparities, are determined by what happens to various groups in primary and secondary education. But access is not the only determinant of equity at the tertiary level. Recent household survey data from Argentina illustrate in a powerful way how even open access tertiary education systems can be deceiving from an equity perspective. Despite the appearance of democratic access for all secondary education graduates, academic outcomes are strongly influenced by socio-economic origin. The proportion of students from the poorest two quintiles who actually graduate from public universities is only a fifth of those who enter as first year students under the open access policy. There are only six countries which have introduced or raised fees risk to experience an increase in access disparities in the absence of effective and well targeted student aid mechanisms. Even in a wealthy country like Scotland, the concurrent establishment of tuition fees and elimination of the maintenance grant in 1998 has resulted in a noticeable decline in enrolment among low income students.

The World Bank in Support of Tertiary Education

In continued pursuit of its mandate to help developing and transition countries reduce poverty and improve living standards through sustainable growth and investment in people, the World Bank has renewed and deepened, its commitment to enhancing the contribution of advanced education to economic and social development worldwide. Through effective partnerships with other multilateral institutions, national governments, non-governmental organizations and the private sector, the Bank aspires to apply its financial resources and extensive knowledge base towards increased efforts in the tertiary education and science and technology sectors, which will help create the foundations of democratic knowledge-based economies and societies. Since it began to lend for the education sector in 1963, the World Bank has had a prominent role in assisting countries in their efforts to expand tertiary education and improve the quality of institutions and programmes. From 1992 to 1998, lending for tertiary education has averaged US$ 481 million a year. The Bank is currently implementing tertiary education projects or projects with a tertiary education component in 28 countries, supporting the following types of interventions:

- Vision development, strategic planning and consensus building (national level, institutional level).
- Financing reforms (allocation of recurrent budget, competitive funding, cost sharing, student loans, scholarships).
- Governance and management reforms (policy body, mergers/federation, adoption of academic credit system, MIS).
- Quality improvement (strengthening of existing programmes, evaluation/accreditation system, innovation in programme content/ delivery, innovation in academic organization, information and communication infrastructure).
- Institutional diversification (establishment and/or strengthening of polytechnics or technical institutes).
- Science and technology development (strategy development, capacity for monitoring and evaluation, reform of resource allocation mechanisms, competitive funding, promotion of research in priority areas, joint public/private sector technology development, capacity for metrology, standards, testing quality, intellectual property rights).

In the 1970s and 1980s, much of the support provided in World Bank tertiary education projects was organized in a piecemeal fashion, with a narrow focus on the establishment of new programmes or discrete quality improvement measures for existing teaching and research activities. These projects occasionally created well-equipped academic oases that became unsustainable over time but the Bank was rarely able to offer the long term comprehensive support for tertiary education that is required for successful reform efforts and effective institution building. An internal review of implementation experience with tertiary education projects undertaken in 1992 and an assessment of recent interventions in this sub-sector have offered insights into more productive ways

of supporting tertiary education reforms and innovations. The most salient lessons about the relative effectiveness of different approaches in support of tertiary education reform and development can be analysed along the following three dimensions:

- Comprehensiveness of the intervention strategy.
- Attention to the political economy aspects of reform.
- Reliance on positive incentives to promote change.

System-Wide Approaches

The degree of comprehensiveness of the Bank's support strategy is an important predictor of outcome. Policy measures and investments which are not integrated into a broad reform programme based on a global vision and strategy for change are less likely to bear their fruits. For example, implementation of an ongoing project in Argentina has been quite successful because it has accompanied a well articulated reform programme sanctioned by a new Higher Education Law promoting the introduction of internal and external evaluation mechanisms, including a national accreditation system, increased autonomy for the public universities in the area of human and financial resources management, support for quality improvements throughout the university system and institutional strengthening of the Ministry of Higher Education and the public universities, as well as a new funding formula. Even when it comes to technical aspects of quality improvement reforms, there is a need for a comprehensive approach reflecting the inter-relatedness of academic programmes and tertiary education institutions.

High quality instruction in engineering, medicine, agriculture and in applied social sciences, for example, requires sound training in the natural sciences, mathematics and even in humanities whose importance to economic development is less obvious and, therefore, less attractive in terms of donors support. Advanced scientific training and research requires strong undergraduate programmes and a large diversified tertiary education system so that undergraduate and postgraduate do not compete for scarce staff and financial resources. Centres of excellence cannot be maintained if they must bear the burden of accommodating most of the increasing social demand for tertiary education as well. The various institutional components of the tertiary education sub-sector, both public and private, constitute a system. How they relate to each other and to the system as a whole needs to be taken into account.

Reforms of the financing of public tertiary education, especially introduction of tuition and other fees, are difficult to successfully implement without expanding educational opportunity through equity measures. They also require significant devolution of government control in matters affecting institutional costs as well as incentives for institutions to engage in cost-saving and income generating activities. Student loan schemes may work well technically but at the same time fail to promote improved efficiency and cost effectiveness of tertiary education. In Venezuela, for example, the Bank supported the reform of

the public student loan agency, FUNDAYACUCHO, with a project aiming at increasing coverage, improving the financial susceptibility of the agency and raising its management efficiency.

Although the operation was a great success from the viewpoint of disbursements, its real impact was very limited because the project was not part of a comprehensive reform of tertiary education financing. By contrast, a similar operation in Jamaica has had a more positive impact because the reform of the Student Loan Bureau has supported parallel efforts to improve the financial situation of the University of West Indies through increased cost-sharing. The preference for comprehensiveness does not mean, however, that all aspects of a reform can or should be packed into a single operation. This is where the time dimension plays a crucial role in the implementation of a system-wide approach. Sequencing provide the tools for responding and adjusting to evolving challenges and long term involvement through a series of complementary operations has proven essential to ensure structural change in a viable and sustainable manner. Investments were most successful in countries like China, Korea, Indonesia and Tunisia where a series of projects have supported a sector-wide strategy.

A long term approach has greatly enhanced the probability of genuine and sustainable reforms. The sustained cooperation with the Indonesian government has led to a real change of paradigm in the way the tertiary education system is governed, financed and developing. Similarly, two successive and complementary sector-wide projects in Tunisia have contributed to substantial improvements in the tertiary education system.

Political Economy Aspects

Until the beginning of the 1990s, very little attention was paid to the political economy aspects of tertiary education reforms. The Bank worked under the assumption that it was enough to design a technically sound reform programme and reach agreement with top government officials to be able to introduce change successfully. But when it came to actual implementation, political reality often proved stronger than the technocratic vision.

In Sub-Sahara Africa, for instance, a number of education adjustment loans in the late 1980's and early 1990's included tertiary education reform measures aiming at containing expenditures, enrolment growth and subsidies. But implementation experience has not been encouraging. The proposed reform programmes, often including too many conditionalities that never materialized, have been opposed by various interest groups and have touched off student rioting in many countries.

Launching and implementing tertiary education reforms and innovations has been more successful when decision-makers have managed to build a consensus among the various constituents of the tertiary education community. This has been the case in recent tertiary education sector activities with Argentina, Chile, Brazil and other countries such as Jordan, Tunisia and Romania.. Furthermore, policy dialogue, stakeholder consultation and consensus building are not discrete

activities which are useful only at the beginning of a reform. Rather, there is a need to maintain and renew attention to the political dimensions of tertiary education reform as country conditions evolve. Failure to do so can expose the project to diminishing commitment or even reversal of policies as a result of electoral deadlines and changes of governments or key leaders. The Bank has been less than successful in supporting the implementation of politically sensitive reforms such as moving from negotiated budgets to formula funding, reducing subsidies or introducing tuition fees. In several countries, for example Argentina, Tunisia and Jordan, the government has not been able to fulfil its commitment to implement a transparent funding formula, even though it had been agreed at the time of project preparation. In Senegal, the authorities have recently gone back on their decision to streamline the scholarships programme in such a way as to ensure that only socially and academically deserving students would be eligible. In Hungary, the government reversed a decision to charge tuition fees for repeating students.

Reliance on Incentives

The third set of lessons is linked to the nature of the policy instruments used to promote reforms and innovations. The extent to which projects rely on positive incentives rather than mandatory edicts to stimulate change has a great influence on their outcomes, as institutions and actors tend to respond more readily to constructive stimuli. In that respect, the World Bank has had positive experience with a number of policy instruments, including competitive funds, accreditation mechanisms and management information systems.

Well designed competitive funds greatly stimulate the performance of tertiary education institutions and can be powerful vehicles for transformation and innovation. In Argentina, the Quality Improvement Fund (FOMEC) has encouraged universities to engage in strategic planning for the strengthening of existing programmes and the creation of new interdisciplinary masters. Within universities, faculties that had never worked together have started to cooperate in the design and implementation of joint projects. In Indonesia, a series of projects since 1993 have been successful in stimulating ownership for the new paradigms in tertiary education among the entire academic community. In Egypt, the Engineering Education Fund was instrumental in introducing, for the first time, the notion of competitive bidding and peer evaluation for the allocation of public investment resources. The Fund promoted in an effective manner the transformation of traditional engineering degrees into more applied programmes with close linkages to industry.

A fundamental prerequisite for the effective operation of competitive funds, as well as one of their significant benefits, is the establishment of transparency and fair play through clear criteria and procedures and an independent monitoring committee. In countries with a relatively small or isolated academic community, it is desirable to draw from a regional or international pool of peer reviewers to reduce opportunities for complacency and subjective evaluation among a limited

group of national colleagues, following a long standing practice in the Netherlands for example. The new Competitive Fund in Jordan has detailed guidelines described in an Operations Manual and relies on international peer reviewers for projects of national interest.

In Chile a second wave of tertiary education reform is being supported by a competitive fund for diversification (development of the non-university sector) and quality improvement (of all tertiary institutions). Venezuela, Mexico and Brazil are encouraging the formation of advanced human capital in science and technology through competitive funding mechanisms. In all of these cases, international peer reviewing experts figure prominently. In some cases, there may be a compelling argument to open several windows or to set up compensatory mechanisms to create a level playing field between strong and weak institutions. In Indonesia, three different windows were designed to serve universities accordingly to their actual institutional capacity.

In the latest tertiary education project in China, the top universities are required to form a partnership with a university in a poor province to be able to compete. In Egypt, the Competitive Fund in the Engineering Education Reform project had a special window for technical assistance to help the less experienced schools of engineering prepare well- formulated proposals. Also, proposals which included a partnership agreement between a stronger university and a weaker one received additional points for evaluation purposes. In Chile, a special window has recently been opened to provide preparation funds for universities requiring assistance in strategic planning and sub-project formulation.

8

Exploration in Educational Research

The education literature, of the prevalence of students' misinterpretations of key statistical concepts leads me to question how statistical thinking is embedded in academic settings. That is, how do students experience statistical reasoning in the academic setting and how does this reasoning relate to that of working statisticians in practical situations? According to one statistician the components of statistical thinking are: understanding the dynamics of the real world problem, moving towards a statistical model and using statistical tools.

If students in a university statistics course are not provided with experiences enabling them to understand the relationship between the statistical model and the "real life" situation, they may not be adequately equipped to use the statistical tools with which they are presented and so may be unable to appreciate the statistical reasoning process. For example a lack of understanding about variation could lead students either to trust implicitly their own intuition about data, without attempting to use statistical reasoning to evaluate the findings more critically, or to distrust universally any statistical statement–without the means to assess its reliability. It is important for me in this investigation to understand how students' concepts of statistical knowledge relate to the contexts in which they apply this knowledge. The ways that students experience learning statistics are also likely to contribute to the difficulties that many students have with the subject and to their objections to studying it. How students understand statistics is related to how they interpret it in the context in which it is presented.

This is an aspect of students' learning of statistics which has received scant attention in the research literature and, in my experience, is rarely accorded importance in the teaching of statistics courses in higher education. Skills and

facts are easily identified in the design and assessment of statistics courses at university. However, students' conceptions of the subject matter as a whole, approaches to learning it and their perceptions and beliefs are rarely evaluated in any formal sense. These aspects of students' learning of Statistics are major topics for exploration in my project as it is a prime postulate of my research that these are critical dimensions of students' emerging statistical knowledge.

In learning Statistics, a student's thinking and problem solving is accompanied by affective elements: feelings, beliefs, desires and attitudes which will affect how long she will persist with a task or a problem, how easily distracted she is likely to be, how long she will remember facts and skills and other important factors which can facilitate or hinder her cognition. Affect is therefore of importance and interest in my investigation which seeks to understand learning Statistics from the activity theory perspective. In this framework, affective elements are not extraneous or secondary to intellectual processes but indivisible from them.

Vygotsky describes this inseparability of intellectual processes and affective elements as follows:

- Thought itself is engendered by motivation, *i.e.*, by our desires and needs, our interests and emotions. Behind every thought there is an affective-volitional tendency, which holds the answer to the last 'why' in the analysis of thinking. A true and full understanding of another's thought is possible only when we understand its affective-volitional basis.

Hence, the affective components of students' cognitive processes give those processes their "attitudinal colour". Statistics is included in the curricula of many disciplines at university in order that students may use it and appreciate its relevance to their chosen fields. In Western countries, statistical thinking is bound up with technology, commercial concerns and other matters of cultural importance. The social interests and preoccupations of a culture are not, however, automatically assimilated by individuals, but are monitored.

Students regulate their thinking and actions according to their evaluations of a learning task. What is culturally accepted may be personally abhorrent. There may even be a conflict between studying statistics and a student's perception of her own self realisation or values. For example, in the following survey response one student expressed the idea that real psychologists don't do mathematics.

She wrote:

- I don't even see the point. In psych why must maths infiltrate itself??? Studies have shown that those who have high maths abilities have low or poor communication and perception skills–shouldn't psychologists be exceptionally perceptive and able to communicate well? It seems that if there aren't silly numbers to justify things then they aren't plausible in our computer/maths/science promotive society.

Leont'ev makes an important distinction between meaning and personal sense, which relates to conflicts such as the one expressed by this student. To

Leont'ev, meanings are connected with the reality of the outside objective world, the life of society. Leont'ev calls them the "crystallization of social experience". Personal sense, on the other hand, is connected with the reality of the person's own life and motives. That is, personal sense involves the incorporation of socially constituted meanings into the psychology of the individual.

Personal sense, according to Leont'ev "does not have its own 'supraindividual', 'nonpsychological' existence". Sometimes, there is a mismatch between societal and personal meaning. According to Leont'ev, it is affect that signals to the individual the fit between the two.

Lerman argues that the valuing of decontextualised, intellectual processes, divorced from personal and social elements, is expressive of oppressive discourse. It is this privileging of abstract thought, such as academic mathematics, that is disempowering for some students. Sierpinska and Lerman refer to the vested interest mathematicians have in maintaining the status of mathematics in society. This idea was passionately expressed by a participant in my pilot study for Study Two.

She wrote:

- Maths is an exercise in agony, because the people who teach it make one feel as though maths belongs in a higher plane of evolution. Even though the number system is for everyone, and the concepts are there for everyone, the feeling that you do not deserve to know anything runs rampant. Maths, in short, is a lofty pain and a real headache to study.

In summary, the issues that I am concerned with in this project relate to the ways that students orient themselves to and interpret the task of learning Statistics and to how personal dimensions and individual actions are linked to the wider social, institutional and cultural arenas surrounding them.

ETHNOGRAPHIC RESEARCH

Qualitative and ethnographic research developed in education in the late 1970s. Ethnographic researchers drew on theory and methods in anthropology and sociology, creating a distinction between ethnography of education (work undertaken by anthropologists and sociologists) and ethnography in education (work undertaken by educators to address educational issues). Other forms of qualitative research drew on theories from the humanities and other social and behavioural sciences, adapting this work to educational goals and concerns, often creating new forms (*e.g.*, connoisseurship, a field method approach, interview approaches, and some forms of action research).

In the early development of these traditions, educational researchers struggled for acceptance by both other professionals and policymakers. This phase was characterized by arguments over the value of qualitative methods in contrast to the dominant paradigms of the time–quantitative and experimental approaches. Qualitative and ethnographic researchers argued that questions important to education were left unexamined by the dominant paradigms. Some qualitative

researchers argued for the need to include and represent the voices of people in their research, particularly voices not heard in other forms of research involving large-scale studies.

Questions asked by qualitative and ethnographic researchers generally focus on understanding the local experiences of people as they engage in their everyday worlds (*e.g.*, classrooms, peer groups, homes, communities). For example, some researchers explore questions about ways in which people gain, or fail to gain, access to ways of learning in a diverse world; others focus on beliefs people hold about education and learning; while still others examine how patterns learned within a group are consequential for participation in other groups and situations.

A broad range of perspectives and approaches exist, each with its own historical tradition and theoretical orientation. A number of common dimensions can be identified across these perspectives and approaches. Qualitative and ethnographic researchers in education are concerned with the positions they take relative to participants and data collected. For example, many qualitative and ethnographic researchers engage in observations over a period of time to identify patterns of life in a particular group.

The theoretical orientation chosen guides the design and implementation of the research, including the tools used to collect (*e.g.*, participant observation, interviewing, and collecting artifacts) and analyze data (*e.g.*, discourse analysis, document analysis, content analysis, and transcribing video/audio data). Theory also guides other decisions, including how to enter the field (*e.g.*, the social group, classroom, home, and/or community center), what types and how much data to collect and records to make (*e.g.*, videotape, audiotape, and/or field notes), who to interview (formally and/or informally), how long to remain in the field (*e.g.*, for ethnography, one or more years), and what literature is relevant. It also influences relationships researchers establish with people in local settings, which in turn influences what can be known. Some theoretical perspectives guide researchers to observe what is occurring from a distance by taking the role of passive observer, recording information for analysis once they leave the field. Such researchers often do not interview participants, preferring to "ground" their observations in patterns in the data, without concern for what members understand. These descriptions are called *etic,* or outsider descriptions, because the observer is not concerned with members' understandings.

This approach is in contrast with ones in which researchers join the group and become active participant-observers, at times participating directly in events. Such researchers also make videotape records that enable them to step back from what they thought was occurring to examine closely what resulted from those actions. Those not using video or audio records reconstruct events by constructing retrospective field notes, drawing on their memories of what occurred to create a written record to analyze when they leave the field. Just which type of approach and position researchers take depends on their research goal (s) and theoretical orientation (s) as well as what participants permit.

ADVANCE TO RESEARCH QUESTIONS

Research questions in a qualitative study are generated as part of the research process. Qualitative and ethnographic researchers often begin a study with one or more initiating question (s) or an issue they want to examine. Qualitative and ethnographic research approaches involve a process of interacting with data, reflecting on what is important to members in the local setting, and using this to generate new questions and refine the initial questions.

This interactive and responsive process also influences the data that are collected and analyzed throughout the study. Therefore, it is common for researchers to construct more detailed questions that are generated as part of the analysis as they proceed throughout the study, or to abandon questions and generate ones more relevant to the local group or issues being studied.

For example, in one study of a fifth-grade classroom, the initial research questions were open ended and general: (1) What counts as community to the students and teacher in this classroom? (2) How do the participants construct community in this classroom? and (3) How is participating in this classroom consequential for students and the teacher? As the study unfolded, the research questions became more directed towards what the researcher was beginning to understand about this classroom in particular.

After first developing an understanding of patterns of interactions among participants, the researcher began to formulate more specific questions: (1) What patterns of practice does the teacher construct to offer opportunities for learning? (2) What roles do the social and academic practices play in the construction of community in this classroom? and (3) What are the consequences for individuals and the collective when a member leaves and reenters the classroom community? This last question was one that could not have been anticipated but was important to understanding what students learned and when student learning occurred as well as what supported and constrained that learning. The shifts in questions constitute this researcher's logic of enquiry and need to be reported as part of the dynamic design of the study.

ADVANCE TO DESIGN AND DATA COLLECTION

In designing qualitative studies, researchers consider ways of collecting data to represent the multiple voices and actions constituting the research setting. Typical techniques used in qualitative research for collecting data include observing in the particular setting, conducting interviews with various participants, and reviewing documents or artifacts. The degree to which these techniques are used depends on the nature of the particular research study and what occurs in the local group.

Some studies involve in-depth analysis of one setting or interviews of one group of people. Others involve a contrastive design from the beginning, seeking to understand how the practices of one group are similar to or different from another group. Others seek to study multiple communities to test hypotheses from the research literature (*e.g.*, child-rearing practices are the same in all

communities). What is common to all of these studies is that they are examining the qualities of life and experiences within a local situation. This is often called a situated perspective.

ENTERING THE FIELD AND AHEAD ADMITTANCE TO INSIDER KNOWLEDGE

Entering the research setting is one of the first phases of conducting fieldwork. Gaining access to the site is ongoing and negotiated with the participants throughout the study. As new questions arise, the researcher has to renegotiate access. For example, a researcher may find that the outcomes of standardized tests become an important issue for the teachers and students.

The researcher may not have obtained permission to collect these data at the beginning of the study and must then negotiate permission from parents, students, teachers, and district personnel to gain access to these scores.

Qualitative research involves a social contract with those participating in the study, and informed consent is negotiated at each phase of the research when new information is needed or new areas of study are undertaken.

At such points of renegotiation, researchers need to consider the tools necessary and the ways to participate within the group (*e.g.*, as participant-observer and/or observer-participant, as interviewer of one person or as a facilitator of a focus group, or as analyst of district data or student products). How the researcher conducts observations, collects new forms of data, and analyzes such data is related to shifts in questions and/or theoretical stance (s) necessary to understand what is occurring.

USED OF RESEARCH TOOLS

One of the most frequently used tools, in addition to participant observation, is interviewing. For ethnography and other types of field research, interviews occur within the context of the ongoing observations and collection of artifacts. These interviews are grounded in what is occurring in the local context, both within and across time. Some interviews are undertaken to gain insider information about what the researcher is observing or to test out the developing theory that the researcher is constructing. In contrast, other forms of qualitative research may use interviews as the sole form of data collection. Such interviews also seek meanings that individuals or groups have for their own experience or of observed phenomena. These interviews, however, form the basis for analysis and do not require contextual information from observations. What the people say becomes the basis for exploration, not what was observed.

Other tools used by qualitative and ethnographic researchers include artifact and document analysis (artifacts being anything people make and use). The researcher in a field-based study collects artifacts produced and/or used by members of the group, identifies how these artifacts function for the individual and/or the group, and explores how members talk about and name these artifacts. For some theoretical positions, the artifacts may be viewed as a type of participant

in the local event (*e.g.*, computer programmes as participants). Some artifacts, such as documents, are examined for links to other events or artifacts.

This form of analysis builds on the understanding that the past (and future) is present in these artifacts and that intertextual links between and among events are often inscribed in such documents. In some cases, qualitative researchers may focus solely on a set of artifacts (*e.g.*, student work, linked sets of laws, a photograph collection, or written texts in the environment–environmental print). Such studies seek to examine the range of texts or materials constructed, the patterned ways in which the texts are constructed, and how the choices of focus or discourse inscribe the views that members have of self and others as well as what is possible in their worlds.

Although some qualitative studies focus solely on the documents, field-based researchers generally move between document analysis and an exploration of the relationship of the document to past, present, and future actions of individuals and/or groups. These studies seek to understand the importance of the artifact or document within the lives of those being studied.

CONTINUING DATA ANALYSIS

While conducting fieldwork, researchers reread their field notes and add to them any relevant information that they were not able to include at the time of first writing the notes. While reviewing their field notes, researchers look for themes and information relevant to the research questions. They note this information in the form of theoretical notes (or write theoretical memos to themselves) that may include questions about repeated patterns, links to other theories, and conceptual ideas they are beginning to develop. They also make methodological notes to reconstruct their thinking and their logic of enquiry. Sometimes they make personal notes that reflect their thoughts and feelings about what they are observing or experiencing. These notes allow them to keep from imposing their own opinion on data, helping them to focus on what is meaningful or important to those with whom they are working.

Researchers constantly use contrast to build interpretations that are grounded in the data, within and across actors, events, times, actions, and activities that constitute the social situations of everyday life. Many qualitative (particularly ethnographic) researchers examine material, activity, semiotic (meaning-carrying), and/or social dimensions of everyday life and its consequences for members. The analytic principles of practice that they use include comparing and contrasting data, methods, theories, and perspectives; examining part-whole relationships between and among actions, events, and actors; seeking insider (*emic*) understandings of experiences, actions, practices, and events; and identifying through these what is relevant to the local group.

COVERAGE RESEARCH FINDINGS

The final step in qualitative and ethnographic research is writing an account. The researchers make choices about how to represent the data that illustrate

what was typical about the particular group being studied. Another choice might be to highlight actions of the group that were illustrative of their particular patterns of beliefs. In some studies, several cases are chosen to make visible comparisons across different activities within the group, or across different groups that may have some activities in common. For example, researchers who study classroom interactions might bring together data from different classrooms to make visible principles of practice that are similar in general terms such as asking students to understand various points of view. However, in each classroom, the actions of juxtaposing points of view will be carried out differently due to the different experiences within each classroom.

Researchers also select genres for writing the report that best enable the intended audience to understand what the study made visible that was not previously known or that extended previous knowledge. The researcher does not seek to generalize from the specific case. Rather, qualitative or ethnographic researchers provide in-depth descriptions that lead to general patterns. These patterns are then examined in other situations to see if, when, and how they occur and what consequences they have for what members in the new setting can know, do, understand, and/or produce. In qualitative and ethnographic studies this is often referred to as transferability, in contrast to generalizability.

EXPERIMENT

In scientific inquiry, an experiment is a method of investigating causal relationships among variables. An experiment is a cornerstone of the empirical approach to acquiring data about the world and is used in both natural sciences and social sciences. An experiment can be used to help solve practical problems and to support or negate theoretical assumptions.

CONTROLLED EXPERIMENTS

Many hypotheses in sciences such as physics can experiment causality by noting that, until some phenomenon occurs, nothing happens; then when the phenomenon occurs, a second phenomenon is observed. But often in science, this situation is difficult to obtain.

For example, in the old joke, someone claims that they are snapping their fingers “to keep the tigers away”; and justifies this behaviour by saying “see - it’s working!” While this “experiment” does not *falsify* the hypothesis “snapping fingers keeps the tigers away”, it does not really support the hypothesis - *not* snapping your fingers does keep the tigers away as well.

To demonstrate a cause and effect hypothesis, an experiment must often show that, for example, a phenomenon occurs after a certain treatment is given to a subject, and that the phenomenon does *not* occur in the *absence* of the treatment.

A *controlled* experiment generally compares the results obtained from an experimental sample against a *control* sample, which is practically identical to the experimental sample except for the one aspect whose effect is being tested (the independent variable).

A good example would be a drug trial. The sample or group receiving the drug would be the experimental one; and the one receiving the placebo would be the control one. In many laboratory experiments it is good practice to have several replicate samples for the test being performed and have both a positive control and a negative control.

The results from replicate samples can often be averaged, or if one of the replicates is obviously inconsistent with the results from the other samples, it can be discarded as being the result of an experimental error (some step of the test procedure may have been mistakenly omitted for that sample). Most often, tests are done in duplicate or triplicate. A positive control is a procedure that is very similar to the actual experimental test but which is known from previous experience to give a positive result.

A negative control is known to give a negative result. The positive control confirms that the basic conditions of the experiment were able to produce a positive result, even if none of the actual experimental samples produce a positive result. The negative control demonstrates the base-line result obtained when a test does not produce a measurable positive result; often the value of the negative control is treated as a "background" value to be subtracted from the test sample results. Sometimes the positive control takes the quadrant of a standard curve.

An example that is often used in teaching laboratories is a controlled protein assay. Students might be given a fluid sample containing an unknown (to the student) amount of protein. It is their job to correctly perform a controlled experiment in which they determine the concentration of protein in fluid sample (usually called the "unknown sample").

The teaching lab would be equipped with a protein standard solution with a known protein concentration. Students could make several positive control samples containing various dilutions of the protein standard. Negative control samples would contain all of the reagents for the protein assay but no protein.

In this example, all samples are performed in duplicate. The assay is a colorimetric assay in which a spectrophotometer can measure the amount of protein in samples by detecting a colored complex formed by the interaction of protein molecules and molecules of an added dye. In the illustration, the results for the diluted test samples can be compared to the results of the standard curve (the blue line in the illustration) in order to determine an estimate of the amount of protein in the unknown sample.

Controlled experiments can be performed when it is difficult to exactly control all the conditions in an experiment. In this case, the experiment begins by creating two or more sample groups that are *probabilistically equivalent,* which means that measurements of traits should be similar among the groups and that the groups should respond in the same manner if given the same treatment.

This equivalency is determined by statistical methods that take into account the amount of variation between individuals and the number of individuals in each group. In fields such as microbiology and chemistry, where there is very little variation between individuals and the group size is easily in the millions,

these statistical methods are often bypassed and simply splitting a solution into equal parts is assumed to produce identical sample groups.

Once equivalent groups have been formed, the experimenter tries to treat them identically except for the one *variable* that he or she wishes to isolate. Human experimentation requires special safeguards against outside variables such as the *placebo effect*. Such experiments are generally *double blind*, meaning that neither the volunteer nor the researcher knows which individuals are in the control group or the experimental group until after all of the data have been collected.

This ensures that any effects on the volunteer are due to the treatment itself and are not a response to the knowledge that he is being treated. In human experiments, a subject (person) may be given a stimulus to which he or she should respond. The goal of the experiment is to measure the response to a given stimulus by a test method.

Scientific controls are a vital part of the scientific method, since they can eliminate or minimize unintended influences such as researcher bias, environmental changes and biological variation. Controlled experiments are used to investigate the effect of a variable on a particular system. In a controlled experiment one set of samples have been (or is believed to be) modified and the other set of samples are either expected to show no change (negative control) or expected to show a definite change (positive control).

Examples of Controls

In testing a drug, it is important to carefully verify that the supposed effects of the drug are produced only by the drug itself. Physicians achieve this with a double-blind study in a clinical trial: two (statistically) identical groups of patients are compared, one of which receives the drug and one of which receives a placebo. Neither the patients nor the doctor know which group receives the real drug, which serves both to curb bias and to isolate the effects of the drug.

In experiments where crop yield is affected (*e.g.*, soil fertility), the experiment can be controlled by assigning the treatments to randomly selected plots of land. This mitigates the effect of different soil composition on the yield.

Positive and Negative Control

The simplest forms of controls are positive and negative controls. Positive controls confirm that the procedure is effective in observing the effect (therefore minimizing false negatives). Negative controls confirm that the procedure is not observing an unrelated effect (therefore minimizing false positives). A positive control is a procedure that is very similar to the actual experimental test, but which is known from previous experience to give a positive result.

A negative control is known to give a negative result. The positive control confirms that the basic conditions of the experiment were able to produce a positive result, even if none of the actual experimental samples produce a positive result. The negative control demonstrates the base-line result obtained when a

test does not produce a measurable positive result; often the value of the negative control is treated as a "background" value to be subtracted from the test sample results, or be used as the "100%" value against which the test sample results are weighed.

For example, in an enzyme assay to measure the amount of an enzyme in a set of extracts, a positive control would be an assay where you add some of the purified enzyme, and a negative control would be where you do not add any extract. The positive control should give a large amount of enzyme activity, while the negative control should give very low to no activity.

Necessity of Controls

Controls are needed to eliminate alternate explanations of experimental results. For example, suppose a researcher feeds an experimental artificial sweetener to sixty laboratory rats and observes that ten of them subsequently die. The underlying cause of death could be the sweetener itself or something unrelated. Perhaps the rats were simply not supplied with enough food or water; or the water was contaminated and undrinkable; or the rats were under some psychological or physiological stress, or any other number of variables that may interfere with the experimental design many of which may not be readily obvious.

Eliminating each of these possible explanations individually would be time-consuming and difficult. Instead, the researcher can use an experimental control, separating the rats into two groups: one group that receives the sweetener and one that does not. The two groups are kept in otherwise identical conditions, and both groups are observed in the same ways. Now, any difference in morbidity between the two groups can be ascribed to the sweetener itself—and no other factor—with much greater confidence.

In other cases, an experimental control is used to prevent the effects of one variable from being drowned out by the known, greater effects of other variables. For example, suppose a programme that gives out free books to children in subway stations wants to measure the effect of the programme on standardized test scores. However, the researchers understand that many other factors probably have a much greater effect on standardized test scores than the free books: household income, for example, and the extent of parents' education. In scientific parlance, these are called confounding variables. In this case, the researchers can either use a control group or use statistical techniques to control for the other variables.

NATURAL EXPERIMENTS

The term "experiment" usually implies a controlled experiment, but sometimes controlled experiments are prohibitively difficult or impossible. In this case researchers resort to *natural experiments*, also called *quasi-experiments*. Natural experiments rely solely on observations of the variables of the system under study, rather than manipulation of just one or a few variables as occurs in controlled experiments.

To the degree possible, they attempt to collect data for the system in such a way that contribution from all variables can be determined, and where the effects of variation in certain variables remain approximately constant so that the effects of other variables can be discerned. The degree to which this is possible depends on the observed correlation between explanatory variables in the observed data. When these variables are *not* well correlated, natural experiments can approach the power of controlled experiments.

Usually, however, there is some correlation between these variables, which reduces the reliability of natural experiments relative to what could be concluded if a controlled experiment were performed. Also, because natural experiments usually take place in uncontrolled environments, variables from undetected sources are neither measured nor held constant, and these may produce illusory correlations in variables under study.

Much research in several important science disciplines, including economics, political science, geology, paleontology, ecology, meteorology, and astronomy, relies on quasi-experiments. For example, in astronomy it is clearly impossible, when testing the hypothesis "suns are collapsed clouds of hydrogen", to start out with a giant cloud of hydrogen, and then perform the experiment of waiting a few billion years for it to form a sun. However, by observing various clouds of hydrogen in various states of collapse, and other implications of the hypothesis (for example, the presence of various spectral emissions from the light of stars), we can collect data we require to support the hypothesis.

An early example of this type of experiment was the first verification in the 1600s that light does not travel from place to place instantaneously, but instead has a measurable speed. Observation of the appearance of the moons of Jupiter were slightly delayed when Jupiter was farther from Earth, as opposed to when Jupiter was closer to Earth; and this phenomenon was used to demonstrate that the difference in the time of appearance of the moons was consistent with a measurable speed.

A natural or Quasi-experiment is a naturally occurring instance of observable phenomena which approximate or duplicate the properties of a controlled experiment. In contrast to laboratory experiments, these events aren't created by scientists, but yield data which nonetheless can be used (commonly through the use of instrumental variables) to make causal inferences. Natural experiments are a common research tool in fields where artificial experimentation is difficult, such as economics, cosmology, epidemiology, and sociology.

An example of a natural experiment occurred in Helena, Montana during the period from June 2002 to December 2002 when a smoking ban was in effect in all public spaces in Helena including bars and restaurants. Helena is geographically isolated and served by only one hospital. It was observed that the rate of heart attacks dropped by 60% while the smoking ban was in effect.

A second example is discussed in Angrist and Evans (1988). The authors wish to estimate the effect of family size on the labour market outcomes of the mother. The correlations between family size and various outcomes do not tell

us how family size causally affects labour market outcomes because both labour market outcomes and family size may be affected by unobserved variables such as preferences and because labour market outcomes may "reverse" cause family size (for example, a woman may defer having a child if she gets a raise at work).

The study notes that two-children families with either two boys or two girls are substantially more likely to have a third child than two-children families with one boy and one girl. The sex of the first two children, then, forms a natural experiment: it is as if an experimenter has randomly assigned some families to have two children and others to have three or more. The authors are then able to credibly estimate the causal effect of having a third child on labour market outcomes.

OBSERVATIONAL STUDIES

Observational studies are very much like controlled experiments except that they lack probabilistic equivalency between groups. These types of experiments often arise in the area of medicine where, for ethical reasons, it is not possible to create a truly controlled group.

For example, one would not want to deny all forms of treatment for a life-threatening disease from one group of patients to evaluate the effectiveness of another treatment on a different group of patients. The results of observational studies are considered much less convincing than those of designed experiments, as they are much more prone to selection bias. Researchers attempt to compensate for this with complicated statistical methods such as propensity score matching methods.

In statistics, an observational study draws inferences about the effect of a treatment on subjects, where the assignment of subjects into a treated group versus a control group is outside the control of the investigator. This is in contrast with controlled experiments, such as randomized controlled trials, where each subject is randomly assigned to a treated group or a control group before the start of the treatment.

Reasons for Uncontrolled Experimentation

The assignment of treatments may be beyond the control of the investigator for a variety of reasons:

- A randomized experiment would violate ethical standards. Suppose one wanted to investigate the abortion–breast cancer hypothesis, which postulates a causal link between induced abortion and the incidence of breast cancer. In a hypothetical controlled experiment, one would start with a large subject pool of pregnant women and divide them randomly into a treatment group (receiving induced abortions) and a control group (bearing children), and then conduct regular cancer screenings for women from both groups. Needless to say, such an experiment would run counter to common ethical principles. (It would also suffer from

various confounds and sources of bias, *e.g.*, it would be impossible to conduct it as a blind experiment.) The published studies investigating the abortion–breast cancer hypothesis generally start with a group of women who already have received abortions. Membership in this "treated" group is not controlled by the investigator: the group is formed after the "treatment" has been assigned.

- The investigator may simply lack the requisite influence. Suppose a scientist wants to study the public health effects of a community-wide ban on smoking in public indoor areas. In a controlled experiment, the investigator would randomly pick a set of communities to be in the treatment group. However, it is typically up to each community and/or its legislature to enact a smoking ban. The investigator can be expected to lack the political power to cause precisely those communities in the randomly selected treatment group to pass a smoking ban. In an observational study, the investigator would typically start with a treatment group consisting of those communities where a smoking ban is already in effect.
- A randomized experiment may be impractical. Suppose a researcher wants to study the suspected link between a certain medication and a very rare group of symptoms arising as a side effect. Setting aside any ethical considerations, a randomized experiment would be impractical because of the rarity of the effect. There may not be a subject pool large enough for the symptoms to be observed in at least one treated subject. An observational study would typically start with a group of symptomatic subjects and work backwards to find those who were given the medication and later developed the symptoms. Thus a subset of the treated group was determined based on the presence of symptoms, instead of by random assignment.

Discussion

In all of those cases, if a randomized experiment cannot be carried out, the alternative line of investigation suffers from the problem that the decision of which subjects receive the treatment is not entirely random and thus is a potential source of bias. A major challenge in conducting observational studies is to draw inferences that are acceptably free from influences by overt biases, as well as to assess the influence of potential hidden biases.

An observer of an uncontrolled experiment (or process) records potential factors and the data output: the goal is to determine the effects of the factors. Sometimes the recorded factors may not be directly causing the differences in the output. There may be more important factors which were not recorded but are, in fact, causal. Also, recorded or unrecorded factors may be correlated which may yield incorrect conclusions.

In observational studies, investigators may use propensity score matching (PSM) in order to reduce overt biases. In 2007, several prominent medical

researchers issued the *Strengthening the Reporting of Observational Studies in Epidemiology* (STROBE) statement, in which they called for observational studies to conform to 22 criteria that would make their conclusions easier to understand and generalise.

Field Experiments

Field experiments are so named in order to draw a contrast with laboratory experiments. Often used in the social sciences, and especially in economic analyses of education and health interventions, field experiments have the advantage that outcomes are observed in a natural setting rather than in a contrived laboratory environment. However, like natural experiments, field experiments suffer from the possibility of contamination: experimental conditions can be controlled with more precision and certainty in the lab.

A field experiment applies the scientific method to experimentally examine an intervention in the real world (or as many experimental economists like to say, naturally-occurring environments) rather than in the laboratory.

Field experiments, like lab experiments, generally randomize subjects (or other sampling units) into treatment and control groups and compare outcomes between these groups. Clinical trials of pharmaceuticals are one example of field experiments. Economists have used field experiments to analyse discrimination, health care programmes, charitable fundraising, education, information aggregation in markets, and microfinance programmes.

Of course, the use of experiments in the lab and the field has a long history in the natural and life sciences. Geology has a long history of field experiments, since the time of Avicenna.[1] Social psychology also has a history of field experiments, including work 8 by pioneering figures Philip Zimbardo, Kurt Lewin and Stanley Milgram. In economics, Peter Bohm, University of Stockholm, was one of the first economists to take the tools of experimental economic methods and attempt to try them with field subjects.

In the area of development economics, the pioneer work of Hans Binswanger in the late 1970s conducting experiments in India on risk behaviour should also be noted. Not much work ensued quickly but the use of field experiments in economics has grown tremendously recently with the work of John A. List, Jeff Carpenter, Juan-Camilo Cardenas, Abigail Barr, Catherine Eckel, Michael Kremer, Glenn Harrison, Colin Camerer, Bradley Ruffle, Esther Duflo, Dean Karlan, Edward "Ted" Miguel, Sendhil Mullainathan, David H. Reiley, among others.

Applications

Recent work by Glenn W. Harrison (University of Central Florida) and John A. List (University of Chicago) who has put forward a taxonomy of field experiments. See their paper in the December 2004 issue of the *Journal of Economic Literature* for a complete treatment or List's website for a quicker overview.

Their taxonomy partitions field experiments into three categories ranging from those that most closely resemble traditional laboratory experiments or compensated survey questions to those that are truly "natural" field experiments in the sense that the subjects involved are unaware of any treatment taking place. (Note that artificial experiments in social psychology often use deception, so that subjects are also unaware of the true treatment).

See List's website for many applications of the field experiment method, including the analysis of public good contributions, charitable giving, market anomalies, discrimination, education, health care and microfinance.

METHODOLOGY

Compared with Natural Experiments

A natural experiment is halfway between a field experiment and an observation. In a natural experiment the independent variable is not introduced by the experimenter - for example, the introduction of television on the island of St Helena.

Compared with Non-Experimental Field Data

Caveats

- Fairness of randomization (*e.g.*, in 'negative income tax' experiments communities may lobby for their community to get a cash transfer so the assignment is not purely random)
- Contamination of the randomization
- General equilibrium and "scaling-up"
- Difficulty of replicability (field experiments often require special access or permission, or technical detail— *e.g.*, the instructions for precisely how to replicate a field experiment are rarely if ever available in economics)
- Limits on ability to obtain informed consent of participants

THE BIRTH OF EDUCATION AS A FIELD OF STUDY AND RESEARCH

The study of and research on education traces its roots back to the late 1830's and early 1840's with the revival of the common school and it is the first time that both school supervision and planning were influenced by systematic data collection. These data collection efforts, according to Robert Travers, involved "an examination of the ideas on which education was based, an intellectual crystallization of the function of education in a democracy, and the development of a literature on education that attempted to make available to teachers and educators important new ideas related to education that had emerged in various countries." Horace Mann and Henry Barnard were early pioneers in educational

data collection and in the production and dissemination of educational literature during the mid to late-nineteenth century. Additionally, they held prominent educational leadership positions by being the first secretaries of educational boards of Massachusetts (Horace Mann) and of Connecticut (Henry Barnard). In many ways, the trends in the early history of educational research were components of the trends in American culture of the time.

The founding of Johns Hopkins University as the first research university in 1876 set the stage for new elite research universities to be founded such as Stanford University and the University of Chicago. Additionally, the Morrill Act of 1862 allowed for the establishment of 'land-grant' colleges and universities, many of which would rival the more established elite institutions on the east coast in research and knowledge production, across the United States. As Ellen Condliffe Lagemann points out, research universities quickly became the leaders in creating and disseminating new knowledge, the professionalization of many professions and they became the "spawning grounds" for research on education at the end of the nineteenth century.

During this time period, there was a belief that the social world could be "acted on and changed through scientific practices …and that teaching and the social welfare professions embodied scientific analysis and planning." The restructuring of higher education in the United States from a focus on teaching to a new focus that included both teaching and research activities lead to new schools of thought and approaches to science. Professors at universities were now expected to teach and to plan and conduct original research. Numerous pioneers of American education began their work and research in the major research institutions of the day. Perhaps one of the most well known of these scholars was John Dewey, Chair of the Department of Philosophy at the University of Chicago from 1896 to 1904, who introduced a new approach to the study of education and became a leader in pedagogy.

Dewey's experimental Laboratory School was based more on psychology than on behaviourism which had long influenced educational research activities. John Dewey's progressive education philosophy opposed testing and curriculum tracking and relied more on argument than on scientific research and its evidence. He worked to combine philosophy, psychology and education. Surprisingly, John Dewey never proposed future areas of enquiry or suggested future research directions in his writings and he never published any evidence on the effects his Laboratory School experiment had on children. Dewey's influence on educational practice outside of his Laboratory School was quite limited and overestimated. Ellen Condliffe Lagemann summed up John Dewey's legacy on educational research as follows: "to suggest that Dewey had served as something of a cultural icon, alternatively praised and damned by thinkers on both the right and left, might capture his place in the history of education more accurately than to say he was important as a reformer. Certainly, his ideas about a science of education did not create a template for educational study." In 1904, John Dewey left the University of Chicago for Teachers College at Columbia

University where he remained as a professor of Philosophy until his death in 1952. Within five years of John Dewey's departure, Charles Judd arrived in 1909 to serve as Chair of the School of Education at the University of Chicago. Charles Judd differed substantially from John Dewey in his approach to educational research. Charles Judd, a psychologist, sought to bring a rigourous and scientific approach to the study of education. Judd was a proponent of the scientific method and worked to integrate it into educational research. This was evidenced by the University of Chicago's School of Education reorganization into the Department of Education within the Division of the Social Sciences shortly after Judd's arrival on campus. Judd's preference for quantitative data collection and analysis and his emphasis on the scientific method, with a particular focus on psychology, was one of the leading schools of thought on educational research during the early decades of the twentieth century.

In 1904, the same year that John Dewey left the University of Chicago for Teachers College at Columbia University, the psychologist Edward Thorndike, also of Teachers College, published An Introduction to the Theory of Mental and Social Measurements which argued for a strong positivistic theoretical approach to educational research. Thorndike held a similar epistemological approach to the study of education to that of Charles Judd at Chicago. Thorndike favoured the separation of philosophy and psychology. He did not care for the collection of data for census purposes but rather the production of statistics and precise measurements that could be analyzed. Thorndike became a very influential educational scholar and his approach to educational research was widely accepted and adopted across academia both in the United States and abroad. What Ellen Condliffe Lagemann describes as "Edward Thorndike's triumph and John Dewey's defeat" was critical to the field and to attempts to define an educational science.

EDUCATIONAL RESEARCH AND THE INTER-WAR YEARS

The inter-war years were a time of transformation in educational research. By 1915, the study of education had been established at the university level with 300 of the 600 institutions of higher education in the United States offering courses on education. This time period also experienced an increase at the doctoral level of study which saw enrollments higher than any other discipline other than chemistry. While faculty at institutions such as Harvard, Teachers College and the University of Chicago, which had dominated the educational research landscape decades earlier, continued to make significant contributions to the study of education, there were scholars at many other institutions making additional valuable contributions to educational research scholarship. At the conclusion of World War I the focus on educational reform in the United States began to change to a more social control and efficiency and there was an opportunity for many educationists to provide guidance to public schools in the United States. Disagreement among educational research scholars persisted during this time period and there was little consensus on the aims of education.

With the population of the United States growing rapidly and the demographic make-up of its people changing due to the arrival of immigrants from across the globe coupled with the migration of African Americans from rural areas and the Southern states to the urban cities in the Northeast and Midwest the student bodies at public schools were diversifying at a rapid pace. The arrival of new immigrants to the United States coincided with the "testing movement" that emerged during World War I when the United States Army was testing its recruits. The most prominent psychologists of the time, including Edward Thorndike, were either involved with or supported the Army's testing. The testing movement attracted both psychologists and sociologists alike and it was the sociologists, primarily in the Department of Sociology at the University of Chicago, who challenged and actively researched the racial differences in intelligence quotients. Otto Klineberg from Columbia University also played a leading role in studying racial and cultural differences in intelligence quotients and their measures.

Educational Research - Post World War Ii and the Future

Educational research continued to flourish in the years and decades after World War II ended. During this time period, the growth in schools of education and in the number of courses on education at institutions of higher education continued to rise. Additionally, more academic journals with a focus on educational issues emerge as a means to disseminate new knowledge. These exciting changes in educational scholarship were not confined to the ivory towers in the United States. Even as Europe was rebuilding, the study of education across the continent was on the rise and in the United Kingdom, for example, the rise of professional graduate degrees in education was significant. Scholarly debates on the aims of education as well as epistemological discourse persisted.

In the decades after World War II, and in particular at the start of the 1960's, a post-positivist movement in educational research starts taking shape. While positivistic approaches to educational research continued to be put forth during the post-war years and continued to be favoured by many social scientists, we start to see the introduction of, and in some case the reemergence of, other epistemological approaches. Constructivism, functionalism and postmodernism theoretical frameworks, among others, have offered strong criticisms of positivism. Vigourous debates on the virtues of the various theoretical perspectives about knowledge, science and methodologies have played a very important role in educational research. Frequently, these critical discussions and analyses have found both a platform and a captive audience in the field of comparative education. These philosophical debates continue today both in and outside of academia.

The United States federal government also began to take a much more active role in educational research in the post-war years. Specifically, in 1954, the United States Congress passed the Cooperative Research Act. The Cooperative Research Act was passed as a means for the federal government to take a more active role in advancing and funding research on education in academia.

Additional legislation and federal initiatives during the 1950's and 1960's that supported and/or funded educational research and provided a means for the dissemination of new educational knowledge included the National Defence Education Act of 1958 and the establishment of the Educational Resources Information Center (ERIC) in 1966.

These are but a few of the many examples of the new role the federal government was playing in educational research during this time period. To be sure, the federal government has continued to play a significant role in educational research since this time period. Since the 1970's, according to Robert Travers, "virtually every bill authorizing particular educational programmes has included a requirement that the particular programme be evaluated to determine whether the programme was worth the money spent upon it." For a long period of time, public focus on education and schools focused on resource allocation, student access, and the content of the curriculum and paid relatively little attention to results. This new "evidenced-based movement" is one that remains with us today. Patti Lather describes the evidenced-based movement as "governmental incursion into legislating scientific method in the real of educational research" and that the federal government's focus on evidence-based knowledge is much more about policy for science than it is about science for policy. The federal government has a vested interest in and support for "applied research" over "basic or pure research". This, of course, is challenging for social scientists and educational researchers who are positivist in their approach to science and knowledge.

A distinctive form of research emerged from the new assessment or evaluation movement in recent decades. Educational assessment, in many ways, is a form of "action research". Action research does not aim to produce new knowledge. Instead, action research aims to improve practice and in the context of education it aims to improve the educational practice of teachers. Action research, as Richard Pring points out, "might be supported and funded with a view to knowing the most effective ways of attaining particular goals – goals or targets set by government or others external to the transaction which takes place between teacher and learner." Action research proponents, Yvonna Lincoln and Egon Guba, highlight that "the call for action…differentiates between positivist and postmodern criticalist theorists."

A new and interesting approach to educational research can be found today at the University of Chicago. The Department of Education at the University of Chicago was closed in 1997 to much surprise around the world. Despite the closing of the Department of Education, a sprinkling of educational research activities by faculty and available education course offerings can be found in a variety of academic departments and professional schools. In addition, the University of Chicago has also operated the North Kenwood/Oakland Charter School under the Center for Urban School Improvement since 1998. Campus interest in urban schools and educational research led to the creation of a new Committee on Education in 2005, with a home in the Division of the Social

Sciences, chaired by Stephen Radenbusch who joined the faculty in the Department of Sociology and whom the University lured from the School of Education at the University of Michigan.

The University of Chicago Chronicle highlighted the arrival of Stephen Radenbusch and noted that the Committee on Education "will bring together distinguished faculty from several departments and schools considered to be among the best in the world into common research projects, seminars and training programmes. The committee will engage faculty and graduate students from such areas as public policy, sociology, social service administration, economics, business, mathematics and the sciences to collaborate on the most critical issues affecting urban schools." The interdisciplinary focus of Chicago's Committee on Education and its Urban Education Initiative plans to create a "Chicago Model" for urban schools that will "draw on and test the best ideas about teaching, learning, school organization, school governance, teacher preparation, and social service provision."

While interdisciplinary research and collaboration is no stranger to the University of Chicago, it is a new and innovative approach to the study of education. The Committee on Education at the University of Chicago is highly quantitatively driven and data focused. If this interdisciplinary approach to educational research is successful and is modeled by other institutions of higher education, both in the United States and abroad, it will be interesting to see if a positivistic approach similar to that found at Chicago is followed or if a more relativistic approach is pursued. Either way, interdisciplinary collaboration may very well be the next chapter in the history of educational research.

STATISTICS EDUCATION AT UNIVERSITY

"Statistics" said Jane, a mathematics student:

- Really fits my logic and it feels right. Maybe the men have finally invented a good mathematical study for a change.

The comment highlights the tension between knowledge as culturally endorsed and the individual's personal appraisal of it, a tension not always as agreeably resolved as in Jane's case. Statistical literacy and appreciation, particularly an understanding of data gathering, presentation and interpretation, are important components of undergraduate and graduate education in many fields. They are also essential for modern living in technologically developed countries.

Statistics is often a compulsory unit of university courses, such as Psychology, Economics, Business or a Health Science, because it is an important tool for analysing the"uncertainties and complexities of life and society". Despite this importance, many students in statistics courses are reluctant to study it.

Indeed, Cotts maintains that it is:

- Almost unarguable that the introductory statistics course is the most widely feared course on most university campuses.

This conviction is echoed in several studies.

RESEARCH IN STATISTICS EDUCATION

There has been considerable research into statistics education in the last thirty years. This interest in the teaching and learning of statistics is not only a part of the general growth spurt in mathematics education but is also a field of research in its own right. Four international conferences, the International Conferences on the Teaching of Statistics have been devoted to grappling with issues relating to statistics education.

Much topical literature concerns the teaching and learning of statistics at all levels of education. For example, in a special edition on teaching statistics of the Journal of Educational and Behavioural Statistics, Becker reviews 530 articles and dissertations which are documented on electronic databases such as the Educational Resources Information Centre. Almost all these articles (97%) were published after 1970 and about one third have appeared since 1990.

A bibliography on the available literature published between 1987 and 1994, concerning the teaching of probability and statistics, is provided by Sahai, Khurshid and Misra. This shows that the body of work being done in this area is extensive and growing.

Truran and Truran review Australasian research into the learning and teaching of stochastics, for the period 1992-1995, which reflects the current importance of probability, statistics and combinatorics to education in my region of the world. Electronically based journals such as the Journal of Statistics Education and Web discussion groups facilitate the extensive exchange of ideas and the development and distribution of research among participants in statistics education.

The quantity and scope of the published literature, as well as the ongoing discussion and research, is a strong indication of current interest in statistics education.

The literature and discourse concerns many topics. One important topic concerns teaching strategies. For example, Lan writes about self monitoring and Smith discusses the use of writing assignments in teaching statistics.

The role of computers, multimedia materials and other technological aids in teaching statistics is an area of increasing prominence. Individual differences, such as gender differences in learning statistics, have also been investigated. For instance, Clark proposes that female students prefer statistical questions which have a people-orientation. Preliminary results by Forbes however, suggest that, in statistical examinations in New Zealand, the form and type of examination questions, for example, whether questions involve essay writing or require using Calculus, are at least as important as their contextual embedding in determining gender preferences and performance.

As with any other field in education, what constitutes statistical knowledge and how to enhance the quality of students' learning of it, are key problems in statistics education. A major theme that recurs in the literature on higher education in statistics pertains to teaching and assessing statistics. There is general consensus in the community of statistics educators that reforms are

urgently needed in the teaching of introductory statistics courses at university. Some university educators describe teaching activities and/or methods of assessment intended to address the problems - to make statistical concepts meaningful and useful - to relate statistics to the real world.

For example, Anderson and Loynes and Pfannkuch discuss what abilities and skills are needed by practising statisticians and how to teach these. Garfield Giraud and Keeler and Steinhorst describe new ways of teaching statistics using co-operative learning activities. Garfield and Hubbard present frameworks for developing assessment instruments and procedures which are appropriate for measuring students' understanding and applications of concepts in probability and statistics.

Many university educators report on the success of their statistics courses by describing enhanced performance and/or increased student satisfaction. However, as some researchers point out, little is known about how such courses would transfer to other settings. Becker concurs with this view. In her review of the literature and other resources on teaching statistics, she focuses mainly on statistics education at the university level. She reports that this literature is largely anecdotal, with less than 30% of the reviewed articles describing the results of empirical studies. Hence, extensive information on instructional strategies and resources are available to tertiary educators in statistics, either in print or through electronic media. However, research empirically evaluating the teaching and learning of statistics is less copious. Further, and in stark contrast to the literature on mathematics education, there is a dearth of studies on the teaching and learning of statistics which are framed in terms of theories of education. The literature on teaching and learning statistics at the tertiary level tends to concentrate on the "knowledge craft" of the teachers.

A further and major area of research in statistics education concerns students' understanding of particular concepts in statistics and probability theory. This research shows that many students have difficulties with and misconceptions about statistical ideas. Konold reviewed research showing that the intuitions of adults are at odds with accepted probability and statistical theory. Fischbein and Schnarch investigated the stability of students' misconceptions in probability across different age levels from grade 5 to college students. Surprisingly, they found three different outcomes: some misconceptions grew stronger with age, some grew weaker and one remained stable.

Vallecillos and Batanero presented a study on the persistence of conceptual errors in university students' understanding of levels of significance in tests of hypotheses. There are also parts on students' understanding and misunderstanding of concepts relating to probability and statistics in books on the teaching of statistics such as Green; Holmes; Hawkins, Jolliffe and Glickman. These attest to the importance of conceptual understanding as a goal of statistics instruction.

In summary, research on student learning in statistics shows that students have difficulties with statistical concepts and that they are often anxious about

and have poor attitudes to learning statistics. The body of research on teaching and learning statistics is large, vibrant and growing. However, much of it still lacks the systematic methodology and theoretical foundations that are often of the most lasting value in any field of education. Further, a concern with describing teaching strategies or students' attitudes or their misconceptions, while indicating an awareness of the problems facing statistics education, does not offer a model for understanding the impasse–a first step towards alleviating it. My approach is to explore the issues underlying this impasse within a theoretical framework which relates students, their actions and the context.

STRATEGY TO APPROACH IN STYLES- AND STRATEGIES-BASED INSTRUCTION

Traditionally, the emphasis has primarily been on the teaching side of second language instruction, rather than on the learner side. It has been assumed that if teachers do their job of teaching well, students would certainly learn and retain the language as well. Yet, it became clear that if students are not learning or are not motivated to learn, it may not matter how well the teachers are teaching. With this realization an effort has emerged to improve language teaching methodology by adding a component that focuses on the learner.

As the "domain" of language teaching has become more learner-focused and interactive, there has also been a heightened emphasis on helping students take more responsibility for meeting their own language learning needs. Students are being asked to self-direct the language-learning process and become less dependent on the classroom teacher. However, what may well stand in the way of learners' genuine success at language learning is an insufficient awareness of how various strategies may help them learn and use a foreign language more effectively.

Given that language learning and the use of what is learned inevitably involve considerable memory work, as well as ongoing and meaningful practice, a systematic and purposeful approach to learning can help to ease the burden. And the classroom teacher can perform a key role in this effort as learner trainer. Styles- and strategies-based instruction is a name that has been given to a form of learner-focused language teaching that explicitly combines styles and strategy training activities with everyday classroom language instruction. The underlying premise of the styles- and strategies-based approach is that students should be given the opportunity to understand not only what they can learn in the language classroom, but also how they can learn the language they are studying more effectively and efficiently. Research seems to suggest that there are a wide variety of strategies that learners can use to meet their language learning and using needs.

Styles- and strategies-based approach to teaching emphasizes both explicit and implicit integration of language learning and use strategies in the language classroom. This approach aims to assist learners in becoming more effective in their efforts to learn and use the target language. SSBI helps learners become

more aware of what kinds of strategies are available to them, understand how to organize and use strategies systematically and effectively given their learning-style preferences, and learn when and how to transfer the strategies to new language learning and using contexts. SSBI is based on the following series of components:

Strategy Awareness-Raising

In this phase, the goal is to alert learners to presence of strategies they might never have thought about or may have thought about but had never used. The SSBI tasks are explicitly used to raise the students' general awareness about: 1) what the learning process may consist of, 2) their learning style preferences or general approaches to learning, 3) the kinds of strategies that they already employ, as well as those suggested by the teacher or classmates, 4) the amount of responsibility that they take for their learning, or 5) approaches that can be used to evaluate the students' strategy use. Awareness-raising activities are by definition always explicit in their treatment of strategies.

Strategy Preparation

In this phase, the goal is to determine just how much knowledge of and ability to use strategies the given learners already have. There is no sense in assuming that students are a blank slate when it comes to strategy use. They most likely have developed some strategies. The thing is that they may not use them systematically, and they may not use them well.

Strategy Training

In this phase, students are explicitly taught how, when, and why certain strategies (whether alone, in sequence, or in clusters) can be used to facilitate language learning and use activities. In a typical classroom strategy-training situation, the teachers describe, model, and give examples of potentially useful strategies. They elicit additional examples from students based on the students' own learning experiences; they lead small-group or whole-class discussions about strategies (*e.g.*, the rationale behind strategy use, planning an approach to a specific activity, evaluating the effectiveness of chosen strategies); and they can encourage their students to experiment with a broad range of strategies.

Personalization of Strategies

In this stage, learners personalize what they have learned about these strategies, evaluate to see how they are using the strategies, and then look to ways that they can transfer the use of these strategies to other contexts.

In SSBI, it is the curriculum writers' and the teachers' role to see that strategies are integrated into everyday class materials and are both explicitly and implicitly embedded into the language tasks to provide for contextualized strategy practice. Teachers may:

- Start with the established course materials and then determine which strategies might be inserted,
- Start with a set of strategies that they wish to focus on and design activities around them, or
- Insert strategies spontaneously into the lessons whenever it seems appropriate.

These strategies-based activities are designed to raise awareness about strategies, to train students in strategy use, to give them opportunities to practice strategy use, and to encourage them to personalize these strategies for themselves. Teachers also allow students to choose their own strategies and do so spontaneously, without continued prompting from the language teacher.

Strategy Practice

In this phase, students are encouraged to experiment with a broad range of strategies. It is not assumed that knowing about a given strategy is enough. It is crucial that learners have ample opportunity to try them out on numerous tasks. These "strategy-friendly" activities are designed to reinforce strategies that have already been dealt with and allow students time to practice the strategies at the same time they are learning the course content. These activities should include explicit references to the strategies being used in completion of the task. In other words, either students:

- Plan the strategies that they will use for a particular activity,
- Have their attention called to the use of particular strategies while they are being used, or
- "debrief" their use of strategies (and their relative effectiveness) after the activity has ended.

INVESTIGATING INDIVIDUALS IN SOCIALLY SIGNIFICANT PRACTICES

My project is an example of what Chaiklin describes as "investigating individuals in socially significant practices".

These are studies in a "yet-to -be-embodied" research tradition that:

- Try and develop an account of the actions of individuals participating in a societally significant practice while it is occurring, by an analysis that locates the practice in a social, societal and/or historical perspective.

Chaiklin proposes that many different theoretical perspectives and methods could contribute to such research. He summarises such projects as having five common characteristics, which I assert are fulfilled by my project.

- Firstly, these studies take concrete, meaningful, societal practices as the direct object of study. That is, the setting for the study is an actual example of human practice and would be so whether or not it was the object of study. In my research I have tried to get as close as possible to the participants and interpret situations which were as typical as possible for students learning Statistics.

- Secondly, the practice takes place in a recognisable societal institution, in this case an institution of higher learning.
- Thirdly, such studies have a definite theoretical interest. Insights into the practice of learning Statistics are important elements of my research but so are insights into and extensions of the activity theory framework in which I have interpreted this practice.
- Fourthly, there is the idea that knowledge and actions are social as well as individual. I draw on Vygotsky's and Leont'ev's idea of "co-knowing" to illustrate this notion–statistical knowledge is socially created and re-created through the consensus of practitioners. This collective development of knowledge is facilitated through tools, for example, by the easy and quick exchange of ideas through electronic media. In my university, Statistics is continually being transformed by each lecturer and committee involved in organising the curriculum and setting the examinations. Students come to know Statistics through interacting with peers and teachers, as well as through their own reflections and actions. Socially organised constraints and resources surround individual actions.
- *Finally, the particular practice selected for such studies has:*
  - Significant consequences for the people participating in these practices.

Learning Statistics can change the lives of those participating in this practice. In some cases, the outcomes could be instrumental in determining students' academic and career paths. For example, students' examination results in Statistics could determine whether they would attain a degree or be accepted into Psychology Honours. Learning Statistics can also powerfully affect students' personal development as was the case for some of the participants of Study One.

EDUCATIONAL TECHNIQUES AND METHODOLOGY

"Educational Techniques and Methodology" offers a comprehensive exploration of effective strategies for teaching and learning. This invaluable resource delves into various pedagogical approaches and instructional methods designed to engage students, promote deep understanding, and foster academic success. Through a combination of theoretical insights and practical applications, educators gain a deeper understanding of how to create dynamic and inclusive learning environments that cater to diverse learning styles and preferences. "Educational Techniques and Methodology" addresses assessment methods and techniques for evaluating student learning effectively. Educators learn how to design authentic assessment tasks that measure students' understanding, skills, and competencies in meaningful ways. From formative assessments that provide ongoing feedback to summative assessments that evaluate overall learning outcomes, the book offers practical guidance on how to assess student progress and inform instructional decision-making. Through reflective exercises and case studies, educators are encouraged to critically examine their teaching practices and explore innovative approaches to enhance student learning outcomes. Overall, "Educational Techniques and Methodology" equips educators with the knowledge and tools needed to create dynamic, student-centered learning environments that inspire curiosity, foster critical thinking, and promote lifelong learning. Delving into effective pedagogical practices, this book offers a comprehensive exploration of educational techniques and methodologies, providing educators with valuable insights and practical strategies for enhancing teaching and learning.

Dr. Anil Kumar, is associated with academics for more than 17 years. He holds the degree of M.A. from MGS University, Bikaner, and B.Ed. from Kota University, Kota, and M.Ed. from IGNOU, New Delhi and Doctor of Philosophy in Education from VBP University, Jaunpur (U.P.). He is supervising research work in the field of Education. At present he is working as Associate Professor, Faculty of Education, Tantia University, Sri Ganganagar (Rajasthan). He has been continuously involved in research work with a bulk of publication in renowned journals and attended various conferences and seminars and workshops to share and impart his knowledge and view. 06 Ph.D. Research Scholars have been awarded Ph.D. degree under his supervision. He has participated and presented papers in more than twenty (20) International/National Seminars and already three (03) books are published by him.

ACADEMIC UNIVERSITY PRESS

4378/4-B, Murarilal Street, Ansari Road, Daryaganj, New Delhi-110002
Phone : +91-11-23281685, 41043100, Fax: +91-11-23270680
E-Mail: academicuniversitypress@gmail.com

ISBN-978-93-6284-300-5