

A Competence-based Perspective on Foreign Language Teaching at the Master Level: Insights from John Dewey's Educational Philosophy

Svetlana MUNTJAN¹
Lilia VALENTINOVA²

Abstract

The article determines the specific bundle of competences and strategies underpinning the modern process of foreign language teaching by building upon the selected elements of educational philosophy of John Dewey. The article proceeds by reconstructing some of these elements and subsequently by drawing implications for the competence development strategies in the foreign language teaching. The main implications include adopting cutting-edge teaching technologies and fine-tuning and individualizing teaching materials. A key aspect of teaching technologies to be adopted is assisting students in developing their own self-study strategies.

Keywords:

Sociocultural competence, Dewey, learner-centred approach, self-study, learning strategies, teaching activities

¹ Svetlana MUNTJAN Head of Department of Foreign Languages at the Tavria State Agrotechnological University, Melitopol, Ukraine Email: muntjan@yandex.ru

² Lilia VALENTINOVA senior teacher at the Department of Foreign Languages at the Tavria State Agrotechnological University, Melitopol, Ukraine Email: lilia-valentinova@yandex.ru

Introduction

The importance of sociocultural competence for modern university graduates is universally acknowledged. Yet, in practical terms, developing this competence still poses a number of challenges related to the background competence of master students and availability of academic hours. While information sources on relevant topics abound, there remains a deficit of comprehensive training packages that are finely attuned to the unique needs of master students in agricultural universities. This article explores the development of sociocultural competence at the master-level foreign language teaching.

The competence-based approach was founded in 1965 by N. Chomsky. According to this approach, students are supposed to assimilate the learning matter in the form of integrated wholes rather than separate and loosely connected pieces of information. The advocates of this approach point out the dramatic rate of societal change and the futility of making timely adjustments in the learning matter. Instead, they lay stress on the formation of key competences that would enable students to deal with this change as it proceeds. In the course of learning, generic competence development is thus supposed to take priority over the particular contents of professional education.

The competence-based approach is fully endorsed in the Ukrainian strategy of modernization of foreign language teaching. Indeed, the contemporary foreign language teaching is centred around the formation of key competences [1, 2, 3]. Yet, important gaps in the understanding of these competences still persist. It is unclear what competences are to be regarded as the key ones and how they are to be accommodated into the learning matter.

The present article will seek to address these gaps from the perspective of the multifaceted theoretical legacy of John Dewey, an outstanding proponent of philosophical pragmatism and a major reformer of the American educational system of the early XXth century. John Dewey laid the pragmatist foundations for educational reform, and his concerns about the potential inadequacies of the educational process gain particular relevance in the modern Ukrainian context. The basic point of departure of Dewey's educational philosophy is the need to integrate the learning matter into the "life continuum", or the social experience, of the students. The key contention of the present article is that a closer look at the relevant social experience may be helpful in

discovering the competences to be accentuated in the foreign language teaching.

Accordingly, the objective of this article is to build on selected elements of Dewey's educational philosophy in order to determine the specific bundle of competences and strategies underpinning the modern process of foreign language teaching. More specifically, the article elaborates on the importance of teaching approaches and strategies, including self-study and autonomous learning. The article proceeds by reconstructing some of these elements and subsequently by drawing implications for the competence development strategies in the foreign language teaching.

John Dewey's educational philosophy: a background

Dewey saw the philosophical meaning of education generally in what he called "renewal by transmission", i.e. the preservation and further development of societal civilizational attainments [4]. This societal problem is laid upon the office of education. It is through the education that people become effectively socialized and enabled to fully participate in the societal problem-solving process. Dewey warned, however, of the corrupting effects of rigid traditions and ceremonies that potentially convert the educational process from "renewal by transmission" into static and rigid systems of ceremonial domination. Dewey contrasted the natural plasticity of students' cognitive capacities with the potential rigidity of the educational process. He argued that the educational process must make students into effective problem-solvers rather than recreate the coercive ceremonial relationships existing in the broader society [5].

As a leading proponent of pragmatism, Dewey rejected the so-called dualisms of the idealist philosophy, such as those between body and spirit, matter and consciousness, facts and values. He explained them as reflections of excessive social stratification resulting in the disintegration of the social experience. In pedagogical terms, a relevant idealistic dualism is that between body and reason. According to Dewey, idealistically oriented teachers wish that their students bring their reason to a class but not their body, as the latter distracts them from learning. Dewey designated this kind of education as scholastic and remote from practical life, i.e., from the actual social experience. He argued that the really useful education appeals not merely to the reason, but to the

personality of the student as a whole. According to him, the students must be able to see the value of the learning matter from the standpoint of the whole social experience of which they continuously partake.

Dewey was particularly critical of all forms of learning by heart. He put forward the original idea that learning by heart is the kind of learning that destroys the ability for critical social inquiry and that is ideally suited for the preservation of the current patterns of ceremonial domination. Ceremonial societies are indeed generally known to favour learning by heart, but this kind of learning is clearly inappropriate for the purpose of making students into critical inquirers.

To Dewey, one of results of learning must be the creation of common interests and understandings among people, i.e., a better integration of the social experience. This is the reason why Dewey was wary of raising “egoistic specialists” and pleaded for general education that makes people more aware of the things they have in common with each other. Among these things, the language ranks foremost, and Dewey was explicit about it. Indeed, to take an extreme example, even if the language is used for conflict escalation, conflict participants still maintain a common interest in making their conflict strategies mutually understandable. Particularly the foreign language helps students to recognize their common interests with the diverse professional stakeholders from all over the world.

Implications for the study of foreign languages

Based on the above ideas, there are a number of general implications that can be drawn for the study of foreign languages.

- First, this study must occur within, rather than outside, the life continuum of the students. The students must be encouraged to bring to the class the diverse practical problems that they are facing, and be encouraged to explore how they can solve these problems by utilizing their foreign language competences. They must not be required to “forget” their problems for the time of the class.
- Second, it is crucial to make sure that the teaching strategy is free of any vestiges of ceremonial domination, primarily that of teachers over the students. The teacher must act not as the master, but as the facilitator of intrinsically driven inquiry.

The students must be given wide-ranging autonomy in the selection of the learning matter and learning strategies. As mentioned above, ceremonial domination often reveals itself in various forms of learning by heart. Speaking in pragmatist terms, it must be examined whether this learning has ceremonial or instrumental value in particular instances.

- Third, the interaction between teachers and students must be guided by their mutually recognizable common interests. Within these common interests, the students must be able to set their own learning goals that ultimately converge upon becoming effective citizens in the global and increasingly dynamic society. The common interests between teachers and students thus develop into a still broader interest of meaningful citizenship that is shared by a much broader community. This is how education contributes to the integration of the social experience in the sense of Dewey.

All of these implications have an important bearing on the basic pedagogical distinction between the learner-centred and teacher-centred teaching approaches. The former approach is traditionally associated with the learners' autonomy in setting learning goals and methods as well as with the extensive collaboration with the teacher in the development of learning schedules. This approach envisages the fullest use of the available learners' knowledge before the teacher starts explaining the new material. It requires taking seriously the learners' sociocultural and lifestyle habits in order to reflect on them from the international perspective. The emotional states and moral values of the learners likewise become important informational inputs into planning the teaching process. Overall, the learner-centred approach boils down to constraining the teacher domination of the learning process and to empowering the learner.

It is common knowledge that particularly the Western experts tend to overemphasize learner independence in the teaching process. They advocate learner autonomy and learner independence and argue for transferring the control over the teaching process from the teacher to the learner. The reasons for this standpoint basically include improving learning motivation and effectively appealing to learners' sense of responsibility. To us, it is obvious that this understanding of the learner-centred approach works best if the teaching occurs within the foreign

language environment and/or with adult learners who already have experience of responsible decision making.

In contrast, the teacher-centred approach supposedly contradicts the humanization of the educational process and constrains the learners' creativity. As a traditional teaching method, the teacher-centred approach involves frontal teacher-learner interaction with teachers usually nominating students for the assignment of specific tasks. Yet, many Ukrainian experts embrace the teacher-centred approach because of its ability to ensure continuity and conformity of the teaching process with the unified national standards.

Taking the Deweyian perspective, we believe that the dispute between the advocates of learner-centred and teacher-centred approaches is largely misplaced and that the best combination of these approaches is contingent on specific learning conditions. The basic condition that is shaping the foreign language teaching process today is the fact that, unlike a few decades ago, the professional environment of modern university graduates is rapidly changing in response to the ongoing processes of globalization and internationalization. This environment presents many opportunities of professional development, but it is a major source of risk as well. Apart from professional competence and language skills, competitive job seekers in the globalized labor markets must have cross-cultural competence, cultural awareness, and a meaningful sense of global cultural community. Developing this competence requires the full-scale cooperation between learners and teachers, with neither side dominating the other.

Practical implications

In line with Dewey's idea of the students' life continuum, we believe that the viable approach to creating the sociocultural competence is to maximize the use of authentic information in the teaching process. More specifically, this means that master students must be supplied with up-to-date training packages, including print and electronic media, and must have access to the web. Particularly the latter condition will enable the utilization of tremendous learning resources dispersed among a wide variety of global organizations and institutions. The high quality of teaching technology becomes even more imperative upon considering the trivial number of academic hours allocated to language study in non-language universities.

Several conditions must be met if the sociocultural aspect of language study is to be successful. Master students must be in possession of sufficient background competence and communicative skills obtainable during the bachelor study. The key aspect of master-level training is to build on the existing foundation of students' factual and lexical knowledge. Therefore, it is crucial that the master-level language study accommodates materials that have not been stripped of interesting factual and professional content. No less essential is the practice of considering grammatical schemes in the context of situational and contingent problem-solving. Class activities must be aimed at transferring information gained through self-study into the shared contextual space of the class.

Given the wide variation in the background competence among students, an individualized approach to teaching requires helping the students to develop independent abilities to search and process information on sociocultural topics. Due to their advanced training stage, master-level students are largely able to devise individual learning strategies. This ability is an important asset that potentially activates intrinsic motivation in studying a foreign language.

We follow Koryakovtseva in taking out-of-class activity to involve organization and implementation of students' cognitive activity without direct guidance on the part of the teacher [6]. This activity is advantageous in many respects. First, it enables time-saving optimization of the learning process. Second, it motivates students to engage in search behavior. Third, it makes for the creative and innovative character of the learning process. Given the significant amount of learning strategies outlined in the extant literature (cf. 62 strategies in Oxford (1990)), we lay emphasis on the formation of strategies of the following types: inductive/deductive inference strategies, comparative analysis strategies, summarization strategies, memory strategies (grouping, structuring, logical linking), compensation strategies (meaning guessing, constrained-language coping), social strategies (cooperation in learning, account of sociocultural traits of learning partners) [7]. All these strategies create a shared mental context for effective self-study.

Skills of independent learning allow master students to significantly enhance their professional competence. The competence thereby added is further reinforced through topical oral communication in the class. The receptive lexical and grammatical skills contribute to this process by effecting learning flexibility that engenders free student

discussions in various organizational settings, such as panel discussions and debates. Panel discussions are based on primary involvement of few students with advanced communicative skills. By contrast, debates involve the delivery of several oral reports featuring conflicting viewpoints that are subsequently discussed by the rest of students [8]. According to our experience, discussions have been particularly lively while considering the theme of 'Leadership styles'. Accordingly, we take discussion to be an advanced training method suitable for highly independent students advocating conflicting perspectives on an issue of common professional interest. Our understanding thus differs from those studies that consider discussion as a mere conversation with no reference to language competence. In our understanding, discussion helps students with various language competences to be meaningful parts of the shared learning space [9].

Concluding remarks

The competence-based approach in the foreign language teaching involves reorienting the teaching process toward the social experience continuum shared by both teachers and students and toward the development of student competences rather than the accumulation of passive knowledge. The basic practical strategy for implementing this approach is the integration of competence development into the teaching process. The work of John Dewey, an outstanding American philosopher and educator, helped to disentangle the challenges of the competence-based approach in a number of respects. The competence-based approach has been shown to most fully correspond to the Deweyian vision of the general societal meaning of education as "renewal by transmission". It is only within this approach that it becomes possible to effectively counteract the institutionalized patterns of ceremonial domination that exist in the broader society and are projected onto the classroom situation. Finally, the competence-based approach enables the most effective integration of what Dewey called the social experience across a number of divides, such as those between teachers and students and those between different cultures. The spanning of these divides is the major prerequisite for effective self-study that expands the students' learning horizons and creates a stimulating intellectual atmosphere in the classroom.

References

- Загальноєвропейські Рекомендації з мовної освіти: вивчення, викладання, оцінювання. – К.: Ленвіт, 2003.
- Програма з англійської мови для професійного спілкування. Колектив авторів: Г.Є. Бакаєва, О.А.Борисенко та ін. – К.: Ленвіт, 2005.
- Ніколаєва С.Ю. Цілі навчання іноземних мов в аспекті компетентнісного підходу // Іноземні мови. – 2010. - №2. – С.11-17.
- Dewey, J. (1921) *Democracy and Education: An Introduction to the Philosophy of Education*. New York: Macmillan.
- Dewey, J. (1922) *Human Nature and Conduct: An Introduction of SocialPhilosophy*. New York: Carlton House.
- Коряковцева Н. Ф. Современная методика организации самостоятельной работы изучающих иностранный язык. – М.: АРКТИ, 2002.
- Oxford, R. L. (1990) *Language Learning Strategies: What Every Teacher Should Know*. Boston: Heinle and Heinle Publishers.
- Jordan, R. R. (1997) *English for Academic Purposes: A Guide and Resource Book for Teachers*. Cambridge: Cambridge University Press.
- Ur, P. (1983) *Discussions that Work*. Cambridge: Cambridge University Press.

The Role of Gender Studies in Transforming Society

Cătălina-Daniela RĂDUCU^{1 2}

Abstract

Gender Studies represent a relatively new field in the area of humanistic and social sciences. Since their apparition, they generated major changes in academy and society; they are responsible, in part, for the way that social universe looks nowadays, continuing to shape its form, drawing attention to gender inequalities and designing ways to remedy injustice. The purpose of our study is to analyze the evolution of gender studies focusing on the way they contributed to the dramatic transformations of society in the last decades. We try to find, at the same time, some answers to the questions that this academic field must face in the light of the new challenges imposed both by its evolution and by the development of society.

Keywords:

gender studies, women's studies, society, gender relations, academism, activism, gender mainstreaming

¹ Cătălina-Daniela RĂDUCU - Postdoctoral Fellow, Romanian Academy, Iași Branch, Project *Knowledge-Based Society: Research, Debates, Perspectives*, POSDRU 56815, email address: cathy_raducu2001@yahoo.com

² „ACKNOWLEDGEMENT: This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU/89/1.5/S/56815”

Introduction

Gender studies represent, along many other disciplines in the area of social sciences, a new field, whose existence barely counts four decades. Despite their short existence, gender studies represent one of the most innovative and dynamic fields in this area. They constitute a very good example of the way in which the academic sphere has been transformed in the last half of century, through the acknowledgement of the fact that knowledge production is a historic, cultural and social phenomenon. Therefore, gender studies imposed, facing many difficulties, though, as we shall see in the present paper, the acknowledgement of gender as a category that intervenes at all levels and stages of the knowledge process, thus challenging the traditional view of knowledge production and contributing to the transformation of the traditional academic sphere. At the same time, gender studies through their contribution to a better understanding of gender construction and meaning, contributed to the transformation of social and cultural practices, thus transforming society in a durable and efficient way. Gender studies imposed, in the course of their short existence, gender as a fundamental category for the analysis of all forms and systems of human organization and social life.

The present paper intends to analyze the evolution of gender studies and the influence they had in the transformation of the academic sphere and of the society in general. We will draw, in the following pages, a short sketch of the evolution of gender studies in the last decades and will discuss a few of the challenges that this new field in social sciences has to confront with in the light of new developments both in the academic sphere and in the social life.

Gender studies evolution and impact on society

Gender studies appeared, for the first time, in the 1970, in American universities, in a time when the simple fact of saying that women can become an object of study was conceived as a radical act (Boxer, 1998, 10).

Initially, the field was called “Female Studies”, but this name was soon abandoned for the more comprehensive “Women’s Studies”, that highlighted the fact that the field contained studies *conceived by* women, *concerning* women and *belonging to* women. At the beginning of the 7th

decade of the last century, the foundations were set for a new field of study, in universities: “For the first time, women were not only learning about themselves, but were actively *creating* and *owning* knowledge based on their own personal and political experiences” (Ginsberg, 2008, 10). Women’s studies were born out of the protest that women in the universities directed against the way academic knowledge was silent about women and their scientific contribution and ignored the power relations in academia and society: „From its very inception, Women’s Studies had a very clear purpose and that was to transform the university so that knowledge about women was no longer invisible, marginalized, or made ‘other’.”(Ginsberg, 2008, 10-11).

The first days of women’s studies were not the brightest ones. Universities did not welcome courses and related activities: those who wanted to teach or hold seminars and attend meetings had no available spaces where to develop these activities: “It was not unusual for programs to form around meetings in bathrooms and broom closets.” (Ginsberg, 2008, 11). Lectures and courses were promoted through word of mouth or ways that did not involve high expenses, since there was no specific finance for these activities from the universities. Moreover, many of the scholars „worried that their association with the field would reflect poorly on them and derail their careers in their various disciplines (e.g., they would not get tenure)” (Ginsberg, 2008, 11). This was happening because those who held courses or gave lectures were not specialized in women’s studies, but in other disciplines and were doing this work in their spare time, as an additional activity to the one they were paid for by their universities. Their lack of specialization was often looked upon and the main argument against their acknowledgement as specialists in the field was that they could not teach a discipline that did not exist, and even if it existed, they could not teach a discipline that they did not know.

All these obstacles were meant to discourage those who aimed to open the universities and the universal knowledge as to include women and their scientific contribution and to transform society in order to balance the relations of power based on gender.

The first thing that the feminist scholars were criticizing was the apparently neutral character of scientific knowledge concealing the male domination in every field. Inspired by the women movements of their time, the feminists in universities grounded their undertaking on two epistemological premises. The first of them was *empiricism* – the will to

make women visible in all fields of knowledge, as subjects and agents of knowledge. The second one was *positionism (standpoint)* – according to which the experience, identity and knowledge of individuals are determined by their physical and social position, therefore the feminist phrase: *who I am is what and how I know*.

The papers published at that time witness the strife of women scholars to impose women as subjects and agents of knowledge in all fields of research and study and in the academic hierarchies, as well. For instance, in her article *Women's Studies – Renaissance or Revolution?* Adrienne Rich argues that women's studies not just broaden the curricula, they rather challenge the very disciplinary grounds of universities. She highlights the radical potential of women's studies in universities, considering them a 'pledge of resistance' (Rich, 1976). In *Toward a Woman-Centered University*, Adrienne Rich argues that universities encourage and maintain certain power relations that allow men to dominate women both in respect to knowledge production and in respect to their positions in universities (Rich, 1975). Thus, universities are built on hidden assumptions that keep alive a conception according to which men's work in academia becomes a sacred value justifying the emotional and economic exploitation of women. Women accede only to inferior positions in hierarchy, they are secretaries, teaching assistants, etc. while men get easier promotions on high positions where they are allowed, in the name of their sacred work, to benefit from the invisible work and merits of women. In other words, the men-centered structures of universities constantly reaffirm women's status as means for the ultimate goal of men's work (a goal translated into carriers and professional success, whose existence and availability are only male).

Despite all difficulties, the field grew stronger, especially thanks to the scholars, the students, the administrative staff or the sponsors who wanted to make visible the field of women's knowledge and that of knowledge about women. Women's studies were gradually introduced in universities, they began to be accepted and, as a consequence, to be financed by universities. Moreover, joining the feminist activists, the women scholars succeeded in obtaining national financed programs promoting gender equity in education at all levels and programs combating gender discrimination in schools.

Until the first half of 1980s, women's studies focus on explaining the reasons for women's subordination and oppression; starting with the second half of the 1980s, women's studies perspective changes radically.

Different women movements appear, such as the black women movement. These movements contest the centrality and the universal character of the notion of 'woman', and the field of women's studies has to identify and justify the use of this notion.

This happened because, although it was never explicitly defined as such, the concept of 'woman' used so far in the theorizations of the field contained characteristics that could have led to a new form of domination, this time inside the category of women in general. The studies developed until that moment focused on a restrictive understanding of the notion of 'woman'; they implicitly referred to the category that the women initiators of the women's studies belonged: white, middle-class, heterosexual, usually Christian etc. It was inevitable that the other women who did not belong to this category would feel left aside, unrepresented and it was also inevitable that they would make their voice heard, because their interests and needs were not addressed by the dominant perspective in the women's studies of their time.

The problem of the difference between women thus became extremely important in the development and evolution of gender studies. The claim for difference became a vivid subject of debate in gender studies over the 1980, the 1990 and in the 21st century, and made an important shift in the problematic of gender studies: the problem of the power relations inside the categories corresponding to each gender became as important as the initial problem concerning the power relations between genders. In the 1990s, as well as nowadays, women of different color, religion, social condition, sexual orientation claimed to be treated differently, according to their needs and interests. Therefore, gender studies, though still focusing on gender differences between women and men that explained the male domination and women oppression, begin to explore the power relations and the inner hierarchies of each gender.

Gender studies had to address two important challenges in the 1990s. Ironically, the first major challenge was imposed by the positive effects that the women movements, supported by the academic feminism have had for society. Since the results of the feminist struggle became visible (women and gender studies were already accepted in universities, women became visible in the academic space, women's and gender programs were financed by the universities, women rights programs and programs for combating discriminations were on a large scale financed by the states, legislation was modified by introducing

antidiscrimination and pro-equality measures, etc.), the germs of a new discourse entered the public sphere. The discourse continues nowadays and claims that since women's problem was solved, there is no need for a field of studies dedicated to them anymore. Of course, progress is undeniable but reality shows that until *de facto* equality between women and men is reached and until the various problems of the different groups of women are solved, there is still much work to do and this work is impossible in the absence of a specific field of study and research dedicated to them.

The second challenge of the 1990s was linked to the introduction of sexuality as a field of study and research. The scholars soon noticed that women's studies became a too restrictive area when faced with the necessities and interests of the different categories of individuals. Thus, 'gender studies' became the 'official' name of the field, in an attempt to include and address all those differences. The problem regarding which name is better for a field addressing women's issues, whether gender studies is a more appropriate name or not, was then and still is a subject of debate and we will approach it in the following pages. The challenge of introducing sexuality studies within the field of women's studies and the broadening of the scope of women's studies reflects the evolution of society in those times and is another reason to argue that gender studies were the result of the politics of their days and that they both reflected and influenced the evolution of society. For instance, in those times, sexuality studies addressed only questions related to the rights of the homosexual persons, but they gradually extended their scope to include studies and research concerning bisexual, trans-sexual or trans-gendered persons.

Gender studies in the 21st century witness a development that their initiations would not have dreamt of: at the beginning of the century, there were over 1500 gender studies programs in more than 700 universities, only in the United States (Howard, Allen, 2000, XXIII). Their development in the last three decades in the United States was easier thanks to the flexibility of universities in respect to curricular changes, especially in the newest universities, where 'tradition' did not action as an obstacle. The United States functioned as a model for many countries from all over the world: „very soon there were women's studies programs in Japan, Australia, New Zealand, Finland, Sweden, India, South Korea, Taiwan, and the Philippines. By the 1980s, there were programs in all countries in Western Europe, plus Thailand, South

Africa, China, the Caribbean, and Uganda. Finally, after the change from communism in Eastern Europe, programs were instituted in Slovakia, the Czech Republic, Russia, Ukraine, and others, in addition to Malaysia, Vietnam, and other African nations.” (McFadden, 2005) In Romania as well, gender studies were introduced after the fall of communism, in the 1990s, and they developed relatively fast: now, there are offers of gender studies programs in the main universities, especially at the M.A. and graduate level (a discussion on gender studies in Romania and the way they influenced the Romanian society can be found in Răducu, 2010).

Gender studies development continues, simultaneously with the development of society: now gender studies are broadly available, thanks to the introducing of computers and internet. The possibility to teach gender studies courses and to have discussions on-line allows a greater opening of the field to the population and to the community. Those persons that did not have the chance to follow gender studies programs in universities now have at their disposal a virtual space for collaborative learning, for dialogue and for sharing resources with a minimum effort. But perhaps the greatest advantage is that on-line gender studies now get out of the rigid academic sphere, re-shaping the relation between the student and the professor and re-shaping communities also, helping people interact and giving a voice to those that could not speak before because of the lack of resources or because of spatial, temporal, cultural or social limits.

We can say, therefore, that from the beginning of the 1970s until now gender studies had an essential impact on the transformation of society. We can say now, when we look behind, that the academic and social space would not have been the same without gender studies and the social movements that grounded them.

Ever since their beginning, gender studies established the purpose to transform, besides the academic sphere, the social universe. They appeared as a natural result of the politics of their time, especially the women movements, and were designed to support and enforce the results of the women’s movement: „From the beginning, the goal of women’s studies was not merely to study women’s position in the world but to change it” (Boxer, 1998, 13). The pioneers of gender studies assumed since the creation of the field the mission to study women and other groups with a view to obtain „positive social change in the rest of the world for women and other oppressed groups alike” (Ginsberg, 2008, 10). By imposing the acceptance of women as subject of study and

the acknowledgment of women's contribution to the scientific knowledge, by helping introducing anti-discrimination legislation and gender equality policies in education, by supporting women's and other groups' movements, gender studies accomplished their initial goal and proved their influence to the transformation of our society.

Present challenges for gender studies

However, in spite of the undeniable results mentioned above, there still are theoretical, methodological and practical challenges that gender studies have to face. Although the field consolidated its position in the academic sphere, it has to offer answers to a series of questions that appeared in the course of its development, such as:

Questions concerning the statute of gender studies: Are gender studies a discipline or rather they are an interdisciplinary field that includes different approaches concerning all the implications of gender as a social construct?

Questions concerning the way gender studies should be taught in universities: Should they be taught as separate courses or rather gender problematic should be mainstreamed as broadly as possible in all disciplines, so that a true sensitivity to gender is obtained? If they remain separate from the other disciplines, are they not they risking closing themselves in a marginal area of the academic sphere? Could gender mainstreaming lead to a change of perspective and a real acceptance of the gender problematic, given the suspicion with which gender is still treated in traditional disciplines?

Questions related to the relation between academism and activism: Are gender studies a 'bridge' between the academic discourse and the reality they investigate? Are there still a 'pledge of resistance', do they still have a real influence over society or rather they are positioned in a rigid academic sphere, where educated and usually well socially and economically positioned persons change general and technical opinions, distanced from the real subjects of their theories?

Finally, *questions related to the very nature of the field's name:* Is it possible that under the generic phrase of 'gender studies' the initiators'

message and goals were somehow attenuated, dissolved? Is it possible that including women's studies, sexuality studies, men's studies and other researches under the generous 'umbrella' of gender studies the importance of gender relations is minimized? Is it possible for a kind of disinterest to hide under the 'polite' name of 'gender studies'?

We will begin our discussion with the last debate. In the 1990s, under the double pressure imposed both by the limitation of a universalistic approach of the women's problems (which would have hidden in fact a real domination of a group of women over other groups of women), and by the introduction of other types of studies than those strictly related to the women-men relations (such as homosexuality or trans-sexuality), the initially well-formed field of women's studies had to accept a broadening of its scope and approaches. Starting that moment, an infinite series of debates concerning the proper name for the field was initiated; the debates continue today. The programs and courses in universities received new names; alternative denominations appeared – *feminine studies, feminist studies, gender studies, sexuality studies, gender and feminist studies, feminist, gender and sexuality studies*, etc. Many programs added the phrase 'gender studies' to their initial denomination, others simply abandoned the initial denomination of 'women's studies' and adopted that of 'gender studies'.

The main challenge is related to the subject matter of the field and to the concern that the inclusion of women problematic in the larger field of 'gender studies' could lead to an 'abandon' of the aims and efforts of the feminist movement that set the grounds for and supported these studies along their short and troubled history. The most frequent formulated concerns relate to the fear that broadening the scope could also lead to a dilution of the preoccupation with women's status and with gender power, a dilution that would have negative consequences for women in general. Eliminating women from the name of the field could lead to an abandon of their specific problematic, a loss of contact with the women movements and a loss of the advantages earned along decades of women movements.

Although we admit that it is difficult to formulate a point of view in this matter, as difficult as it is to give a 'verdict' referring to the name of the field, our opinion is that the term 'gender' in itself is not so 'guilty' as it has been considered by those who feared that women's problematic would be dropped in a second or third place of academic or scientific interest. We believe that the term 'gender' indicates a transition from the

feminist approaches treating women as an isolated group, to more complex approaches concerning the relations between and within genders. We believe that the fact that gender studies include now men's and masculinity studies, and studies concerning the relations the two gender developed historically and continue to develop in new cultural, social, economic contexts is not, in its essence, threatening for women.

It is a new approach concerning both women's and men's specific ways of thinking and communicating in their social interactions as well as a new approach concerning the roles that each of them might play in the public and private spheres. In our opinion the use of the term 'gender' signals both the awareness of the fact that gender identities, roles and relations have a certain cultural specificity and the awareness that inequalities between and within genders is, eventually, the result of a social process that can be contested and must not be not taken as a biological given.

Finally, we think that the phrase 'gender studies' can be used for a generic field including specific areas of study. These range from *women's studies* (with a clear focus on women's problematic), to *gender equality studies* (with insistence on historic and social inequalities and on the measures needed to correct injustices suffered by individuals or groups of individuals), *sexuality studies* (reflecting the problematic of persons with different sexual orientations, transsexual or trans-gendered persons). *Men's studies* should also be included here, as most recent approaches show that there are men's problems deserving a separate space of study, such as power relations inside the men's group that need to be addressed in a specific manner, etc.

This would not lead, we think, to a decrease in interest for women's problems, except in the situation when the women themselves would wish this to happen. The history of the field shows us that when social reality must be changed, because it disadvantages a category of individuals, there are ways and means to do it. Moreover, nowadays, when much of the way has been covered, women do not have reasons to consider themselves threatened by a simple denomination. The reality shows us that it is in their power to continue the road. Therefore, we believe we have reasons to be optimistic in this matter.

Besides the concern about a dilution of women or other disadvantaged groups movements' political message, there is a justified concern that institutionalizing gender studies could lead to a withdrawal in a 'ivory tower' of the gender studies scholars and to an abandon of the

pioneers' aim to build a bridge between the academic and social spheres. This is, as we have shown elsewhere (Răducu, 2010), a genuine reason of concern in post-socialist countries, where, for instance, formal equality was intensely promoted for decades by the communist discourse, although in substance it never existed. In these situations, the voice of non-governmental organizations is weak, the politicians do not assume the theme of gender equality in their programs and the gender scholars rarely communicate in an efficient manner with the activists.

However, gender studies assumed since their appearance and continue to assume today their mediating role between theory and activism, personal and political, academy and community. By providing the students with the theoretical and methodological instruments for analyzing gender relations, by encouraging them to apply these critical analyses to the family, the group or community they belong to, gender studies do not lack the activist part. In this matter, as well as in the matter concerning the name of the field, gender studies could detach from the society they intended to transform from the very beginning only in the unfortunate case when the gender studies scholars would wish to withdraw in a theoretical sphere of a rigid discourse over distant realities. But our opinion in this matter, is optimistic again.

Trying to answer the second problem indicated, we will answer, indirectly, to the first problem, concerning the statute of gender studies. Thus, gender and gender relations can be studied as well in the frame of a discipline in itself, as in the frame of a research area and a complex, multidisciplinary and interdisciplinary field of study. Moreover, gender studies, as we have seen in the first pages of our paper, generated transformations in the traditional academic disciplines, developing new perspectives and approaches within those disciplines. Since gender studies evolved in an independent field of research and study with implications over other disciplines and fields of study, we think that a bi-directional approach of this question could bring the most benefits, both to the gender studies and to the academic sphere in general.

Thus, our opinion is that gender studies need to continue to develop as a separate field of study, because they build the theoretical framework in which gender relation research takes place. Also, because they are still regarded with suspicion and tend to be 'pushed' in a marginal area in the corpus of academic disciplines and fields of research, it is equally important to mainstream a gender perspective in other disciplines. Human knowledge is historically and socially

determined; therefore, transgressing the borders between disciplines is most of the times inevitable. Gender, too, is implied in all forms of human organization and social life, and becomes more and more an 'inevitable' element in scientific research, especially in social sciences.

Conclusion

In the light of all the things shown in this study, we have reasons to believe that irrespective of theoretical, methodological and practical challenges they meet, gender studies are a solid field of research and study, in a continuous development and transformation and their evolution on a long term in the academic sphere is granted by the transformations that are taking place in society on a global scale. We may thus speak, at this moment, of a bi-directional transformation: from the gender studies towards society and from the society towards gender studies. Therefore, the interdependent relation between gender studies and society can generate benefits both for the academic field of gender studies and for the social reality they investigate and permanently seek to transform, with a view to re-balance the power relations between and within genders and to remedy injustice where it appears or is maintained by social practices.

References:

- Boxer, M. (1998) *When Women Ask the Questions: Creating Women's Studies in America*. Washington. DC: Johns Hopkins Press
- Ginsberg, A. E. (2008) Triumphs, Controversies, and Change: Women's Studies 1970s to the Twenty-First Century. In Ginsberg, A. E. (ed.). (2008). *The Evolution of American Women's Studies. Reflections on Triumphs, Controversies, and Change* (pp. 9-37). Palgrave: Macmillan.
- Howard J.A., & Allen C. (2000) Feminisms at a Millenium. *Signs: Journal of Women in Culture and Society*. vol. 25, XIII-XXX.
- Răducu, C. (2010) Studiile de gen și impactul lor asupra societății românești de după 1989. In Ana Gugiuaman (ed.) *Idei și valori perene în științele socio-umane* (pp. 103-113). Editura Argonaut: Cluj-Napoca.

- Rich, A. (1976) Women`s Studies – Renaissance or Revolution?. In *Women's Studies* 3, 2, 35-47.
- Rich, A. (1975) Toward a Woman-Centered University. In Rich, A. (ed.). *Women and the Power to Change* (pp.15-46). Florence Howe, New York: McGraw Hill.

Research Studies

Analysis of the Educational Statistics of the Sudanese Secondary Certificate

A.W. Mohamed ISSAM ¹

Abstract

In the past two decades many voices were raised calling for corrections, restructuring and improvements of the Sudanese educational system. Complaints focused on the weakness of university students. However, their weaknesses were traced back to the elementary, secondary and high secondary educational phases. Economic restraints, lack of facilities and meager numbers of qualified instructors were pointed out as the prime causes. However, in this paper we hypothesize that besides those reasons basic impediments were entrenched in the whole educational system. The metamorphosis the Sudanese educational system has foreseen in the early 1970s and 1990s have distorted students knowledge base. Changing the system into purely Arabic based and partially abolishing English language weakened a well-known and strong system Sudan has known for generations. The current paper analyzes those effects and review some parameters to contain them. It is concluded here that interventions by lowering the students' pass mark rates have affected their knowledge and education capabilities. That gave them false sense of passing from an educational phase to another. However, the result is worse knowledge transferred into another educational phase and then to the university. The result is like Zeno or Achilles paradox and a worse graduate education.

Keywords:

Sudan, Education, Metamorphosis, Knowledge Transfer, Pass Marks, Weakened Structure, Destabilization.

¹Professor Dr. Mohamed A. W. ISSAM - Professor of Economics, Alneelain University, Khartoum-Sudan. P.O. Box 12910-11111, Email address: issamawmohamed@yahoo.com

Introduction

Our administrative environment rarely recognizes the impact of statistical analyses and findings as a measure of importance in designing and implementing policies. In a way, this is mainly due to either of the lack of awareness of its usefulness and importance, or the fear of its ability to uncover facts and hidden failures. In fact it would not be strange to notice that policies are designed or decisions taken depending on rough estimates or upon the hunch of the administrative leaders. Decisions about complex issues are sometimes, taken on subjective grounds or on what the officials think is right rather than adhering to any dependable scientific reasoning. This kind of culture, naturally, breeds disregard to data collection and data analysis. Officials working in the fields of statistics in various ministries and departments are often left on the margins. Their meager efforts are in most cases, used to decorate the official reports to give them some essence of complexity. Under these circumstances, not many candidates would, willingly be ready to get into the profession of statistics. It is a field that is hard to study and that does not secure a defined and rewarding future. It would be hard to persuade a good student to choose statistics as a major for it would not be easy to give a convincing answer to the important and continually asked question: where would we find employment, and what could be the nature of my job, if I got any? This obscurity leads to a rarity of qualified statisticians. Of these, just a few stay to work at home. Those of them who are more experienced and qualified would rather leave the country to work somewhere else where the nature of their services can be gainfully utilized. The research community usually finds it hard to conduct conclusive research as the kind and quality of data available is marked with so many shortcomings. Most of the data available, with the exception of that of the population censuses, is outdated, limited in scope and coverage, lacks replication in time and so represents only a single point of time. Annual statistical reports that used to be, issued by some government ministries are, in some cases, marked with chronic deficiencies like:

- Data Inconsistencies, e.g., consecutive annual reports are marred with inconsistencies and errors.

- Abrupt changes: contents and coverage of reports change as years pass by. That renders it hard to create a researchable time series about variables of interest.
- Lack of a strategy: composition of contents of those reports does not seem to follow any pre-designed strategy.

In fact, those deficiencies are a natural by-product of a civil service culture that does not have a heritage of promoting policy-oriented research or comprehending statistical analysis. The aforementioned annual reports are usually compiled by employees who are not trained as statisticians and so are naturally, unaware of the necessity of precision or the role a good package of data could play. Besides, lack of policy-oriented research lead to indifference to the substance and quality of the annual reports and to the lack strategy in their compilation. A research- promoting environment would lead to continued monitoring and development of dependable and expanded data systems and, thus, produce precise and improved annual reports. This would lead to a perception and acknowledgment of the services a statistician could provide and a clear definition of his role and status. Improvements of this sort would tempt quality youngsters to profess statistics as a career with an anticipated positive impact of energizing the research community. An energetic and active research environment would surely play an active role in minimizing the random and erratic nature of our policy and so optimization of use of the limited resources and of the time needed to achieve results. That is the path that should be followed by countries if they intend to flourish and be compatible with the age of globalization and use the means of the twenty-first century.

As previously mentioned, a key to real social and economic development is the optimal use of available resources. The most important of these resources is the human factor, which is the target of economic development. Precise and updated multidimensional information about the varying aspects and activities of this human factor should furnish the basis of prime importance when laying-out policies, tracing and evaluating their impact. It is the cycle of prior and posterior research: Creating data systems to enhance research; then research findings and accruing needs will eventually make it necessary to develop better and more varied data systems.

This research aims at discussing an important element of the human development, i.e., Education. Educators, as part of the official

aforementioned culture is universally said to be indifferent to research results about the impact of varying education policies. They mostly base their judgments on personal experience, and had strong views about what worked and what did not. In fact, if policies are designed to improve the education system by optimizing the use of the available resources, those policies should be formulated after ample research based on adequate and dependable statistical data. Statistical analysis is used in a descriptive format; that is preliminary description of the current situation and of progress throughout a given period. Advanced research should dig deeper into the available data to relate different variables and create models conditioned that they should use the proper hypothesis (Akaike 1974). Research is needed to assess policy issues like the continuing efforts to:

- Improve the performance of education
- Enhance equality of opportunity of education.
- Find out and analyze the sources of discrepancies in student academic achievement and the effects of diversity within and between various communities of the society.
- Analyze the problem of dropouts at varying levels of education, measure the frequency of its occurrence and see whether such variables as family background, regions and gender of the students are of significant effect on it

The World Conference on Education for All (EFA), which was held in 1990 at the city of Jomtien in Thailand, agreed to *universalize* primary education and massively reduce illiteracy before the end of the decade². To answer the question of how well countries have delivered on their commitment to provide basic education to their people. The EFA forum launched a global assessment of the progress made towards the goal of education for all. The exercise's purpose was not just to measure success and failure but also to *reinforce the dialogue on basic learning needs in order to develop strategies for the twenty-first century*. In December 1999, delegates of 47 countries in sub-Saharan Africa met in Johannesburg to assess progress made toward the goals of the Jomtien Declaration on Education for All (EFA) and to set the agenda of the twenty-first century. The EFA 2000 assessment was launched to find out what has been achieved so far and to analyze what has enabled and obstructed

² Assessment of Basic Education in Sub-Saharan Africa (1990-2000), EFA 2000.

progress toward the goals of Jomtien. The assessment reports encompassed 18 statistical indicators and policy reviews of basic learning needs at early childhood, primary school, youth, and adulthood.

Progress in all areas is often made through learning from past errors and successes. The EFA assessment provided useful lessons as to where progress can be or has been, achieved. Some of these are:

- Improve the management of complex programs: Jomtien goals proved to be complex to implement, particularly those aimed at improving institutional management, organizing partnerships between agencies and quality and effectiveness of learning process
- Concentrate on activities that have a good chance of being successful. Programs that have realistic goals, based on earlier successes and failures, were emphasized to avoid wastage of time, energy, and resources.
- Avoid fragmentation of activities. The holistic aspect of basic education in a lifelong learning perspective needs to be kept intact while distinctive elements are simultaneously defined.
- Reinforce partnerships: The diversity of those involved in improving basic education is an unprecedented asset in Africa. It is essential, however, that support to basic education be efficiently consolidated and managed in order to ensure participation of all major stakeholders at various levels of society.
- Take into account technical economic and institutional capacities: Social, technical, political and economic environments have had a profound impact on basic education achievements. Education systems need to adapt to changing conditions so that they continue to function as well as possible and to teach children skills and knowledge that are directly relevant to development in these areas.
- Develop policies on illiteracy, poverty, and population. Correlation analysis have shown a significant relationship between illiteracy, poverty, population growth and educational achievements.
- To develop policy-relevant information systems: Relevant and efficient information systems based on statistics, research

and analysis, which are essential for sound policy and management of the whole education sector.

So, if educational policies are mainly designed to increase the productivity of the educational system and improve the quality and extent of the academic achievement of the student. However, the question remains as to what extent do our educational policies comply with the aforementioned policy targets of the EFA? A historical narration of the developments in our education system is summarized as follows: Richer countries achieve better results because they have more resources for their education systems. In poor countries funds are mostly consumed by teacher salaries. So if extra funds are available, they are allocated to expand coverage to meet the high demographic burden and to areas of the country that are short of the service. The case in Sudan is typical of that. Up to the fifties, education opportunities were very limited and restricted to some urban areas.

The probability of a child getting even a primary education was so low to the point of being negligible. Then, the education system was engineered to meet the needs of the government for civil servants of various levels and specializations. The quality of education was high and extra incentives were provided to persuade enrollment of students. As of the mid fifties, when the country gained independence, there was a change of policy. The objective was to expand coverage while preserving the quality of education. Class size was controlled, student teacher ratio was kept at a reasonable level, books and other instructional materials were provided, salaries were satisfactory and teacher training was a priority.

At the beginning of the seventies a drastic change in politics occurred. Expansion of education became a priority and was considered a hot political issue, which, resulted in drastic increases in numbers of public schools of various levels distributed all over the country. Schools were opened wherever the local citizens were willing to contribute to build them. In most cases, the locals were poor and so schools were built without paying much attention to standard specification. The limited resources allotted by the government were not up to the challenge and that lead to a gradual deterioration of quality of education and school environment. Besides, at those critical times the temptation of high paying jobs at the Gulf States drained the education system of its experienced and trained teachers. Those factors, added to the impact of

rising inflation rates, resulted in a highly noticeable negative impact on the whole of the education system.

In the nineties a second "*Education revolution was declared*". Coverage was expanded to measures unknown before. To minimize government expenditure that should have accrued due to that expansion, most of the cost burden was transferred to the parents. Services that were previously being granted, like boarding houses, were no longer made available. That was a serious policy decision, which was incompatible with the principle of equal opportunity education and with the special needs and living conditions of some communities of the country. Besides from that, the school system was decentralized and the cost burden of its dramatically slashed budgets was transferred to the regions. Political priorities and revenues vary from region to region and so does the attention to the school system. School conditions, which were already bad, got worse. Many regions could not even afford to pay teachers their meager salaries. We began hearing about teachers striking or leaving schools to try to earn a living somewhere else. In such conditions of worsening school environment and low teacher moral, it would not be strange to find a good number of illiterates among pupils with over 6 years of schooling. That is a complex issue. Demand for education is high all over the country and it has to be met somehow. The complexity arises when trying to set priorities for allocating the limited resources of a poor country with varying needs. For the education sector, the choices available were to expand coverage with the aim of giving a wider range of youngsters some education, retain quality at the expense of restricting access to education and retains a certain measure of both. As an ad-hoc result of those policies, the choice was limited to the first option. Other policy issues included changes in the duration of different levels of schooling. It used to be four years for each of the primary, intermediate and secondary levels of schools. Then duration of primary was increased to six years, leaving intermediate and secondary levels with three years for each. A third change was applied lately resulting in the abolition of the intermediate level altogether, extension of primary to eight years, while leaving secondary with three years as before ; thus slashing one year off the length of duration of basic schooling. Such changes are dominantly accompanied by changes in curriculum and didactic material that were sometimes entered not solely for educative improvements, but rather to comply with certain political reasoning. One can be certain that none of the aforementioned policies,

or changes in policies, was ever subjected to proper scientific analysis to justify their implementation, or to verify the impact of that implementation. There is neither the will nor the ability to undertake and carry out genuine research. The quality and coverage of statistical data available does not help researchers who want to try.

The primary purpose here is to try to lay out grounds for this proper scientific analysis, and the role statisticians and statistical analysis can play in this context. It is important here to underline a basic fact: educators should be the prime researchers in their field. They are the only ones who know best about the student and school environment. They are the ones who, normally, should lay out the grounds for changes in educational policies, and assess the impact of their implementation. Thus, naturally, they will be the ones most qualified to decide if research is needed, set down its objectives and areas of coverage. The role of the statistician emanates as a helping force. He should assume his role in close co-ordination with the educational researcher and help him achieve the planned objectives of the research³. Statistics Myth 2 assumes that the researcher and the statistician are two different and unrelated persons, and that forethought as to what statistical procedures utilized to analyze the data from a study is unnecessary. This myth causes researchers to settle for less sophisticated data analytic methods and/or do a poor job of interpreting the results of the methods they use. Additionally, limited knowledge of basic statistical concepts can also lead to inappropriate interpretations of statistical results. Chief among these misinterpretations of statistical results is the common misunderstanding of statistical significance testing, the assumption that a statistically significant result is necessarily a noteworthy result. The role of a statistician is not just to labor in producing annual statistical reports that rarely anybody cares about or understands and which are usually out of date. That should integrate his role within the system to be part of ongoing research activity or pioneer in initiating that research activity.

In a widely publicized speech⁴, the Vice President of the Republic addressed a meeting held on the mid of April 2003 at the Ministry of Education. In that speech important questions were raised which are summarized as:

³Larry G. Daniel (1996) Kerlinger's Research Myths, ERIC Clearinghouse on Assessment and Evaluation.

⁴ Al- Anbaa Newspaper, April 22, 2003

- What are the educational indicators of last year? What are the highlights the conference of education and whether its findings and recommendations duly embodied into the educational strategy?
- What are the results that accrued from the field visits and workshops that were held?
- To what extent have we moved towards achieving the prime goal of mandating primary (basic) education, and what measures are being taken by the states toward that goal' and if help is needed from the central government ?
- What about the prerequisites like: student seating, school environment, textbooks, teachers and teacher training?
- Concentrating on the subjects of the availability of textbooks and teacher training. He targeted teacher training as an issue of prime importance and enquired whether they are problems in that area and whether those problems originated solely from financial reasons. If that financial problem is solved, what about other issues like: availability of the qualified training staff, whether the school system can afford giving full-time leaves for its prospective trainee teachers to attend fulltime at the training centers, and whether these training centers (supposedly 73) are sufficiently qualified to undertake that job.
- Whether policies are laid out to deal with problems like dropouts of students from the educational system, adult illiteracy and vocational training?
- Would the ongoing educational process help in building up a sense of national identity, national unity and cohesion?
- If there is a policy of promoting non-instructional activities like sports and other student activities, or whether the instructional course burden is so heavy that it leaves no space available for such activities.
- If there is an ideal course curriculum, should it be standardized all over the country, or allowed to differ to accommodate for differences between the different local communities.

- If exams should remain to be the only means of assessing student attainment or should we look for other alternatives.

In fact, that speech was remarkable as it addressed major educational issues and contained so many exclamation marks that needed answers. Such questions should not be answered in a hurry for educators may assume they know the answers. Those questions should be properly answered not by adhering to the meager statistical data available. But they need to be properly dealt with through serious and proper scientific research, to accumulate knowledge about the real facts of life all over the country and build strategies in accordance with that knowledge.

Education is one of the most complex topics of social science. Its complexity arises from the variety and multidimensionality of factors affecting the educational processes and influencing the students' academic achievement. According to MacKinnon (1996) this complexity places great demand on the scope and quality of data, the sophistication of model specifications and content, and the accuracy of the estimation techniques.

Three research approaches will be necessary to lead to reliable research findings and to effectively-guide educational policies:

- Introducing experimentation,
- Building theories of educational processes,
- Improving non-experimental data and its analysis.

Fromm and Taubman (1968) conclude that: If experiments are well designed, implemented, analyzed and replicated, they will provide explanations that are close to casual as possible in social science. Such experiments can provide the most accurate results for the effect of a particular variable in a given context. Experiments can also play another, and perhaps more important, role in social science research-namely, helping to validate model specifications for non-experimental data. Experimentation and data collection need to be directed toward both the building of theories and the improvement of our assumptions in analyzing non-experimental data. In the long run, policy analysis will largely be dependent on improving non-experimental analysis since experiments can never be counted upon to solve all the complex and contextual effects present in education”.⁵

⁵ Moving Educational Research Toward Scientific Consensus, (1990) David W. Grissmer& Ann Flanagan. RAND.

We proceed by discussing the important issues of educational indicators and of modeling the educational processes with special reference to the students' educational attainment as a variable being influenced by a multiplicity of complex factors.

Models

The model is primarily used to make inferences about differences among what is termed as treatments or factor level effects. A regression model relates a response variable, Y , to an independent variable, X , or a number of such independent variables using the relationship:

$$Y = \beta_0 + \beta_1 X_1 + \xi_1$$

where the parameters β are the regression coefficients that specify the nature of the relationship and E the random error. When using and analyzing multiple it is important the proper interpretation and inferences for its partial coefficients, reflecting the contribution of particular coefficients over and above those made by others in the model.

A model is an imitation of a real world system or process. The effect of changing certain inputs can be studied before a decision is taken to implement them in the real world. "To build a model of a system or a process, a set of mathematical or logical assumptions about how it works needs to be developed. The complexity of a model is determined by the complexity of the relationships between the various model parameters.

In order to produce the model and determine suitable parameters, data needs to be considered and judgments need to be made as to the relevance of data".⁶ Models are used for predicting the effects on the response (or dependent) variable when alterations are entered on any of the independent variables when their significant contributions are confirmed by the analysis.

⁶ Stochastic Modeling , (2001) Faculty of Actuaries (Textbook).

Indicators

Indicators in the educational system serve similar purposes to those that are used to monitor the economy, population growth, health problems or other social phenomena. Statistical⁷ indicators are used to monitor complex conditions that we would probably judge imprecisely or miss altogether in day-to-day observations. The overriding purpose of indicators is to characterize a system through components-how they are related and how they change over time. This information can then be used to judge progress toward some goal or standard, against some past benchmark, or by comparison with data from some other institution or country. Thus, an indicator is a statistic that, if calculated from a sample data, can be used as a good estimator of a population parameter. Jaeger⁸ concluded that indicators are anything but clear and consistent. Review of a dozen definitions has produced much that is contradictory and little that is concise and illuminating (P. 285)".

He recommended that: all variables that:

(1) Represent the aggregate status or change in status of any group of persons, objects, institutions or elements under study and that

(2) Are essential to a report of status or change of status of any group of the entities under study or to the understanding of the condition of the entities under study, should be termed indicators. "I would not require that reports of status or change of status be in quantitative form, for narrative is often a better aid to comprehension and understanding of phenomena than is numeric report "(pp. 285-287).

An education system can be conceived as having underlying properties that are not directly measurable. We can talk about the quality of the teaching force but cannot find a direct way to measure it. At best, several statistics can be combined into an indicator that gets at our notion of teacher quality. An indicator of teacher quality might be some aggregate of years of academic training in the discipline taught; possession (or lack of) a credential in the subject matter taught; years of experience and background and so on.

7.) Shavelson, Richard J., Lorraine, M. McDonnell & Jeannie Oakes, (1991) What are educational indicators and indicator systems, ERIC Clearinghouse on Assessment and Evaluation -

⁸ Jaeger, R. (1978) About educational indicators. In L.S. Shulman {ED.}, Review of Research in Education

Education indicators are statistics that reflect important aspects of the education system, but not all statistics about education are indicators. Statistics qualify as indicators only if they serve as yardsticks. They should qualify to be informative about the current status of the whole system. Numbers of schools being opened, or numbers of students enrolled or of those admitted to higher levels of education are informative but they do not really give needed information about the sufficiency of the education system. Proportions of enrollment or admittance to higher levels of education relative to the relevant population of applicants will do the trick. So an indicator can be defined as “an individual or composite statistic that relates to a basic construct in education and is useful in a policy context”.⁹

INDICATOR SYSTEMS

A single indicator cannot always give efficient and sufficient information about an education phenomenon. Indicator systems are designed to attain the purpose of generating more expressive and accurate information about ongoing conditions. These systems are not just a collection of indicator statistics, but they measure distinct components of the system and provide information about how the individual components work together to produce the overall effect. Indicators at the national level should in general, represent the important components of the educational system. “In addition to monitoring outcomes, indicators should reflect the characteristics of students and communities served by schools, the financial and human resources (especially teachers) available to the schools, and other educational inputs. Moreover, they should reflect the adequacy of the curriculum and instruction received by students, the nature of the school as an organization in pursuit of educational excellence and equity, and other educational processes. Finally, indicators must be related to one another so that their relationships, and changes in these relationships, can be ascertained to suggest possible explanations for observed changes in outcomes.”¹⁰

Expectations from social indicators are generally modest. They describe and state problems more clearly, signal new problems more

⁹ Ibid

¹⁰ Ibid

quickly, obtain clues about promising educational programs, and the like. But, they can not substitute for well-designed, in-depth evaluations of social programs (Gagné, and Ouellette 1998). They do not, however, provide the level of rigor or detail necessary, and, most importantly, they do not give us any scientifically verifiable information about the variant variables that are contained in them. If the main issue behind educational policies is to end up with the target of improving the academic achievement of the student or his education outcome, then inequality in that achievement needs to be studied. Inequality in family background is probably the key to inequality in educational outcomes. Then inequality in aggregate family background ought to be the key to understanding differences in achievement means (indicators) between states, or between areas or groups within the state (McFadden, 1984). On the other hand, if inequality in school quality is the key to understanding inequality in individual outcomes, then aggregate school quality ought to explain variation between and within the states. That leads to an important deficiency in our system, lack of data. For in order to run in-depth studies to try and find out verifiable answers to these questions about the reasons of inequality, data is needed about the school and about the socio-economic background of the student. Availability of such detailed and dependable data will render it possible to build up models to find out and examine the impact of various factors on the student achievement. A macro level model containing the factors of significant effects on the student outcome at the national level is useful but not enough. There are no reasons to assume existence of homogeneity between and within the states, unless it has been proved. The effects of these factors may not be the same in all states or regions. Other factors, specific to certain states may emerge, and the variable composition of the models may differ. So models will be more useful if they were built for different states, (or even regions) to suit and contain variation of conditions at the local level. This may be likened to” a meta analysis” in which each state’s data provide an independent study of the varying correlates. Here the need is to examine the student’s socio-economic backgrounds, and home educational resources as predictors of student proficiency. Yet our models simultaneously include indicators of educational quality: class size, pupil-teacher ratio, school climate, course-taking opportunities, teacher qualifications, teacher’s experience, cognitive stimulation in the classroom, and so on. Our findings, if

reasonably consistent across states, will show both of the “the home effects” and “the schooling effects” explanations.

Perhaps more importantly, the analysis within the states bears more directly on controversies surrounding accountability and kind of priority given to the school system at the state level. An important question arises here: What if we controlled the indicators of student background and educational quality, will the state variation vanish? We assume it will be because we are speaking here about an average not individual student outcome.

This hypothesis makes sense, in that state- level policies (regulations, incentives, aid, etc.) can presumably affect student outcomes only through affecting specific educational resources and practices at a more local level, i.e., within schools and classrooms. If those local practices were fully controlled in our models, then there would be no direct role for state policy to affect student achievement.

This is a hypothesis that is to be tested when the data is rightly available. However, for now we can assume that state differences in average student outcomes, almost entirely, reflect variation in measurable aspects of student background and school quality.

If we define a state’s average student outcomes as being the average scores the state’s students attain in various papers or subjects of a national-level exam, then student background and school quality will be our focal correlates of proficiency. They are the correlations, if statistically verified should be the ones that should be the focus of manipulation by educational policies. It is important to investigate state-to-state variation in the provision of key educational recourses, in particular those resources found consistently related to student outcomes. Thus, it becomes of particular interest to study equality of access as a function of student’s geographic background. Here a historical review and comparison of outcomes of students of the Northern State (previously, province) and those of the Capital city of Khartoum may be a representative example. Up to the late seventies, students of the Northern Province used to be the top achievers in the national intermediate level exam, so they commanded a high rate of admittance to the secondary schools. Students of Khartoum used to be mid-level achievers in comparison with these and other rural students. The case is reversed nowadays. Outcomes of students of the Northern State are low and declining, while those of Khartoum are continually staying at the top. The director of education of the Northern State stated

in a newspaper¹¹ that for the past years only 2% of the university students population came from the Northern Province and that student dropouts from schools is becoming a phenomena. He states, as an example that out of 1750 students enrolled for the first grade in a school, only 511 made it to the eighth grade.

Education is a cumulative process, and the quality of inputs made available to schools of primary and secondary levels is directly related to the quality of student outcome. Urban students, especially those of Khartoum's are more likely to have access to favorable course-taking opportunities, school climates, qualified teachers, availability of teachers for various fields, better infrastructure and better equipped classrooms. Even within Khartoum, school conditions vary according to the living conditions of the quarters that they are meant to serve. School conditions in places like Riyadh, Amarat, or even Arkawet, differ to a great extent from those of the poor areas of Mayo or Salama.

This analysis of diversities and inequalities in our school system reveal the need to deal with important limitations and shortcomings in our present educational data systems. To be of practical use to policy-makers, and researchers, our educational data systems should pay wide attention not only to the outcome side, but also to the input side. Data on the student's socioeconomic backgrounds, and key educational resources, is completely lacking. The early statistical reports of the Ministry of Education used to include some financial data about budgets allocated to education at the country level. Its coverage, though limited, was indicative of the degree of priority given to education, and kind of progress made. Nevertheless, for reasons unknown, that part of data was excluded from the latter reports, and this was one of the miserable counter developments of the educational data system. Of course, there were and still are the obstacles of inflation, decentralization and inflated figures allotted in governmental budgets that never materialize. Additionally, there are means and methods to handle such obstacles and produce data of immeasurable importance and inclinations. In fact, for us to be able to investigate state-to-state diversities in student outcomes throughout the years, it becomes important to have this kind of detailed data about resources that are actually being annually allotted to the educational sector in each state from various sources. Allotment of expenditure accrued to various elements of the educational processes are

¹¹ Al-Anbaa Newspaper of 2/5/2003

basic for assessing the quality of education and verifying student outcomes. It shows which states, if any, includes in its educational budgets allocations for items other than teacher salaries and benefits. Allocations for items like libraries, laboratories, instructional materials, maintenance, student activities etc., produce important indicators about the quality of education.

Data about the socioeconomic background of the student: parental education, occupation and income, family demographics, housing, neighborhood and other social conditions is important and will also provide key sets of indicators.

Those varied sets of data are expected to provide more refined indicators that help in identifying specific targets and strategies for interventions aimed at reducing inequality and thereby improving overall levels of student proficiency.

We can sort out some inequalities that affect student outcomes and that need to be reflected in the educational data system as follows:

Differences as a Source of State Inequalities in Student Proficiency

Existence of social and geographic inequality in achievement constitutes a troublesome and enduring aspect of schooling. Large achievement gaps between students of high and low socioeconomic status and between urban, rural and nomadic students have to be thoroughly scrutinized and revealed.

Home Environmental Inequality

From one standpoint, the school is an essentially neutral learning environment passively allowing sharp inequalities in home circumstances to translate into similar inequalities in learning outcomes. Families have long been known to vary substantially in their capacities to provide educational environments that foster school readiness and reading literacy. Such differences are linked to social status indicators, including income, parental occupation, and parental education. Parents of high social status are more likely than parents of low social status to have the

resources and skills needed to support their children's academic learning.¹²

If this explanation were sufficient to understand observed achievement gaps, variation in student achievement between schools would simply reflect the varied home environments of students attending those schools. Policy interventions should aim at alleviating impacts of these inequalities or minimizing their role.

School Environmental Inequality

From an entirely different standpoint, schools are a much active force, subjecting essentially similar children to dramatically different learning experiences and thereby actively recreating in each new generation a wide intellectual inequality that conforms to the wide inequalities in earnings and occupational prestige.¹³ If inequality of schooling were the sole determinant of inequality of educational outcomes, inequality in school mean achievement would reflect school differences in policy and practice. "Not surprisingly, those who have emphasized the school as a casual agent in creating educational inequality, while often endorsing compensatory educational policies, have called for sweeping structural reforms in the provision of schooling. These include the elimination of tracking, school finance reform that would equalize spending across rich and poor districts, and a recasting of teacher preparation to foster more favorable expectations and more cognitively stimulating instruction for currently disadvantaged students. If the 'school effects' explanation were correct, such reforms would reduce or eliminate differences between schools in achievement."¹⁴

The debate reviewed above leaves school differences in student mean outcomes open for different interpretations. An elevated school mean may be viewed by one observer as reflecting an advantaged student composition; another may attribute this success to excellent school governance, organization, policy, and instructional practice. Those who want to study school effects should try to measure key aspects of both student composition and school process to assess the relative contribution of each and to isolate those contributors to achievement

¹² Stephen W. Raudenbush,(1992) School of Education and survey research center, University of Michigan

¹³ *ibid*

¹⁴ *Ibid*

that reformers can modify. Casual inference in such studies should be conservative because student composition and school processes are inevitably correlated. Thus, if student composition or school process is not measured well and is still included in the analysis, estimators of both will be biased.

So given this difficulty of conducting sound studies of school effects, then holding schools accountable for their mean achievement levels would not be a wise decision. School means that are not adjusted for student composition will typically convey an overly negative picture of school process in those schools with the most disadvantaged students, and vice versa.

The student's contribution is an important factor in the educational process. If he lacks the motive to learn and relate his self-esteem and self-confidence to his level of educational attainment, then the traditional educational system can not be of much help. Students are key factors in shaping school outcomes, and therefore a central focus of educational productivity. Schooling is something that is done to the student. Education is something that the student should essentially do for himself. No matter how good the school system is, or how interested the family is in educating its children, all those resources will just be wasted if the student doesn't have the motive. So the student's motive and perception of his gains from education are elements to be considered and included in an efficient model for assessing academic attainment and school outcomes. If our objective is to find a production function a mathematical expression of the relationship between inputs and outputs in education, then students are the key factors in shaping our thinking about productivity. Sudan as well as many other countries, made major efforts to expand educational provision since the 1950s with the idea that "more education would lead to many other social goods, such as increased economic success, greater social harmony, less poverty, less crime, and the like."¹⁵ "More is spent on education, yet economic and social circumstances do not seem to improve."¹⁶

¹⁵Benjamin Levin, (1993) Students and Educational Productivity, University of Manitoba- Educational Policy Analysis Archives, Volume 1.

¹⁶ Ibid

David Monk (1990)¹⁷ uses in his book the production function as the basic element in studying productivity in schools. He defines a production function as a model, which links conceptually and mathematically outcomes, inputs, and the processes that transform the latter into the former in schools. He notes that production functions are important for improving both technical and allocative efficiencies. However, despite their potential benefits, Monk recognizes the major obstacles that face the creation of production functions for education. However, neither outcomes or inputs nor processes are simply understood as proved in the following: *in education, outcomes are multiple, jointly produced, and difficult to weigh against each other*. The outcomes of education are not all translatable into a standard metric, such as money, which makes it very difficult to give them a relative value. A further difficulty with outcomes has to do with the level at which they should be measured. At various times researchers have been interested in outcomes of individual students, classes of students, schools, school districts, states, nations, ethnic groups, age groups, gender groups, and all sorts of other subsets of the population. Monk is aware of the difficulties in dealing with both micro and macro analysis. He concludes that there is no one best approaches.”¹⁸ Monk mentions that the proper level of analysis depends largely on the nature of the phenomenon being studied. Some phenomena are district rather than school or classroom phenomena and have effects that are felt throughout school districts.

The inputs of schools: buildings, teachers, textbooks, etc., are relatively easy to recognize. A teacher might spend time providing tutorial instruction to a student. Nevertheless, the student may or may not be attentive to that instruction. This is a case where resources being provided did not flow to that inattentive student as illustrated by Heathfield and Wibe (1987): “Time is another significant problem in studying educational productivity. It seems reasonable to believe that students will learn at different rates. Yet this seemingly innocuous conclusion creates enormous difficulties for analysis, since it means that different resources at different times and in different arrangements may be necessary for different students. Indeed, there could be a unique production function for each child, or even several production functions

¹⁷ Monk, D (1990) *Educational Finance: An Economic Approach*, New York: McGraw-Hill.

¹⁸ Ibid

for each child under different circumstances.”¹⁹ Some children may be made by family or school to exert so much effort studying at home or preparing homework, but that effort doesn’t show in their exam results. That may not necessarily happen because of their being stupid, but rather may be because they lack the necessary organizational abilities to help them make use of that effort.

A child may have the motive but lacks the *know-how* to promote that motive and thus learning is significantly influenced by factors outside the school. Intelligence of the student is an important factor, but factors like home and background surroundings exert varying levels of significance. Monk also mentions various technical problems in studying productivity in education; limited variation among schools in many attributes, possibility that both inputs and outputs are collinear, and likelihood that inputs and outcomes influence each other. However, we cannot verify that there is a real possibility that some aspects of education are anarchistic by which Monk means that actors are not goal-oriented, so that even if a best way of doing things were known, people would not pay attention to it.”

Though there are so many problems in studying productivity with regards to the difficulty of measuring some of its components, but still there is faith in the existence of a production function, with the *outcomes-as-standards strategy* as a new means of gaining insight into the function’s properties”²⁰. Monk goes on to advocate the study of productivity through looking at the properties of the classrooms. This approach is based partly on the belief that teachers will use different instructional methods with different classes of students. He discusses the ways in which these responses by teachers might occur depending on the students, and suggests that teachers may have individual patterns of adjustment that could be studied and defined in terms of their impact.

Educational Production and the Role of the Students

The study of the productivity in education will be greatly hampered if the role of students in generating education outcomes is downplayed. A better understanding of productivity in education requires attention to what students think and do.

¹⁹ Ibid

²⁰ Ibid

The idea of a production function depends on seeing education as being a production process, where inputs are being transformed into outputs in a standard way. In a factory, raw materials are turned into finished products through various production processes. However, in a schooling it is not evident what (or who) the raw materials are, nor who is doing the producing, nor what is the product. An important question here arises as to whether students are raw materials being processed, or workers doing the producing. When formulating a production function, students are seen as producers to the extent that issues as their motivation and efforts are taken into account. Yet they are seen as materials to the extent that studies focus on the antecedents of students, much as we want to make sure that a certain quality of steel went into the manufacture of cars. The implicit assumption is that schooling is something done to the student for him to produce certain outcomes. Interest in students is chiefly a matter of how well our efforts have succeeded. An example of that may appear in the high publicity and amount of attention given to schools and students with top scores in the secondary school exam results, a kind of annual ceremonial practice that beside its cons has so many pros. In fact learning about the students and understanding their perceptions and changes in these perception from generation to generation and from location to location is of utmost importance.

Students are the ones who do the learning. Whatever schools provide, or teachers do in the end it is the student who must use the resources to acquire skills and knowledge. Our production function should reflect the contention that students, unlike other raw materials, are the people who do the learning. Education is not something we do to people, but something that people do for themselves, assisted by the efforts of teachers and school policies. What happens in a class is fundamentally dependent on who the students are, how they make sense of their world, and what they want or do not want to do. So many factors affect their attitudes toward schooling and reason their decisions about the amount and kind of effort, attention and interest they put into their schoolwork. Students are the ones who decide to come to school or not, to take the material seriously or not, to focus on grades or not. However, these decisions are not entirely independent of what schools and teachers do nor are they solely determined by what happens in schools. So, it is dangerous to arrange schooling in a standard and rigid format, for such plans would come to nothing in the face of students

with different backgrounds. A decision to close boarding houses in schools all over the country caused poor students of remote rural and nomadic areas to drop out of the school system. It is not merely that students are shaped by their backgrounds or abilities, though both are important. For every student, background, abilities, and a variety of other circumstances produce a unique biography and personality and a unique way of responding to the world. These elements are expressed in schools as they are in every other aspect of life. The present policy of amalgamating the primary and intermediate levels into an eight-year basic education can be considered as a showcase of the previous discussion. The length of duration of that level has led to an accumulation of diverse sets of students of varying motives, backgrounds and age groups. Since children's behavior in school is influenced by their family, social class background and peer culture, then those students with low family guidance will mostly distinguish themselves from other groups of students within the school. They mostly set out to be disruptive, aggravate teachers and avoid doing what the school wants. They are mostly the ones expected to fare badly in their educational attainment, repeat over the years and stay long at school or drop out at the end. The accumulation of their numbers at schools is likely to produce negative influences on other normal students, bear negatively on the school environment and, as it happens sometimes, may lead to some violent, or even criminal, behavior. Incidence of diversity of the student community within the basic schools is being aggravated by the 8-year amalgamation practice, a policy with its cons and pros that needs to be subjected to serious research to be able to verify its actual impact. This research should include the students, learn directly from them their way of thinking and attitudes toward the school system and reveal their sufferings. This kind of research is vital because the schools of today are dealing with tomorrow's men and women of the country, and if we are aspiring for them to prosper, then we better take utmost care in preparing them to accomplish their mission the best way we can. What the students do and think is central to education and should be central to the way schooling is organized."²¹ Most of the policy attention about schools focuses on such matters as curriculum, teachers, school organization, or school governance. Policies in these areas are presumed almost unthinkably, to yield changes in what students do, think, or learn.

²¹ Ibid.

Consider various sides of the debate over restructuring schools. One approach has been what Fullan (1991)²² calls the ‘intensification’ approach: -stricter curriculum requirements, closer supervision of teachers and students, external examinations, and so on. Here the assumption is that teachers and administrators will be tougher on the students, and that students will respond to the changes by intensifying their own efforts at school. The strategy could be phrased as ‘making them learn whether they want or not’. Put this way, of course, it is clearly unworkable, since we have abundant evidence that though we may be able to influence, we cannot control what students learn. If we could, presumably we would already have taken steps to make sure all students learned what we wanted them to. As soon as we see students as both workers and product, clearly a strategy of intensification will not, by itself, be successful, since it does not take into account the power and range of students’ ideas and motivation.

The Alternative Policy

The main alternative policy currently being proposed is the *professionalization* approach, in which more authority is to be given to teachers to take the steps they see as most desirable. In some versions authority is moved to school communities, which include teachers, parents, and sometimes students (Zeichner, 1992)²³. Nevertheless, if we think of students as central, then this strategy too seems unlikely to succeed. It assumes that teachers know what to do to create more learning and that they will do so if given the authority. Neither assumption seems credible. It is reasonable to think that most teachers have a real concern about students and their welfare. It is not reasonable to think that all teachers have a tremendous store of knowledge about how to educate that they are waiting to unleash with dramatic effect as soon as they are freed from the shackles of bureaucratic restrictions. Nor is it reasonable to think that teachers will, any more than other occupational group, always recognize the best interests of the students or have those interests at heart when they conflict with teachers’ own ideas,

²² Fullan, M. (1991). *The New Meaning of Educational Change*. New York: Teachers College Press/ OLSE Press.

²³ Zeichner, K. (1991). Contradictions and tensions in the professionalization of teaching and democratization of schools. *Teachers College Record*, 92(3) 363-378

needs and interests. If neither intensification nor professionalization is a good strategy given a belief in students as the center of education, what policy alternatives do we have? Many could be derived from available evidence about how people live and learn.

Perhaps most importantly, we would need to pay much more attention to the issue of motivation. If students are the producers of their own learning then their motivation is critical. There is a substantial literature on motivation, both in education and psychology (e.g., Ames & Ames²⁴, 1984, 1989; Deci & Ryan,²⁵ 1985; Hastings & Schwieso²⁶, 1987). Various strategies for the organization of schooling and teaching have been advanced based on this research. Nolen and Nicholls²⁷ (in press), in reviewing the literature, came to the conclusion that the most effective strategies have to do with treating students as capable persons, capitalizing on their knowledge and interests, and involving students in determining goals and methods of learning.”²⁸

By bringing in the above discussion and theories, I didn't mean to intrude into the sensitive and specialized issue of educational policies. Educators of the Ministry of Education are more qualified and equipped to do that, and most probably, their knowledge of these issues goes beyond what I have brought. However, our main concern here is research and a proper research cannot be accomplished or be of practical use if we do not have real knowledge and information about the world we are researching. That knowledge and information cannot be affirmed without the availability of plenty of good and varied statistical data. The above mentioned discussion, mostly theorized by educators, showed the immense opportunities of research and the variety of fields of data collection, and the role that statistical methods of analysis can play to be of help to our educators to uncover our world and design policies to suit it.” Wherever there is a mass of numerical data that admits of

²⁴ Ames, R. & Ames, C. (Eds), (1984), *Research on Motivation in Education*, San Diego: Academic Press, Ames, R. & Ames, C. (Eds), (1989), *Research on Motivation in Education*.

²⁵ Deci, E. & Ryan, R. (1986). *Intrinsic Motivation and Self-determination in Human Behavior* New York; Plenum.

²⁶ Hastings, N. & Schwieso, J. (Eds.), (1987). *New Directions in Educational Psychology*, 2; *Behavior and Motivation in the Classroom*, East Lewis; Falmer Press.

²⁷ Nolen, S. & Nicholls, J. (2002). A place to begin (again) in research on student motivation; Teachers' beliefs. *Journal of Teaching and Teacher Education*

²⁸ Ibid.

explanation, the statistician should consider its analysis his field of endeavor.

To utilize statistical methods to advantage, a person should:

- Be well versed in the subject matter of the field in which the research is to be conducted.
- Know how to organize masses of data for efficient tabulation and how to lay out economical routines for handling data and tabulation.
- Know effective means of presenting data in tabular and graphic form.
- Have some knowledge of the mathematical theory of statistics in order to have assurance there is a fair correspondence between his data and the assumptions underlying the formulas he uses.
- Be acquainted with a variety of statistical techniques, the limitations and advantages of each, the assumptions upon which they are based, the place each occupies in a logical analysis of the data, and the interpretations, which can be made from them.

Statistics, then, boils down to numerical results, the methods and processes used in obtaining them, the methods and means for estimating their reliability, and the drawing of inferences from these results.²⁹ The multiplicity of variables that are supposedly influencing students' academic achievement as detailed in the aforementioned discussion, makes it hard to collect data about them on a census basis or to include it, with all its needed details, in an annual report. Such annual reports and censuses are, for convenience and practicality, intended to cover only the major characteristics and parameters of the whole of the target population. The details are usually covered through sampling procedures, whether sampling surveys or experimental designs. Surveys are used to collect data about certain characteristics from a few randomly selected units of the population. A sample is just a part of a population selected according to some rule or plan whose results are used to make inferences about the population from which it was selected. There are many types of sampling procedures, which differ in accordance with the type of

²⁹Bernard Ostle, (1966) *Statistics in Research*, Second Edition
The Iowa State University Press.

randomization used to select its units. Since data collected from the sampled units does not by mean anything as the aim is to use it to infer about the population. Therefore, such selected observations should be as closely representative as possible to the nature of that population. The more diverse and complicated the nature of that population is, the more complicated will be the sampling procedures followed. Samples are preferred for economic reasons and because they produce efficient results in the shortest possible time. Of course, that means covering the fewest possible units or the smallest sample size. Sample size is by necessity, determined by the degree of diversity in the characteristic, or characteristics, being investigated in that population, sensitivity of the issues being investigated, and the margin of error being allowed in our estimation of the population's characteristics from that sample data.

While samples collect data about the sampled units at a certain point of time, experimental procedures collect data for an interval of time. The function of experiments is to eliminate untenable theories. Since the scientific method is the pursuit of truth as determined by logical considerations, then experimentation, as one of its tools, is used to test some hypotheses and to discover new relations among and between variables. It is important here to mention that no hypothesis, which states a general conviction about a certain set of parameters, should be demonstrated as being true. Inference about it can only be probably true within a stated margin of error.

Experimental design is the plan used in experimentation. It is an application of certain treatments to some experimental units, then analyzing the effects of that application when the data become available. Some kind of experimental design was being used at Bakht-El Ruddy training institute. Any changes that were planned to be introduced onto the ongoing school curriculum with the intention of achieving certain policies like improving school processes or student educational attainment were firstly subjected to intensive and lengthy testing at that institute, before making inferences on the validity of their impact. It was an application of the theorems of experimental design, which though may not be numerical, uses the same procedures and satisfies the requirements of a scientific experiment.

Educational indicators, as mentioned before, do not have inherent ability to explain their existence as they are, or give a valid interpretation of the factors that lead to any changes that are reflected through them. Besides, they do not give any measure of the degree of

error that might be inherent in their calculation. Statistical models are the ones better qualified to do that, and more, to decompose indicators or bring out factors of significance in their composition.

The Concurring Models

Since such models need to cater about as much details as possible to single out the factors or variables of significant effects on the response variable. Therefore, detailed and consistent data are critically required for solid analyses. Annual reports will be inconveniently bulky, hard to process and expensive to produce if they were loaded with many details. Detailed data would more conveniently, and efficiently, be produced through sampling procedures. Sampling surveys can be used to collect detailed data about specific issues of interest and cover all or certain parts of the population.

Therefore, for building meaningful models that can be of any help to the policy makers, a complete review of the present educational data system is in urgent need. The present annual educational statistical report needs to be redesigned and its coverage reviewed, in accordance with a pre-defined strategy. A system of sampling surveys and experimentation needs to be applied to cater for details and attest hypotheses about the validity of policies designed for improvements. That, when accomplished, needs to be accompanied by the introduction of research incentives, research facilities, and a work environment that is hospitable and accommodating to research, and that uses its findings as the base for its future plans...

Building Educational Models

Our proposed system of statistical surveys should aim at finding out within states and between states variations in the characteristics and factors affecting students' attainment as detailed before. If we take examination scores, a measure of that attainment, for a response variable, and home and school correlates as our influencing variable components, then a model could be built for within state and between states comparisons. We can hypothesize that the student's social, economic and geographic background, along with indicators of home literacy and home environment are expected to be closely related to his educational proficiency. Those indicators that are key aspects of school

quality, such as course-taking opportunities, disciplinary climate, teacher quality, class size, classroom conditions, availability of basic facilities in the schools, and may be cognitive stimulation in the classroom, would also predict proficiency.

It is reasonable to expect that geographic location is a key variable which, if measured at the student and school levels, would account for most of the variation between states. Controlling for explanatory variables at lower levels of aggregation, such as the student or the school need not reduce variation at higher levels, such as the state.

It can safely be assumed that states will vary in outcomes for two reasons; (1) Selection processes and, (2) Measuring the effects of state educational policies and practice.

Selection processes arise because patterns of settlement, fertility, and economic conditions produce state variation in the demographic and cultural backgrounds of the students and their families. Educational policies and practices of schools vary as a result of the decentralization policies.

In order to control these factors we hypothesize and explain the greater part of state variation.

The strategy for modeling student attainment has to be built on two stages; a within-state and between-state analysis. Since sampling procedures are the means that are normally used for data collection, the within-state analysis has to use hierarchical linear models that suits the sampling procedure followed. The output for each state is a vector of parameter estimates and their estimated sampling variance matrix. These then, provide input data for the second stage of analysis, which involves findings across the states.

The recently applied policy of closing boarding houses especially of schools at the lower levels of the educational process had an impact that needs to be researched. Villages in some rural and nomadic areas are mostly scantily populated and scattered. A single village of those is, usually not expected to have enough pupils for a school to be viable. Accordingly, schools are installed to serve groups of villages. From here arises the logic behind the opening of boarding houses. Being poor and barely motivated to educate their children, those mostly illiterate families used to be persuaded by the presence of boarding house facilities to venture into enrolling their children into schools. Our hypothesis is: Closure of boarding houses had a drastically negative impact on the progress of education in those areas of the country; that policy lead to

increasing rates of students dropping out of the educational system, and in some cases, the closure of whole schools; thus resulting in a preeminence of illiteracy in those societies. Lack of data, as usual, is a deterrent; but if the idea is agreed upon, the topic could be included as part of a multipurpose sample survey to collect data about it and other variables of importance in assessing and redesigning educational policies.

Another topic that I hypothesize is of prime effect on the motivation of families to educate their children, and on the motivation of students to want to be educated, is the lack, or extreme scarcity, of employment opportunities for those who complete their schooling, and the subjectivity of selection to those scarce job opportunities. That was stated as “a fact of life” by the Minister of Education of the Northern State.³⁰ He stated in a recent press conference that families are continually pulling their children off the educational system because they think it more rewarding to use their services to help them gain a living. That is a dire situation, which, if not dealt with properly and soon, will put the whole future of this country into jeopardy. The principle of Equal Opportunity Education and the future of the whole of the educational system are at stake, and if not dealt with promptly, increasing parts of the future generations are almost surely going to join the ranks of the illiterates’ club. Researching the problem is the only way to handle or sway away such dire situations, and is the most qualified to propose the most suitable solutions that can lead us to the right path to face our problems scientifically.

Of course one sample survey will give us data about the state of affairs at a single point of time. A series of surveys is needed to enable us build a time series analyses and find out to which directions real life affairs are moving. Since, for now, we do not have any dependable data about any point of time, then it is time we begin looking seriously into building a needed efficient and sufficient database. One of the objectives of these surveys should be our need to find out what people really think about education, their motives and the factors behind those motives. It should not be taken lightly that a high portion of the families in the Northern State are pulling their children out of school as stated by the minister. Families of that state were historically known to pay utmost care to education. Literacy rate in that state for those of 15 years of age and over is 71.5%, the highest in the country, whereas that rate is just

³⁰ Al Anbaa newspaper 2/5/2003

49.5% for Northern Sudan ³¹. If education is declining in such a state, then the red lights of danger are on about the future of the educational process in the whole of the country and speedy actions are needed. That minister based his statements upon the findings of a study group that he commissioned, and though we do not know that much about the procedures they followed in their study. However it is, definitely, a step forward in the right direction. Similar measures should be taken in other states, but on experimental basis, to find out the extent of emergence of such a problem and run within and between- state comparisons. The resulting analysis should take care about finding out the factors that can significantly explain such criteria as emergence of dropouts or lack of motives, and help predict their trend in the future, or what we term as models.

Case Study Evaluations³²

Rather than using large samples to examine a limited number of variables, case study methods involve an in-depth, longitudinal examination of a single instance or event. It is a systematic way of looking at what is happening, collecting data, analyzing information, and reporting the results. It produces a better understanding of why a thing has happened as it did, and what might be important to look at more extensively in future research. So, case studies are especially well suited toward generating, rather than testing, hypotheses.

The author discusses six types of case studies, where, for each, he presents the type of evaluation questions that can be answered, the functions served, some design features, and some pitfalls.

Types of Case Studies

Illustrative Case Studies are descriptive; they utilize one or two instances to show what a situation is like. This helps interpret other data, especially when there is reason to believe that recipients know too little about a program. These studies help familiarize the topic to the

³¹Central Bureau of Statistics , , Khartoum ,2002, Population Data Sheet for Sudan by State.

³² Edited from “The Application of Case Study Evaluations” by Lynn Davey, ERIC Clearinghouse on Assessment and Evaluation.

recipients and give them a common language about it. The chosen site should be typical of important variations, and contain a small number of cases to sustain interest. This may be typical when introducing changes in the school curriculum, or entering some changes on the text or policies.

There are pitfalls in presenting illustrative case studies. They require presentation of in-depth information on each illustration. The case(s) must adequately represent the situation or program. Where significant diversity exists, it may not be possible to select a typical site.

Exploratory Case Studies are condensed case studies, undertaken before implementing a large-scale investigation. Where considerable uncertainty exists about program operations, goals, and results, exploratory case studies help identify questions, select measurements constructs, and develop measures; they also help to safeguard investment in larger studies. They are sort of pilot surveys that are used to test procedures, questions, and see if the answers satisfy what is intended behind the study. The greatest pitfall in the exploratory study is pre-maturity: the findings may seem convincing enough to be released inappropriately as conclusions. Other pitfalls include the tendency to extend the exploratory phase, and an inadequate representation of diversity. I hope that the aforementioned statements of the Minister of Education of the Northern State were not depending on findings of an exploratory study. Because if the case was so, then he would have fallen into the pitfall of pre-maturity and jumping to the inappropriate conclusions; and would have generalized resulting findings irrespective of inherent diversities in that state.

Critical Instance Case Studies examine one or a few sites for one of two purposes. A very frequent application is the examination of a situation of unique interest, with little or no interest in generalization. A second, rarer, application entails a highly generalized or universal assertion which is called into question, and we can test it by examining one instance. This method is particularly suited for answering cause-and-effect questions about the instance of concern. The most serious pitfall in this application is inadequate specification of the evaluation question. The importance of probing the underlying concerns in a request is crucial to the appropriate application of the critical instance case study.

Program Implementation Case Studies help discern whether implementation is in compliance with its intent. These case studies are also useful when concern exists about implementation problems.

Extensive, longitudinal reports of what has happened over time can set a context for interpreting a finding of implementation variability. In either case, generalization is wanted and the evaluation questions must be carefully prepared. Policy programs like the cancellation of the intermediate level of schools and introduction of an eight- year basic level is a **case** that needs to be studied and its implementation reevaluated. A requirement for good program implementation case studies is investment of sufficient time to obtain longitudinal data and breadth of information. Multiple sites are typically required to answer program implementation questions; this imposes demand of training and supervision needed for quality control. Replication is always a statistical requirement for the credibility of results and minimization of bias, while multiplicity of sites is a necessity to account for and represent diversities inherent in the groups under study.

Statistical Indicators of Performance

The hectic pace of school life makes it difficult for teachers and administrators to step back and objectively assess the validity of their operating assumptions. In addition, educators tend to judge success anecdotally rather than through formal assessment. A small sign of progress from a recalcitrant student may outweigh months of low performance. Although these victories may be highly satisfying in human terms, today's accountability environment demands that educators collect and analyze objective data before making decisions.

Schools usually collect large amounts of data, but much of it is used for short-term illustrative purposes and then filed and forgotten. That is the kind of good and factual data that can, if aggregated for regions and states, form the base for calculating "performance indicators" which not only satisfy the demands of accountability and serve as a tool for school improvement.

This chapter examines the nature and purpose of educational-indicator systems, and discusses the design of report cards by which schools can inform upper echelons of the educational hierarchy and the public of their performance.

What are Educational Indicators?

An indicator³³ is any statistic that casts light on the conditions and performance of schools. A push for accountability emphasizes test scores, but it is important to note that a comprehensive indicator system should provide a wider range of information.

Some indicators, such as teacher turnover or student mobility, can signal problems that need attention. Some indicators can provide information geared to current policy issues; for example, data on course-taking will help policymakers who want students to take more academic courses.

Other indicators focus on context, such as student demographics, teacher workload, financial resources, and teacher qualifications. Such information can help schools interpret the, sometimes, ambiguous statistics that come from test scores and other outcome measures. Although contextual factors do not provide the bottom-line measure of success that policymakers seek, they do have an impact on student learning and can help explain a school's performance.

Policymakers and school administrators should choose their key indicators by asking three questions: Why is this information important? How much effort is required to track the data? How will we use this information when we get it?

How Do Indicators Support School Improvement?

Indicators should if available, play a central role in educational accountability systems by focusing attention on results, especially the school's performance on standards-driven assessments. Comparative regional ratings can raise public attention and awareness, provide focus and energize schools and communities to work to improve student achievement. At their best, ratings can provide momentum, measure schools' progress and show parents, the public and policymakers about how schools are doing and enhance them to look for measures for improvement. This attention-getting feature is even stronger when indicators are the trigger for incentives, giving practitioners personal as well as professional reasons to focus on the target. However, attention

³³ Lashway, Larry, (2001), Educational Indicators, ERIC Digest

does not always lead to positive action. Educators may attempt to explain away poor results rather than act on them, while parents and community members often report that they are uncertain how to lobby for improvement. Teachers in high-need schools, struggling to educate large numbers of under-prepared students with limited resources, may simply get demoralized by repeated public embarrassment and public criticism of the bad results. This also does not help the student moral or buildup of self-confidence in these schools. The more lasting value of indicators is their role in the school-improvement process. Used thoughtfully and systematically, they allow schools to take charge of their own assessment by identifying strengths and weaknesses and pinpointing which improvement strategies are working. Ideally, a school's indicator system will not be merely a grudging reaction to state mandates but will reflect a school's, and the educational system's, commitment to "an ethic of continuous improvement" Used this way, indicators are merely an extension of what thoughtful professionals should always try to do.

Indicators Misuses

Although indicators hold out the promise of improved decision-making, they can easily lead schools astray. One danger is to collect data indiscriminately. This not only costs effort and money, it swamps decision-makers in a sea of numbers that make it difficult to distinguish the significant from the trivial. Second, raw numbers never speak for themselves, but require careful interpretation. For example, a rise in the eighth -grade scores in a certain year may be due to improved instruction, or it may be due to differences in capability between last year's group and this year's group. The focus in universal primary education implies equity in terms of both access and quality, irrespective of gender, race or any other socio- economic circumstance."³⁴

Access is often measured by 'apparent Intake Rate (AIR), being the number of children entering primary school regardless of age, as a percentage of the of the official entry age:

$$\text{AIR} = \frac{\text{New entrants in Primary Grade 1}}{\text{Population of official school- entrance age}}$$

Population of official school- entrance age

AIR, also referred to as the *Gross Admission Rate*, is used to measure the admission of children in the education system. The *Net*

³⁴ Education For All Assessment ,1998, WGES

Intake Rate (NIR) refers to the number of children of the official school-admission age admitted into the first grade of primary education, expressed as a percentage of the population of official admission age.

NIR = $\frac{\text{New entrants in Primary Grade 1 of official school-entrance age}}{\text{Population of official school-entrance age}}$

The NIR measures the admission of children of the right age, that is, of the official entry age. It also shows the proportion of eligible children who were not admitted. The AIR is used in place of the NIR when data on enrolment by age, of new entrants into the level is not available. The difference between AIR and the NIR gives an indication of the extent of the under-aged and, more often, over-aged enrolment at the first grade.

The *gross enrolment ratio* is the total enrolment, irrespective of age, expressed as a percentage of the eligible school-age population.

GER = $\frac{\text{Total number of students in primary education}}{\text{Population of the official primary school age}}$

The *net enrolment ratio* is the proportion of children of eligible school-age group who are enrolled.

NER = $\frac{\text{Total number of students in primary education of official}}{\text{Primary school age}}$

Primary school age

Population of the official primary school age

The difference is that the former includes also those who are not within the eligible age group.

The survival rate to Grade 5 is the percentage of a cohort of students enrolling in grade 1 of primary education, who subsequently attain grade. Its purpose is to assess the internal efficiency of an education system and the attainment and retention of literacy, which is normally equated with reaching grade 5 of primary education. Further, this indicator reveals the magnitude of the dropout rate before grade 5.³⁵

The Human Development Indicator (HDI) measures the average level at which any given country stands. This composite indicator contains three variables: life expectancy, level of education (measured by the adult literacy rate and school enrollment rate) and the GDP per

³⁵ Education For All—previous reference

inhabitant. There are so many other educational indicators of significant importance like pupil-teacher ratio, class size, ratio of trained teachers, rates of success in the final regional exams, percentages of its students admitted to the secondary level, dropout rates and so on. Those indicators, though ambiguous in some cases, will be of great value if used in a comparative capacity and for a time series. These indicators, if calculated annually for regions within a state, for states within the country and for the country as a whole, will definitely pinpoint problematic areas. Some of them will help directly explain the reasons behind the existence of others.

Apparent Intake Ratio for the regions of Northern Sudan³⁶

<i>Region</i>	% <i>1979/80</i>	% <i>1982/83</i>	% <i>1989/90</i>	Index Of growth
Northern	78.0	88.6	101.8	131
Khartoum	64.7	73.9	75.5	117
Central	49.1	73.7	88.0	179
Eastern	42.1	47.8	61.1	145
Kordofan	33.2	50.5	82.3	248
Darfor	31.8	47.0	69.3	217
Sudan	44.2	55.8	90.3	204

This table shows the progress in the apparent intake ratio (AIR) in the northern regions of the country for the ten years from 1979/80 to 1989/90. The index of growth relates 1989/90 to a base year of 1979/80. Data for the southern regions was excluded because of its in-reliability. Instability caused by internal in-fighting and civil strife there led to mass population movement, and hence instability of the education system. The data of the table and the index of growth, show an impressive improvement in AIR indicator over the ten years period. The index shows that regions with the lowest AIR in 1979/90 (Kordofan and Darfor) are speedily catching up to the national level(90.3) in 1989/90. The AIR for both has more than doubled. It could be noticed that the Eastern region is lagging behind. Though its AIR indicator of 1979/80 was higher when compared to both of Kordofan's and Darfor's, but its speed of progress was lower to the extent that it claimed the lowest AIR(61.1) in 1989/90. The overall index of progress for the

³⁶ Annual Statistical Report, Ministry Of Education, Khartoum

country(204) shows that the overall AIR has more than doubled in 1989/90 compared to 1979/80. Its AIR of 90.3 in 1989/90 indicates that the educational system was successful in approaching achievement of the goal of eliminating illiteracy among that age group of youngsters. Leaving only less than 10% of the 7-year olds out of the educational system is a remarkable achievement that needs to be accomplished in all of the different regions of the country.

Gross enrollment ratio by region in Northern Sudan³⁷

<i>Region</i>	<i>% 1979/80</i>	<i>% 1982/83</i>	<i>% 1989/90</i>	index of growth
Northern	74.5	84.4	75.1	100
Khartoum	67.2	80.7	71.3	106
Central	50.2	68.8	69.5	138
Eastern	27.1	39.6	42.5	159
Kordofan	28.6	39.3	52.2	183
Darfor	22.2	32.6	44.5	200
sudan	38.7	46.2	61.9	160
variation				

Disparity refers to lack of equality or parity among different components. Parity between the sexes, states or regions within the states is a strategic policy issue naturally targeted by educational policy makers. Since primary education is universally considered as a prime right of the people or the basic tool used for achieving the universal target of eliminating illiteracy, then the AIR(apparent intake ratio) and the GER(gross enrollment ratio) are indicators that show our progress toward achieving that goal. Comparing these indicators throughout the years will show the overall progress over time, while comparisons between, and within states, will measure the parity of that progress. The table shows some discrepancy in the ratios of the Northern and Khartoum regions

³⁷ Ibid

for the year 1982/83 that I find hard to explain. Apparently, there is an inconsistent jump in their indicator for that year in comparison with the other two years. That is an example of the inconsistencies that are common in the annual statistical educational reports, a fact which questions their efficiency and makes it extremely hard for researchers to depend upon them for serious research.

Table 3 Apparent intake ratio by sex for regions of Northern Sudan

<i>Region</i>	<i>1961/62</i>		<i>1966/67</i>		<i>1982/83</i>		<i>1989/90</i>		<i>Index</i>
	<i>m</i>	<i>f</i>	<i>m</i>	<i>f</i>	<i>m</i>	<i>f</i>	<i>m</i>	<i>f</i>	
Northern	56.0	28.2	55.1		88.8		105.6		189
			39.2		84.2		97.8		347
Khartoum	71.9	56.3	88.2		73.9		75.5		105
			69.7		61.8		73.2		130
Central	55.3	25.7	59.4		73.7		88.0		159
			32.1		63.6		79.9		311
Eastern	37.6	17.7	51.7		47.8		61.0		162
			33.3		40.5		39.9		225
Kordofan	30.2	10.0	37.6		50.5		82.2		272
			18.9		27.9		57.0		570
Darfor	20.8	4.9	28.5		47.8		69.3		333
			8.7		22.8		41.0		837
Sudan	43.8	18.3	39.7		55.8		90.3		206
			24.2		40.4		61.9		328

The gender parity index is a measure of the size of the girls' enrollment relative to that of the boys. From table 3 above we can see that **while the first three regions** of Northern, Khartoum and Central almost achieved gender parity in AIR, gender disparities remains to be apparent in the other three regions of Kordofan, Eastern , Darfor and in the country as a whole. There are 41 girls of eligible enrollment age who are left out of school in Darfor, 31 in Kordofan and 35 in the Eastern

region. At the national level we find that for every 100 boys of the seven-year age-group enrolled in the first year of primary only 69 girls are enrolled. This is almost the same pattern earlier observed in AIR and GER indicators and may be due to the same reasoning. The expansion of access to primary education has been a major concern since independence. Besides a push out for the elimination of illiteracy, expansion of primary education helps in broadening the human resource base that is more qualified to help in increasing productivity and improving the welfare of the population. The expanded vision of Education for All identifies *primary schooling as 'the main delivery for basic education of children outside the family'*. This 'World Declaration' stresses the importance of universal access and equity of that access. Gender disparities in access to primary education violate the whole essence of primary education and renders it short of achieving its intended national goals. Gender disparities in our country are high. Suitable measures need to be taken urgently to alleviate this and accelerate girls' enrollment. Detailed gender indices are found in Table 4 below:

Table 4: Trends of gender parity indices ³⁸

<i>Region</i>	<i>1961/62</i>	<i>1966/67</i>	<i>1982/83</i>	1989/90
Northern	50	71	99	93
Khartoum	78	79	84	97
Central	46	54	86	91
Eastern	47	64	84	65
Kordofan	33	50	55	69
Darfor	24	31	48	59
Sudan	42	61	72	69

If we just ignore the discrepancies resulting from source data of 1982/83, one notices an overall trend towards mitigating gender disparities. But still the pace of progress is so slow in the last three regions that special efforts need to be exerted towards that end.

Up to now we have been discussing expansion and equity of access to the primary education without paying much attention to the quality of education being offered. In fact there are so many such measures but for reasons of data limitations our analysis will be restricted

³⁸ Ibid

to pupil-teacher ratio which, if based on dependable data, will be a good indicator of educational quality.

Table 5: Pupil- teacher Ratio (p/t ratio) in schools by gender³⁹

<i>Year</i>	p/t ratio m f
1961/62	28.8
	44.4
1966/67	49.4
	50.8
1982/83	33.8
	34.1
1989/90	40.8
	34.3

The erratic nature of the ratios in this table indicates the quality of data from which they were calculated. These inconsistencies are a common feature characteristic to the Annual Educational Statistical Reports which are the only official source of data available, and this will show the amount of hardships any one will face if he ventures into trying to research that data with the hope of getting conclusive results. The bigger problem is that those reports ceased to be issued since 1990 and so the last official data on education available dates back to 19889/90. As the course curriculum of the primary (basic) schools is standardized all over the country, then the curriculum applied in any school will fully represent the whole. I picked one of the basic schools and got the following analysis about the composition of its course curriculum:

Current curriculum in the primary level

<i>Nature of course material</i>	<i>Details of content</i>	<i>% time</i>	<i>% subgroups</i>
Humanities	Religion	18	
	Arabic	28	65
	English	10	
	Others	9	
Sciences	Math	19	26

³⁹ ibid

	Sciences	7	
Student activities	Athletics & arts	9	9
Total		100	100

Governments have taken on the responsibility of unifying the different educational services by developing a national curriculum that focuses on the overall development of the child ; social, psychological, physiological and cognitive. In the 'Education For All World Declaration' it was mentioned that *curriculum aims to introduce creative play, social interaction, 'school readiness' play and stimulating activities such as story telling, songs and access to picture books.*⁴⁰ So, besides the target of learning the basic skills of literacy, and a good persuasive orientation with the bases of the mathematical principles, the curriculum should aim at promoting the child's intellectual, social, physiological and emotional development as a guiding principle. It is not for me to judge, but it seems to me that the composition of the curriculum of our basic schools as detailed in table (3) found above, is off bound relative to the principles laid down by the aforementioned 'World Declaration'. We see from the contents of that table that 65% of the overall course load is reserved for humanities, while only 26% is left to math and science, with as little as 9% for student activities which includes Arts courses. Such an imbalance in the course curriculum is apt to affect a student's social wellbeing, limit his base of educational attainment and consequently affect his future educational achievement. Reserving almost half of the course load for Arabic and Religion does not seem to be reasonable especially if we remember that a majority of those materials will have to be memorized by heart as mandated. The student will have to spend most of his time at home memorizing those materials at the expense of other materials and at the expense of a much needed playtime and socialization with his friends, both being vital to the buildup of his personality. It is worth mentioning here that things that are mandated to be kept by heart will definitely vaporize out of a child's mind as soon as the exams are over. So he will remain with as little real educational content which will accumulate to produce at the end a semi-educated university graduate with insufficient educational caliber.

The impact of these indicators will be greatly explained and verified by the introduction of certain measures of school quality and

⁴⁰ ibid

content. It is a common practice that schools of various levels prepare analytical data about the academic standing of its students as a by product of its annual exam results. Visit any school and you will find on its internal boards plenty of data that can be of great comparative value if put in a certain standardized statistical format. Such formatting will help produce what can be termed as *measures of school quality*.⁴¹

Measures of school quality—current status

Schools can use ‘accountability systems’ to measure school status as well as student achievement. The most common method of reporting school results is in terms of current status, i.e., reporting the school mean score for students in the final general exams. Schools within a state are ranked according to that. This kind of procedure is usually followed as a by product of the secondary school final exam results. Here schools are ranked in accordance with the percentages of their students who meet or exceed a certain performance standard or the percentage of students in each of several performance categories.

In fact, exam results of all grades are available at all school levels and could be utilized as a valuable source for assessing school quality and student performance. If those exams, and their grading systems, were sort of standardized state and countrywide, and if they were closely managed to ensure that schools treat them credibly and seriously as a measure of true performance, then they could be pooled to reflect comparative performance within and between states, and country wise. The following measures, indicators, could be of great value in assessment;

Measures of School Quality-Improvement Over Time

A preferable approach to measuring and reporting school achievement is to place greater emphasis on improvement than on current status. The procedure is to compare test scores between two years but for the same grade, for example third grade math scores in 1998 to third grade math scores in 1999. Such ‘improvement-over-time’

⁴¹ Reporting school quality in standards-based accountability systems by Robert I. Linn
CRESST POLICY BRIEF 3

comparisons based on successive groups of students in selected grades are reasonable for schools with consistent student populations.

<i>year</i>	<i>Fail</i>	<i>Pass</i>	<i>Good</i>	<i>Very good</i>	<i>excellent</i>	Total
1998	10.2	23.3	35.5	20.1	10.9	88
1999	7.2	13.3	40.5	25.1	15.9	102

The table above shows a hypothetical configuration of the final exam results of, say, the third grade for the years 1998 and 1999. It contains percentages of their performance at various levels of grading, and the total number of students in that grade for the two years. It shows an improvement of the third year grades of 1999 over that of 1998.

Longitudinal And Quasi-Longitudinal Reporting Methods

Another way to measure improvement is to track the performance of the students from one grade to the next. The approach using only students with scores in both years of the comparison is commonly referred to as a *longitudinal model*. It has the appeal that school is only held accountable for students who were in the school for the period between the first and second exams. Although this feature of the longitudinal approach may seem an advantage to schools, it has the clear disadvantage of excluding mobile students who change school from one year to the next and students who for some other reason are tested in only one of the years being compared. Therefore, the educational system is not held accountable for these students.

An alternative that avoids this disadvantage is to base the accountability on a comparison of the performance of all students in the school in, say fourth grade in 1999, with that of all students in that school who were tested in the fifth grade in 2000. This approach has been called a *quasi-longitudinal* approach. It has the advantage that all students in that school at the selected grades influence the results in a given year.

Both the longitudinal and quasi-longitudinal approaches require comparable tests across each grade level compared. For these measures to be used as a basis for within and between states comparisons, the annual tests should be subjected to a certain form of standardization. In such a case they will be powerful assessors of school quality, student achievement, and of the accountability of the education system.

Measures of Home Background Effects

There is a definite overriding effect of the socioeconomic status (SES) on student achievement. If students were classified in accordance with their SES status in standard pre-determined sub-classes, and exam results were classified accordingly, then we could measure SES effects and those results could be tabulated in as many forms of tables as needed to reflect the actual state of our world of education and to cover an important side of information that is normally missing and ignored by our education system. If these measures were correlated with other educational indicators like class size, pupil-teacher ratio, teacher qualification, etc., they would explain a lot and be of great help to policy makers. An example of these tabulations could be follows;

STUDENTS ATTAINMENT ACCORDING TO MODE OF LIVING

Mode of living	% dropout	% repeat	% pass	% higher performance	Total no. of students
Urban					
rural					
nomadic					
total					

STUDENT ATTAINMENT ACCORDING TO FATHER'S
EDUCATION

<i>FATHER'S EDUCATION</i>	<i>% dropout</i>	<i>% repeat</i>	<i>% pass</i>	<i>% higher performance</i>	Total no. of students
illiterate Primary or lower below secondary secondary high					
total					

STUDENT ATTAINMENT ACCORDING TO FATHER'S
OCCUPATION

<i>FATHER'S OCCUPATION</i>	<i>% dropout</i>	<i>% repeat</i>	<i>% pass</i>	<i>% high performance</i>	Total no. of students
deceased unemployed nomadic farmer craftsman employee private Business					
total					

Measures of Internal Efficiency

We propose that our eight-year primary education be subdivided into two four-year parts, a first half of four years and a second half of four years. The idea is to consider pupils who pass the first half as literates. Internal efficiency of the primary education is measured by the degree of incidence of dropouts and repetitions among the students. The

higher is the incidence of dropouts and repetitions, the lower becomes the efficiency of the education system.

Definitions and comments: ⁴²

New entrants: New entrants are students who come to school (the school system under study) for the first time. They may or may not have prior education.

Re-entrants: These are students who interrupt their schooling for sometime (a year or two or even more) and later rejoins the system. There is a problem here of how to distinguish a student who comes back to the system as a repeater re-entrant or as a promoted re-entrant. In the absence of good carried over records, preferably computerized, such information cannot be obtained from schools unless a good annual reporting system that includes that type of important data is decreed.

Promotion rate: is the rate at which students move from one grade to the next higher grade. This is obtained by comparing students in grade $g+1$, year $y+1$ with students in grade g year y .

Repeaters: Repeaters are students who attend the same grade they have attended the previous year. Traditionally a repeater is equated to 'failed' students. This creates problem because a student may have left school after spending, a semester or so in school. When he/she comes back to school next year or the following year, he/she is a repeater and must be treated as so to improve the outcome of the model. This is because that student has used the school resources (teachers, books, school space etc) for the period he/she has stayed in the system. Also these students have played the role of preventing others who could have come to school and possibly continue schooling up to the end of the cycle.

Dropouts: Dropouts are students who leave the system for the academic year under consideration. These have to be distinguished from transfers who change schools.

Graduation rate: the graduation rate is defined as the rate at which children complete the final grade of the cycle and are eligible to go to the next higher level.

Educational Inputs: Educational input is defined as the number of pupil-years spent in the cycle. Educationists agree that the quantity of

⁴² An open secret behind inefficiency in the first half of primary grades—Ethiopia, by Tegegn Nuresu Wako-(2001): Report on findings of a pilot study, Awasa.

educational inputs: teachers, classrooms, furniture, books etc. rises as the number of students increase and the years pass by. Hence, they use pupil-years as a measure of inputs. Again this is a crude but non-monetary measure and a convenient one to use [UNESCO:1998].

Educational output: Educational output is defined as the number of pupils who successfully complete a given cycle. This definition does not address the fundamental question of education “how much has a child learned in school?” nor the role of education in the economy of a country or social, political, and cultural life of a society [UNESCO:1998].

A comparison of the student attainment throughout the first four years of the primary education shows out the status of its internal efficiency and measures its success in achieving the minimum goal of producing literate students. It would be good if the final exams at the end of this first sub-group was prepared and managed as a national exam. It needs not be a comprehensive exam that covers all curriculum. Since it may be considered a *literacy test*, then its coverage could be limited to the basic elements of literacy: reading, writing and mathematics. This would furnish a reasonable basis for comparative analysis and pinpoint areas of weakness whether at the school, district, state or country levels. The kind of data needed for the calculation of the above mentioned indicators is aplenty in school records. What is only needed is the creation of a good system of tabulation that would suit the needs of comparative statistical analysis. Tables like the ones proposed above are just examples of how school records could be prepared to be of great benefit to school administrators, policy makers and the research community. It is worth mentioning here that this simple re-editing of school reports would resemble a great breakthrough without incurring any additional expenses. It is a simple rewriting of these reports in a format that needs to be standardized countrywide and included in the annual statistical report of education.

Educational statistical data

The main and only, official specialized sources of educational data of Sudan are the Statistical Reports that used to be published annually by the Ministry of Education. They date back to the late fifties and were reformatted in 1961 to include a wide variety of data and furnished the basis for reports of the years that followed. Then, the size of the whole of the educational system was small (e.g. The total number

of primary schools for girls was 292 and those for boys were 676 in the country as a whole). The founders of the of the system of those Statistical Reports were originally experienced, highly educated teachers who were especially selected for the job and were enabled with high levels of training in the fields of demography and statistics. This, together with the small size of the educational system at that time, enabled them to lay down the beginnings of an efficient educational statistical database. Eventually they started producing annual statistical reports that covered the various aspects of the educational process and could be used as a credible and efficient source of educational data. But, as it usually happens to good things in our country: they start strong, stabilize, and then take on the continuous process of deterioration towards oblivion. That was typical in the case of the system of educational statistical reporting. As of the seventies, their coverage started to dwindle with a tendency toward detailed reporting of a limited, and continually decreasing number of variables of the educational process. Their content does not seem to be built on any pre- designed strategy. The reports ended up as just an accumulation of statistical data that lacks consistency, sufficiency and efficiency. Its continuous change of content and multiplicity of errors in the data for various years reported in subsequent annual reports renders them useless as a base for in-depth, conclusive statistical time-series analysis or for producing a dependable indicator system whose trends could conveniently be used to tap up educational parameters and assessing progress of the educational system. This deterioration may mainly be due to:

- lack of trained and knowledgeable staff,
- marginalization of the role of statistics and unawareness of its uses,
- lack of a research initiating culture in our system of civil service
- and generally an apparent lack of strategy for assessing progress or for availing the necessary tools for evaluation of how things are going in real life.

The sad fact is that those annual reports were stopped from being issued since 1990. The thing is, those reports, though deficient, give a kind of light on the state of affairs that should better be kept on. This action resulted in the total loss of any kind of official data for more

than a decade and to the interruption of the serial reporting and subsequent analysis of any aspects of education however few and limited.

The world of today is living the information age where a few numbers could be more expressive than tons of descriptive reporting and loads of talk. The world of today understands and deals with numeric data, and to relate and deal with it we have to build up an efficient system of reliable data to back up our talk and prove the credibility of our case. Building such a system is not the sole responsibility of statisticians but it rather results from the persistent contributions of all concerned parties. If the role of statistics is recognized, then an initiative need to be taken by the top echelons of educational policy makers to bring up those concerned parties together and to whistle the start. These won't start from scratch, for we already have a long history of annual statistical reporting. We just need to restart or redesign our reports within the terms of a new strategy. The mission of the different concerned parties is to lay down the basic elements of that strategy and write down its components. It came down to my knowledge that the Ministry of Education is planning to build up a network of computer- centers in all states to be connected with a center at its headquarters in Khartoum. The target is to build up a continuous, and hence updated, flow of educational information and the formulation of a central and, as many, regional databases. That plan, if achieved, would resemble a great breakthrough. But for it to work and be of real use a lot of preparatory work is needed to lay down and standardize the composition of that information system. In my field of concern, a new statistical information system should just be an improvement to the old system and an expansion of its content. Continuity of reported data, its consistency accuracy and the span of its coverage are the most needed basic elements for conclusive analysis. The contents of the new educational statistical report should comply with the recommendations of the Jomtien Declaration of the EFA⁴³ and contain the data necessary for building up statistical models for assessment of progress. A task force of expert people of related specializations is needed to work on that and forward their recommendations. Besides, a central board for educational statistics, directly linked to the Minister of Education, should be formulated to undertake the job. Its terms of reference should extend, besides the annual report with its general nature, to the collection of

⁴³ ibid

more intensive data about areas of special interest through use of other means of statistical data collection as detailed below. The scope of its work shouldn't only be restricted to students and schools, but should be extended to cover pre=enrollment student backgrounds, and the fate of a student when he finishes schooling.

Types of educational data

There are three major sources of data that be used to build up the content of an educational statistical database:

Annual data

Similar to all censuses, the annual reporting system can not contain lots of detailed data that covers the multiplicity of variables that compose all of the educational process. For reasons of convenience and practicability it can only cater for basic elements which are themselves aplenty. It would be conveniently sufficient for them to include data needed for the calculation of as much indicators of educational performance as could be. They wouldn't, for example, be expected to contain detailed data about student background and the comparisons that can accrue from it. Including such details in an annual questionnaire renders it hazardous and expensive to fill, administer, collect and process. Other methods, that are detail oriented, can be used to supplement the annual data.

Surveys

Random sample surveys are a credible and dependable means of data collection. A sample is a data set representing a portion of a population. Usually a sample is intended to represent the population from which it comes and to make inferences about it. To be acceptable for statistical analysis, a sample must , not only, represent the population but must also have a measurable reliability. In other words, in addition to the desired estimates of the population characteristics like totals, averages, etc. , the sample should give measures as to the precision of those estimates. Sampling surveys are preferred because they are economic, time saving, easier to administer to minimize errors and can supply as many detailed information as needed. So, the concept of a

sample as opposed to a *population*, is very important. A *sample* is just some part of a population selected according to some rule or plan. If we are dealing with the entire population, our statistical work will be primarily descriptive. On the other hand, if we are dealing with a sample, the statistical work will not only describe the sample but also provide information about the sampled population.⁴⁴

In fact results from sample are as credible as that of a census, and could conceivably be more accurate. This procedure, though so feasible, is not yet used for collecting educational data. The only present source of educational data is the annual report, and that is not enough. A lot of data about details is not, and can not be, included in that report (census) and can more conveniently be collected through a small sample of schools of varying levels and locations. This would enable an enlarged and more varied database to provide for the calculation of all needed indicators and for building statistical models to compare relations between the different aspects of the educational process and test significance of their impact. In fact this is the procedure that can be used to find out the impact of a student's home background on his academic achievement as an example. A dependable and more varied data set is a necessity for in-depth conclusive statistical analysis.

Experimental data

Data from experiments are considered sample data, and the introduction of experimentation in education is of utmost importance. Experiments are usually used to study certain programs or cases or what are termed as case studies. Rather than using a large sample to study a limited number of variables, case study methods involve an in-depth, longitudinal examination of a single instance or event.⁴⁵ It is a systematic way of looking at what is happening, collecting data, analyzing information, and reporting the results. The product is a sharpened understanding of the instance happened as it did, and what might be important to look at more extensively in the future research. Thus, case studies are especially well suited toward generating, rather than testing, hypotheses.

⁴⁴ Bernard Ostle (1966) Statistics In Research, second edition

⁴⁵ Lynn Davey (1991) : The Application of Case Study Evaluations, ERIC Clearinghouse on Assessment and Evaluation

Types of Case Studies

Illustrative case studies: are descriptive and utilize one or two instances to show what a situation is like. This helps interpret other data, especially when it is believed that the recipient population know too little about a program. These case studies help familiarize things and give a common language about a topic.

Exploratory Case Studies are condensed case studies, undertaken before implementing a large-scale investigation. Where considerable uncertainty exists about program operations, goals, and results, exploratory case studies help identify questions, select measurement constructs, and develop measures. The greatest pitfall in the exploratory study is prematurity; the findings may seem convincing enough to be released inappropriately as conclusions. Other pitfalls include the tendency to extend the exploratory stage, and an inadequate representation of diversity.

Critical Instance Case Studies examine one or a few sites for one of two purposes. A very frequent application is the examination of a situation of unique interest, with little or no interest in generalizability. A second, rarer, application entails a highly generalized assertion which is called into question, and which we can test by examining one instance. This method is particularly suitable for answering cause-and-effect questions about the instance of concern.

Program Implementation Case Studies help discern whether implementation is in compliance with its intent. These case studies are also useful when concern exists about implementation problems. Extensive longitudinal reports of what has happened over time can set a context for interpreting a finding of implementation variability. In either case, generalization is wanted. A requirement for good program implementation case studies is investment of sufficient time to obtain longitudinal data and breadth of information. Multiple sites are typically required to answer program implementation questions. This imposes demands on training and supervision which are needed for quality control. The demands of data management, quality control, validation procedures, and analytic model (within site, between sites, etc.) may lead to cutting too many corners to maintain quality. Programs , like model schools at the secondary level that is being implemented for some time is an experiment that need to be assessed by such an application.

Program Effects Case Studies can determine the impact of programs and provide inference about reasons for success or failure. Like the program implementation case studies, this evaluation procedure usually require generalizability and, for a highly diverse program, it may be difficult to answer the questions adequately and retain a manageable number of sites. There are methodological solutions to this problem. One is to first conduct the case studies in sites chosen for their representativeness, then verify these findings through examination of administrative data, prior reports, or a survey. Another solution is to use other methods first. After identifying findings of particular interest, case studies could then be implemented in selected sites to maximize the usefulness of the information. New course curriculums, or adjustments in the ongoing curriculum, need to be subjected to such critical evaluation to test their impact and assess their success or failures.

Cumulative Case Studies aggregate information from several sites collected at different times. The cumulative case study can be retrospective, collecting information across studies done in the past, or prospective, structuring a series of investigations for different times in the future. Retrospective cumulation allows generalization without cost and time of conducting numerous new case studies; prospective cumulation also allows generalization without unmanageably large numbers of cases in process at any one time. The techniques for ensuring sufficient comparability and quality and for aggregating the information are what constitute the “cumulative” part of the methodology. Two features of the cumulative case study are the case survey method, used as a means of aggregating findings, and backfill techniques. The latter are helpful in retrospective cumulation as a means of obtaining information from other accessible sources. Opinions vary with respect to the credibility of cumulative case studies for answering program implementation and effects questions. One authority notes that publication biases may favor programs that seem to work⁴⁶. Others are concerned about the problems in verifying the quality of the original data and analysis.

The case study, thus, is a method of learning about a complex issue through extensive description and contextual analysis. The is an articulation of why the instance occurred as it did, and what may be

⁴⁶ Berger, Michael A(1983),. “ Studying Enrollment Decline via the Case Survey”. Educational Evaluation and Policy Analysis,

important to explore in similar situations. Evaluators considering the case study as a design for evaluation must first decide what type of evaluation question they have and then examine the ability of each type of case study to answer it. The crucial next step is in determining whether the methodological requirements of the chosen case study method can be met in the question at hand.

Though we have our own experience in the field of educational statistical information systems, and since we are being faced by inherent infrastructural problems that need to be dealt with promptly for the good of the whole of our educational system, it would be of great help in that context to look into the experiences and practices of some other countries and see if we can make use of them. Zambia is one those countries that sought the help of NESIS to strengthen its education statistical information system. A document about that experience gives a step-by-step account of the implementation of the NESIS program, commencing with the diagnosis of the information system and proceeding through the action that Zambia has taken so far.⁴⁷

The Education Problem

The education system in Zambia has been characterized by continuous growth and increasing demand. The absence of a sound information base meant that it had become more difficult to plan, manage and monitor the education delivery system.

The lack of access to a pool of information had resulted in a situation where several departments were engaged in uncoordinated data collection. Often, the type of data collected met just the internal information needs and was not for general consumption. This resulted in duplication of efforts and repetitive demands on the responding school offices. Furthermore, since the data collected was gathered from a variety of sources which were stored separately, each department and user only obtained a segmented view of the overall situation. Worse still, the Ministry had no capacity to process its own data. This resulted in delayed submission and processing of data, incomplete or inaccurate information and poor retrieval.

⁴⁷ Manasseh Nkamba, (1996). “ strengthening NESIS : the Zambian experience”, National Education Statistical Information Systems.)

All these problems created difficulties with planning, policy formulation, management of the system, accountability and decision making. In response, the Ministry of Education, in collaboration with the NESIS program of the ADEA/ WGES⁴⁸ and the Swedish International Development Agency (Sida), developed the Zambian National Education Statistical Information Systems (NESIS) project.

Context of the NESIS project

As a result of its participation in the NESIS project, Zambia's current education policy focuses on two areas:

- strengthening the information base necessary for the routine management of the education system
- strengthening the information base for the development of education policies
- The policy states that “in order to facilitate planning, monitoring and supervision at all points of the system, the Ministry will revise the existing data collection instruments and establish and maintain a comprehensive Education Management Information System.”⁴⁹

Implementing NESIS

The NESIS project evolution involved four steps:

- establishing NESIS management structures
- conducting a diagnostic survey in order to identify the information gaps among users and producers of information at both central and regional levels.
- holding an orientation workshop for key people involved in the information management process
- drawing up plans and organizing activities to address the priority areas identified in the diagnostic survey

A policy level Advisory Committee was formed in order to manage the NESIS project. This committee consists of heads of major producer and consumer departments at the Ministry of Education

⁴⁸ ibid

⁴⁹ national policy on education, Ministry of Education, Zambia, 1996 p. 131

(MOE) and elsewhere. Then there is a national team that oversees the implementation of the project. The team can set up technical committees comprising experts from teacher training colleges, universities, schools, the Central Statistics Office (CSO) as well as the MOE to review specific priority problem areas identified in the diagnostic survey. The Advisory Committee reports its policy recommendations to the Permanent Secretary and the Minister. After having set up the Advisory Committee and the national team, the next step was to identify the information gaps among users and producers at both central and regional levels. This meant undertaking a number of activities:

- the national NESIS team adapted the NESIS diagnostic survey questionnaires to the country's needs
- the questionnaires were then sent to the producer and consumer departments in the MOE , institutions or units outside the Ministry (like the Examinations Council of Zambia) and the nine provincial heads
- data from the questionnaires, submitted by various departments, was analyzed
- an orientation workshop was organized for the key role players at both national and provincial levels. The results of the data analysis formed part of the workshop's discussion. The main objective was to reach a consensus on the nature and degree of the problems existing in information management, so as to identify major problem areas

The diagnostic survey helped to identify major bottlenecks and six priority areas for action:

- the development and strengthening of record keeping at all levels
- improving the quality and content of statistical information
- staff development along the entire information cycle
- creating an information system, the quality and content of which is linked to user demand
- computerizing the information system
- strengthening the use of statistics at all levels

School records and data collection instruments

In an attempt to improve the Education Management Information System (EMIS), it was decided to focus initially on the development of effective school records and school survey collection instruments. The reasoning behind this was that:

- accurate and good quality data depends on the sort of records that exist at the source
- the kind of information generated depends on the variables used in the data collection instruments

The main actors involved in the development of school records were:

- Policy-makers, including educational planners, school inspectors, etc.
- National experts from specialized units, such as the Central Statistics Office, the Career Guidance and Counseling Unit, the Examination Council of Zambia and the University of Zambia
- The immediate implementers, which included Education Officers in the districts, school heads and classroom teachers, as well as record managers and clerks within MOE
- The process of developing school records included the following:
 - reviewing existing school records
 - modifying and adapting some of the school records
 - developing new school record forms for those areas which previously had never had any

There were four categories of records: pupil, teacher, material and finance.

<i>Type of record</i>	Comprising
Pupil records	Application form for enrolment into grade one(the first grade) Transfer request(primary/ secondary) Admission register Basic education pupil record cards Secondary school pupil record card

Teacher records	Class register Mark schedule(primary/ secondary) Assessment form School report Teacher record cards
Material records	Inventory forms Stock books Hand- over certificate forms

Finance records revenue forms

Expenditure forms

After the development of the school records and the review of the data collection instrument, the next issue that was tackled was insuring the continued and proper use of the school records and the data collection instruments. To this end, two training manuals, one on the development of school records and the other on the collection of data, were produced.

Two specialized committees worked on the production of these manuals. As had occurred with the compiling of the school record forms, both training manuals were developed with substantial inputs from various key players and experts, with a different author for each text.

The school records training manual aims at training:

- heads of schools on how to administer, manage and use various school forms relating to students, teachers, materials and financial records
- teachers on how to administer, manage and use pupil and materials record forms
- other officers and records officers who may be administering, managing and using the pupil, teacher, material and financial records

This manual has the following components:

<i>Introduction to school records Management</i>	This section gives an outline of the practice of records management in schools. It includes definitions of records management concepts as they have been used in the manual.
Training modules	There are four groups: pupils, teachers, materials and finances
Index	Provides a guide to finding important references in the manual, thereby making it an easy-to-use reference tool

The data collection training manual contains data collection instruments as well as guide lines on the use of such instruments as part of the training component. Unlike school records, which are limited to maintaining information within the school, the data collection instruments serve to:

- capture information on pupils (such as age, grade, nationality, gender, dropouts, etc.) by means of the primary/ secondary annual returns
- collect the details on teachers within a school (qualifications, gender, etc.) using school staff returns
- collect information on school buildings and equipment (state of the facilities, i.e. furniture, kind and amount of equipment, etc.) using the building, furniture and equipment returns
- obtain details of the financial situation of the education system via the expenditure and revenues forms

The manual also includes information on how summaries can be made by using tally sheets.

Testing the data collection instruments

All the school records and data collection instruments were tested in 300 primary and secondary schools within Zambia between February and May, 1996. This was done using a three- stage, stratified, systematic sample design.

The first sample was conducted in the provinces. Zambia's nine provinces were divided into three zones. In total, five provinces were selected. Within each province, the tendency was to select two districts, one rural and one urban. The urban districts which were chosen were all

provincial centers, while the rural districts were randomly selected using circular systematic sampling or simple random sampling. This was done by applying random number tables, namely, the first random digit appearing, between 1 and the serial number of the last rural district. Exceptionally, in the Copper-belt province, two urban districts were selected due to the highly urbanized nature of the province and the large number of schools.

In each urban district (provincial center) and rural district, three strata were created for the different kinds of schools, following which sample schools were selected. The strata were primary schools (grades one to seven), secondary schools (grades Eight to Twelve) and basic schools (grades One to Nine).

In all, ten secondary schools, fifteen primary and five basic schools were chosen from each selected district within the targeted provinces. Using circular systematic sampling, the following total number of schools was covered: $10 \times 2 \times 5 = 100$ secondary schools, $15 \times 2 \times 5 = 150$ primary schools and $5 \times 2 \times 5 = 50$ basic schools. In this way, a sample size of 300 schools was reached. This comprised approximately 10 per cent of government and grant-aided schools in Zambia.

After the sample schools had returned their comments, the training manuals, record forms and data collection instruments were revised. The final versions are expected to be in all schools and Teacher Training Colleges by May 1997. (Reminder: this report is written in 1996).

Other NESIS-related activities

In addition to the development of school records, the review of data collection instruments and the compilation of training manuals on school record development and data collection, other NESIS activities have included:

- the participation of members of the NESIS national team in international conferences aimed at sharing experiences .
- commencing the computerization of the MOE, starting with the Ministry headquarters followed by provincial headquarters and finally the districts

- training data entry operators/ statisticians at national and provincial levels. This training will be extended to district personnel once the districts have been computerized
- training education planners, based at headquarters, in the use of computers for information analysis, etc. The training will eventually also include the regional planners
- the participation of a member of the NESIS national team in four educational development indicators and database training workshops. Two were organized by NESIS/ UNESCO, and the other two were organized by the International Institute for Education Planning (ILEP)
- the NESIS project has helped to foster partnerships between various information users and producers, both within and outside the MOE
- fostering a spirit of project ownership by way of a participatory approach in designing the data collection instruments, the school records and the training manuals

Since education reforms are increasingly focusing on decentralization, there has been an extension on the usage of information. In this way, the NESIS project is helping to strengthen planning, management and monitoring of the education system at local levels.

At both national and local levels, the information generated by the NESIS project is helping to rationalize the allocation of resources while enhancing accountability at different levels of education. The training manuals are currently being used to cover the records management component in the Head Teachers Management Training Program financed by the World Bank. The training manuals are soon to be institutionalized in all pre-service teacher training colleges so that record management will be an integral part of the training courses there. This is to ensure that teacher graduates from these colleges will know how to use records as well as recognize the importance of maintaining good records.

The Way forward

The following activities are envisaged for the future:

- computerization of national, regional/ district offices
- development of user-friendly software package to facilitate data analysis
- continuous training of school heads and teachers in: the administration of the various school records forms for pupils, teachers, and the others; the use of information from pupil and other records for planning purposes
- the Government is planning to establish the Education Sector Investment Program, in which four Ministries (Education, Science and Technology, Sport, Youth and Child Development and Community and Social Services) are to work together on the issues of education and training (in this way the NESIS experience will also be used by other Ministries).

The role WGES-NESIS Program

Through the ADEA Working Group on Education (WGES), Zambia was selected by its NESIS program as a pilot country for developing and testing NESIS modules on school records management and annual school survey. Participation in the NESIS-WGES program, based on the principle of full country leadership, ownership and partnership, facilitated Zambia's development and improvement in management in several ways:

- the most important contribution was the orientation of the policy-makers and technical specialists on the organization and management of self-managed development of national education statistical information systems. This gave a vision of the goal, a manageable strategy and a deep conviction of its feasibility that initiated the whole development process.
- The WGES-NESIS provided technical support and advisory services in the design, organization and management of the diagnostic surveys and national action formulation. The findings of the diagnostic surveys gave all involved a clear understanding of the information needs of policy-makers and

- managers as well as the problems that must be overcome. The result is an action plan that is grounded on a systematically conducted needs analysis and feasibility study.
- The NESIS organizational framework for cooperation between information consumers and producers was introduced, by the establishment of the national advisory committee, consisting of policy-level representatives of the major consumer and producer departments and institutions in the country, and the technical team, consisting of specialists from MOE and other institutions, for implementation of the action plan.
 - The WGES-NESIS program has become an important source for a steady stream of new ideas, application tools, good practices and innovation pertaining to records management, annual school survey, indicators for monitoring education development, database and data-processing, etc. These have had direct impact on the increased productivity, efficiency and quality of our statistical products and services.
 - The WGES-NESIS has also provided direct and indirect training, by way of demonstration effect, in professional skills of information specialists. These skills encompass such areas as oral and visual techniques, organizing technical teams, report design, writing and statistical presentations and preparing briefing meetings. These professional skills as well as the resulting policy relevant, improved quality and punctuality of information products and services have enabled the planning and statistics staff to play an important role in the policy process.
 - The cooperation within and among the WGES-NESIS network of member agencies (UNESCO, SIDA, UNICEF, the World Bank) helped to mobilize resources and expertise for going to scale, beyond the pilot project. The WGES-NESIS team played a key role in this process. This enabled clear and strong links between the pilot projects and the long-term development of the national information systems.
 - In sum, the role of the WGES-NESIS program has been far reaching, by initiating a self-managed developmental process

that will have long-lasting repercussions throughout the Zambian education system.

This valuable document was brought for its totality, with a few omissions, for the following reasons:

- It coincides with, and gives details of the views that I have tapped frequently in this paper about the importance of establishing an efficient educational statistical information system.
- While we remain lagging behind, other countries have taken the issue seriously, and sought help to initiate their own systems
- So many regional and international conferences were held on this issue and it is beyond any doubt that we were represented in them, but it seems to have ended at that.
- The paper makes it clear that there are regional and international organizations (of which we are member) which can provide both financial and technical help to member countries who seek it.
- It is high time that we do away with our culture of disregard to statistical information. Availability of an efficient educational information system will help us do research and assume our natural role in the international community as active contributors and leaders to be reckoned with instead of just being attendees whose presence can hardly be noticed.

Accountability and Assessment Systems

Education policies in our country has swung radically instead of making minor and gradual changes. While politics is the major factor behind those radical changes, research should be given priority in assessing the effectiveness and impact of those policies. The only means of national assessment available at hand so far at the national level are the secondary school exam results. The swing in the assessment policy, emerging from those exam results, is toward greater emphasis on accountability, i.e., holding schools responsible and providing rewards and sanctions, moral at least, for the performance of their students in that national exam. This is a move away from using assessment as an impartial monitor of progress and a means of diagnosing the system's

strengths and weaknesses. In fact, student secondary exam results could offer a great opportunity for research and analysis to find out the real factors that caused some schools, or regions, to fare less than others. It is also a fact that the actual, unadjusted, examination scores could be used as our better attainable means for assessing the impact of the new policy changes introduced onto the education system as a whole and on the kind and quality of student attainment. Such research would furnish some sound grounds for holding schools accountable or accrediting them for their performance. It would indicate to those seriously worried policymakers who hopefully, are looking for answers as to why things are not going the way they were planned, the right path to dig.

There are several goals for testing for promotion or graduation of a student group. These include setting high standards, raising student achievement, ensuring equal opportunity, fostering parent involvement and increasing public awareness and support for schools in their communities. However, testing may have negative consequences for individuals, so policy makers should be sensitive to the balance between individual and collective benefits and costs. By that, we refer directly to the official practices of celebration and accreditation that used to be followed when publishing results of the secondary school exams. An important question needs to be asked here: do those exam results give a true measure of the state of knowledge of the students?

The socioeconomic development of a country is a function of the quality of its education system. Quality education will produce quality graduates who, most probably, will have the ability to grasp, improve and develop on the present state of affairs in their particular fields of specialization, while a poor system will definitely end up with barely educated doctors, engineers, teachers, etc. , who will only add up to the speed of deterioration towards oblivion.

It is time to face up to the facts that show us the true nature of our world. Instead of adjusting the exam results to reflect an artificially high level of performance that doesn't exist, and thus evading accountability, I presume we need to show those results as they are and help ourselves by serious diagnosis. In our state of lack of a dependable educational statistical database, the actual examination results will almost surely, give us a rare and renewable opportunity to compensate for that deficiency and start an in-depth comparative analysis.

The National Center for Research on Evaluation, Standards, and Student Testing (CRESST) of the University of California, Los Angeles

in the USA, commissioned a committee to look into the principles of testing. Of that committee's findings we picked four principles of test use⁵⁰ that is relevant to our situation:

- Tests have validity only in relation to specific purposes.
- Tests are not perfect, and neither are the alternatives.
- No high-stakes educational decision about a test taker should be made solely or automatically on the basis of a single test score; other relevant information should also be taken into account.
- Neither test scores nor any other kind of information can justify educational decisions that are not beneficial to the students.

That report included six findings and recommendations which are:

- Accountability for educational outcomes should be shared among states, school districts, public officials, educators, parents and students; not by students alone.
- Tests should be used for high-stakes decisions about individual mastery only after students have been taught the knowledge and skills on which they will be tested.
- consequences of high-stakes testing are often” either-or “, but this need not be the case; tests and other information can lead to early diagnosis and effective intervention when students have learning problems.
- Some educational practices are typically bad for students, including placement in low track classes (schools in our case) and retention in grade; tests should not be used for such purposes.
- All students are entitled to sufficient test preparation; e.g., familiarity with item format and appropriate test-taking strategies. At the other extreme, educators should avoid narrowly teaching to the test.
- High-stakes testing programs should include a well designed evaluation component, and the consequences of high-stakes

⁵⁰ 1998 CRESST conference proceedings on Comprehensive Systems for Educational Accountability

assessments should be measured for all students and major sub-groups of students.

Examination results do tell lot of things about the quality of education and about the student's academic attainment. They will help pinpoint areas of weakness and prompt educators to try and remedy them, highlighted by the relative educational and student background indicators. Entering major adjustments on the actual test scores to upgrade them wouldn't help achieve that purpose. In fact such changes which are entered due to some political reasoning and pressures, resemble a menace to the credibility of the whole of the educational process and its future prospects for many reasons some of which are:

- it distorts the actual score data and thus diminishes its power of detecting weaknesses of the educational system regionally and at the national level.
- It inflates the egos of students who, in fact, are poorly educated. Poor education results in disillusioned semi-educated engineers, physicians, teachers, etc., and that cycle of deterioration will multiply if not checked and proper measures taken to correct its trend.
- It would reflect negatively on the educational guidance of the future generations and their motivation to work hard. Their response may be like: if a weak student can get scores higher than his actual attainment, then why would he or the others take things seriously. A noticeable phenomenon among the youth of today is that they mostly won't bother to exert the effort needed to perfect whatever jobs they were supposed to do.

Measure of the Examination Components and Scaling to Adjust Them⁵¹

Educational examinations usually comprise a number of components, be they different papers, sections within a paper, or individual questions. These components will contribute to the results of the examination as expressed in the candidates' final rank order,

⁵¹ Robert M. Adams and John Wilmut,(1981), Associated Examining Board, Aldershot, UK, *The Statistician*, (1981), Vol. 30, no. 4

established by the aggregate of the component marks. It is clear, however, that components will not all contribute to the same extent. The first paper of an examination may, for instance, achieve greater weight in determining the final results than another paper, though they were originally meant to have equal contributions. Allocation of a higher number of marks to one component than another will not necessarily result in that component gaining more weight. Neither does the weighting depend on the actual numbers of marks gained for either component, or some average thereof. This is simply because the final rank order of the students will not be affected by the addition of an arbitrary but constant number of marks to each student's actual score in either component. It is also fairly clear that one property of the marks on a component that determines its contribution to the aggregate is their dispersion. This may be illustrated by the following table:

Marks and ranks for six candidates for a two- paper exam

<i>Paper one</i>		<i>Paper two</i>		aggregate	
mark	rank	mark	rank	mark	rank
72	(1)	12	(6)	84	(6)
71	(2)	16	(5)	87	(5)
70	(3)	20	(4)	90	(4)
69	(4)	24	(3)	93	(3)
68	(5)	28	(2)	96	(2)
67	(6)	32	(1)	99	(1)

It can be seen from this table that the marks of paper two, being more dispersed than the other, dominated the aggregate so that the final rank order is the same as that of the paper two marks alone, in spite of the inversion of the rank order of the marks of paper one. This has led several authors, among them Forrest (1974)⁵², to define the achieved weight of an examination component as proportional to the standard

⁵² Forrest, G.M.(1974),The presentation of results. In Techniques and Problems of Assessment (ed. H.G Macintosh) Edward Arnold, London

deviation of the marks gained for it. Thus, if the k components of an examination of respective standard deviations s^1 , then achieved weight of component m , W_m , expressed as a fraction of unity, is given by k

$$W_m = \frac{S_m}{\sum_{i=1}^k S_m}$$

The decisive weakness in this definition is that the resulting achieved weights are not additive. Using the subscript $i + j$ to denote statistics arising from the aggregation of the marks for the i th and j th components, we have

$$W_{i+j} \neq W_i + W_j,$$

because unless the components are perfectly positively linearly correlated,

$$S_{i+j} < S_i + S_j$$

Thus, it becomes clear that any satisfactory definition of achieved weights will have to take account of the several inter-component correlations, as well as the dispersion of the components' marks if the resulting weights are to combine additively.

The k components-with- aggregate coefficients of correlation can be used as measures of the weights achieved by several components, but apart from a simple ordering of components according to the size of this coefficient, it is impossible to make any other quantitative judgments about their relative weightings because of the non-linearity of the scale of correlation coefficients.

Component- with- aggregate Covariance

As is shown above, a suitable measure of an achieved weight of an examination component might be based on a measure of the degree of association between the marks achieved by candidates on that component and their aggregate marks. Another such measure is the covariance, which has the additional property that the variance, $S_t = \sum_{i=1}^k S_{it}$ (1)

$$T=1$$

This can be established by noting that:

$$S_{it} = \sum_{j=1}^k S_{ij}$$

where S_{ij} is the covariance of the marks for the i th and j th components, and $S_{it} = S_i$, the variance of the component i marks.

Dividing (1) by S_i , we obtain

$$1 = \sum S_{it} / S_t \quad (2)$$

hence expressed as a fraction of unity, we have that W_i , the achieved weight of component i is given by

$$W_i = S_{it} / S_t \quad i = 1, \dots, \kappa \quad (3)$$

From (2) it can be seen that we have the desired additive quality that

$$W_{i+j} = W_i + W_j$$

This definition has, in fact, been discussed by Richardson (1941)⁵³ and has been recently used by Fowles (1974)⁵⁴ in her study of weighting in CSE examinations, and by Willmott and Hall (1975)⁵⁵ in a study of O Level GCE examinations.

The achieved weights may be multiplied by the aggregate available marks for the examination; and the resulting measures can then be compared directly with the marks allocated to the various components if it is thought that these allocations reflect their intended weights. The table below shows hypothetical results for a three-paper examination.

Achieved weights in a three-paper exam

<i>Paper</i>	<i>Marks Allocation</i>	<i>s.d.</i>	<i>Covariance with aggregate</i>	Achieved weights	
Written paper	100	21.47	534.42	0.83	124
Coursework	38	5.20	90.91	0.14	21
practical	12	2.29	19.21	0.03	5
	150		644.54	1.00	150

⁵³ Richardson, M.W. (1941). The combination of measures. Appendix to P. Horst, The Prediction of Personal Adjustment, SSRC, New York

⁵⁴ Fowles, D.E. (1974), The CSE : Two Research Studies. Evans / Meuthen Educational, London

⁵⁵ Willmott, A.S. and Hall, C,G, (1975), O' Level Examined: the effect of question choice, Macmillan Educational, London

The variance-covariance matrix for the examination components is:

$$\begin{matrix} 451.14 & 61.59 & 11.69 \\ 61.59 & 27.03 & 2.29 \\ 11.69 & 2.29 & 5.23 \end{matrix}$$

the calculations show that there is a considerable difference between the desired weights (as expressed by the component mark allocations) and the achieved weights.

Scaling to Achieve Intended Weights

Having established that there are discrepancies between the achieved weights and intended weights of the components of an examination, we may want to adjust the component's marks to remove the discrepancies. This can be achieved by simple linear scaling of the candidate's component marks. Suppose every candidate's mark for component i is multiplied by α_i for $i = 1, \dots, \kappa$. Then the covariance of the scaled component marks and the aggregate of the scaled marks, S_{it}° is given by

$$S_{it}^{\circ} = \sum_{j=1}^{\kappa} \alpha_i \alpha_j S_{ij}$$

Where S_{ij} is defined to be S_j squared.

The achieved weights resulting from the scaling, W_i are, following (3), as fractions of unity.

$$W_i = S_{it} / S_t \quad i = 1, \dots, \kappa$$

Where (S_t°) is the variance of the aggregated scaled marks, and is given by

$$(S_t^{\circ}) = \sum_{i=1}^{\kappa} \alpha_i^2 S_t + 2 \sum_{i=1}^{\kappa-1} \sum_{j=i+1}^{\kappa} \alpha_i \alpha_j S_{ij}$$

If the intended weights also are expressed as fractions of unity, say Ω_i , then

$$W_i = \Omega_i \quad i = 1, \dots, \kappa \quad (4)$$

Is a system of κ simultaneous homogeneous second-order equations in α_i . Since only the relative sizes of the α_i are important, one of them can be fixed and the system (4) reduced to $\kappa-1$ simultaneous second-order equations by dividing successive equations thus:

$$\frac{W_i \Omega_i}{W_{i+1} \Omega_{i+1}} = \frac{\alpha_i}{\alpha_{i+1}} \quad i = 1, \dots, \kappa-1$$

$$\alpha_k = 1 \text{ say.}$$

hence the k-1 equations

$$W_i \alpha_i - W_{i+1} \alpha_{i+1} = 0 \quad (5)$$

which can be solved using any suitable algorithm. It is desirable to avoid loss of discrimination in the examination, so that all the α_i should be greater than unity. This can be achieved by dividing each of the α_i in turn by the smallest value of α_i resulting from the solution of (5).

Returning to our previous example, the two equations (5) in the scaling factors in α^1 and α^2 necessary to bring the achieved and intended weights into line, turn out to be:

$$1460.281\alpha^2 - 318.230\alpha^1\alpha^2 - 225.238\alpha^2 + 37.022\alpha^1 - 19107\alpha^2 = 0$$

$$27.029\alpha^2 + 61.593\alpha^1\alpha^2 - 37.022\alpha^1 - 4.968\alpha^2 - 16.570 = 0$$

These were solved using a certain program (numerical algorithm group) and the solution was found to be

$$\alpha^1 = 0.400$$

$$\alpha^2 = 0.700$$

Having set $\alpha^3 = 1$

To make all the scaling factors greater than unity we divide each in turn by 0.400 to obtain the required scaling factors:

$$\alpha^1 = 1$$

$$\alpha^2 = 1.936$$

$$\alpha^3 = 2.503$$

Effects of Scaling

If non-unity scaling factors are applied to the component raw marks in an examination, the candidates' rank order would be sure to change. This means that, if we were to be reporting the examination results on a scale of grades (such as pass/fail, distinction/credit/pass or grades fail/pass/good/very good/excellent, etc.) and these grades are awarded simply on the evidence of aggregate marks, some candidates would not get the same grade after scaling as they did before. The extent of the grade difference between marks obtained before and after scaling is highly dependent on the coefficient of correlation between the aggregated unscaled and aggregated scaled marks.

As an extension to the mathematical development given before, it can be shown that the coefficient of correlation (r) is given by

$$r = \frac{\sum_{i=1}^k \alpha_i S_i^2 + \sum_{i=1}^k \sum_{j=i+1}^k (\alpha_i + \alpha_j) S_i S_j}{\sqrt{\sum_{i=1}^k S_i^2}}$$

When we apply this equation to examples of the type given above we characteristically get very high coefficients. Our present example is typical in that r turns out to be 0.98. However, if then look at the grade differences which will occur if we scale the marks we will find that a large number of candidates would change grades. The difficulty here is that there is no obvious way of arriving at a grading scheme to apply **to the aggregated scaled** marks. We use a conservative scheme to minimize the number of candidates who would change grade.

In the examination results cited as our example, a six-point grading scheme was used. If that was applied on the scaled marks, we find that: 0.5 per cent of candidates change by two grades 15.1 percent of candidates change by one grade while 84.4 percent of the candidates have the same grade.

It is typically satisfactory to use a covariance method for determining the weighting of examination components than relying simply on a ratio of standard deviations. It becomes apparent that the covariance approach can be developed into equations for the determination of linear scaling factors which can be applied to raw marks, and extended to enable the correlation coefficient between aggregated scaled and aggregated un-scaled marks to be calculated.

The Secondary School Certificate Exam Results

The weighting procedure that is described above is rather a complicated statistical procedure that is not applied in handling the components of our Secondary School Examination (SSE) results. The usual practice is to scale up the actual marks attained by the candidates in certain exam components by using a certain up scaling factor that varies in accordance to how the candidates fared in that component's exam. The factor used is in the form of $c \sqrt{X_i}$, where c is a constant and X is the actual mark attained by a student in component i . So if we let $c= 10$, and pass mark in that component is 40 after scaling, then that pass mark

is actually 16. That is because $40 = 10\sqrt{X_i}$, and thus $X = 16$. A candidate who attained 25 in the actual exam results will be scaled up to 50, which **means** that his marks are doubled. Let now include in the following table a hypothetical array of marks (X_i), and see what happens when we scale them by using the factor $10\sqrt{X_i} = y$:

<i>Marks achieved</i> X_i (1)	<i>Scaled marks</i> Y_i (2)	Ratio Y_i / X_i (3)
1	10	10
4	20	5
9	30	3.5
16	40	2.5
25	50	2.0
36	60	1.7
49	70	1.5
64	80	1.2
81	90	1.1
100	100	1.0

It can be noticed that column (2), scaled marks, is less dispersed ($SD^{56} = 30.3$) than column (1) of the 'marks achieved' with a standard deviation of 34.2. This may be explained by the ratios in column (3) where it appears that that scaling system favors the low achievers. That system will grant a student who achieved one mark in math a scaled mark of 10, which is equal to 10 times his actual attainment. A student of 4 marks will get 5 times that in the adjusted results. Thus it could happen that a candidate who, for health reasons, sits for just one of the two papers of math and attains 49 in that paper will be getting 70 in the final results. Such a case is likely to happen and reflect negatively on the credibility of the exams in the eyes of that student, his colleagues, his school and other associates in the community who know about the case.

Such scaling practice, with all its negative side effects on the morale and future attitudes of the students, will also nullify the use of the exam results in attesting the efficiency of the educational system being

⁵⁶ standard deviation

applied, measuring the actual students' academic attainment and highlighting the areas of weaknesses for proper and pragmatic treatment. We have the exam data for the years 1993 up to 2001. Some of those results are clearly scaled and normalized. For them to be of use as a basis for genuine research and analyses, the scaling effects should be removed to leave us with results that are as near as possible to the actual state of nature. We presume that the scaling criteria differ from component to component within and between years. Since no information about those criteria is available, we will have no way other than to rely on assumptions. To limit, as much as possible, the risk of making erroneous assumptions due to the multiplicity of scaling procedures used, we will limit ourselves to a few components. By experience and close observation, it is known that the most problematic areas to most of the students are Mathematics and English then we limit this analysis to those two components. We will hypothesize that those two components, being as problematic as they are, which will result in low levels of mark attainment, will need to be treated with a high level scaling criteria. We can presume that a scaling factor of $10\sqrt{X}$, or something of that dimension is being used. The actual standard of our college students in those two components doesn't comply with the high grades they presumably attained in their secondary school exam results. It seems to me reasonable to assume that a high-level scaling factor is used, thus we have to reverse the scaling process before analysis. Let us hypothesize also that this scaling standard was used as of 1996. Then we can divide our time series into two portions: 1993- 1995 as a control group, and 1996-2001 as a comparative. Our statistical analysis will encompass two statistical systems of analysis:

- Meta –Analysis
- Truncated weighting procedure

The Meta- Analysis

Meta- analysis is a collection of systematic techniques for resolving apparent contradictions in research findings. Meta-analysts translate results from different studies to a common metric and statistically explore relations between study characteristics and findings.

Gene Glass⁵⁷ **first** used the term “meta-analysis” in 1976 to refer to a philosophy, not a statistical technique. Glass argued that literature review should be as systematic as primary research and should interpret the results of individual studies in the context of distributions of findings, partially determined by study characteristics and partially random. Since that time, meta-analysis has become a widely accepted research tool, encompassing a family of procedures used in a variety of disciplines.

Meta-analysis responds to several problems in the educational research. First, important issues are studied by numerous investigators. The amount of information on a given topic is probably overwhelming and not amenable to summary. Even when there are relatively few studies on a given topic, it becomes difficult to determine if outcome differences are attributable to chance, to methodological inadequacies, or to systematic differences in study characteristics. Informal methods of narrative review permit biases to remain easily undetected. Reviewers’ biases can influence decisions about study inclusion, relative weights given to different findings, and analysis of relations between study features and outcomes. These biases can have clandestine effects when reviewers do not systematically seek to reduce them or provide sufficient information for readers to evaluate them.

Meta-analysis typically follows the same steps as primary research. The meta-analysis first defines the review’s purposes. Organizing frameworks can be practical or theoretical questions of varying scope, but they must be clear enough to guide study selection and data collection. Second, sample selection consists of applying specified procedures for locating studies that meet specified criteria for inclusion. Typically, meta-analysis are comprehensive reviews of the full population of relevant studies. Third, data are collected from studies in two ways. Study features are coded according to the objectives of the

⁵⁷ Glass, G.V, McGaw, B. & M.L. Smith (1981), *Meta-Analysis in social research*, Beverly Hills, CA: Sage.

review and as checks on threats to validity. Study outcomes are transformed to a common metric so that they can be compared. A typical metric in educational research is the effect size, the standardized difference between treatment and control group means. Finally, statistical procedures are used to investigate relations among study characteristics and findings.

Criticisms of meta-analysis tend to fall into two categories. A complaint that meta-analysis obscures important qualitative information by averaging simple numerical representations across studies. Other critics argue that research is best reviewed by a reflective expert who can sift insight from varying, and may be contradicting, studies.

The Classic or Glassian Meta-Analysis

Glass set the pattern for conventional meta-analysis: define questions to be examined, collect studies and analyze relations between study features and outcomes. Meta-analytic approaches share three features.

- Classic meta-analysis applies liberal inclusion criteria. Glass argued that one should not disregard studies on the bases of study quality a priori; a meta-analysis itself can determine if study quality is related to the variance in reported treatment effects.
- The unit of analysis is the study finding. A single study can report many comparisons between groups and subgroups on different criteria. Effect sizes are calculated for each comparison.
- Meta-analysis using this approach may average effects from different dependent variables, even when these measure different constructs.

Study Effect Meta-Analysis

Study effect meta-analysis alters the Glassian form in two ways. Here the inclusion rules are more selective. Studies with apparent methodological deficiencies are excluded. Also the study is the unit of analysis, but only one effect size is computed for each study. Although this action reduces the number of data points included in the analysis,

but it preserves the independence of the data and gives equal weight to all included studies.

Tests of Homogeneity

There is an argument that conventional statistical tests are not appropriate for meta-analysis. Homogeneity tests were developed to determine the likelihood that variance among effect sizes is due only to sampling error. If the homogeneity test is significant for a group of studies, a procedure similar to the analysis of variance can be used. Studies are repeatedly sub-grouped according to their features until within-group variation turns out insignificant (Dhawan and Gerdes, 1997). According to Pal et al (1998) many elements can cause variations in effect size: measurement unreliability, range restrictions, reporting errors, within-study statistical adjustments, etc.

Psychometric Meta-Analysis

Hunter and Schmidt's ⁵⁸ approach combines some of the best features of the other approaches. All related studies are combined regardless of quality. The distribution of effect sizes is corrected for sampling errors, measurement errors, range errors, etc.. If still we are left with large variance, then we can regroup the effect sizes into subsets according to pre-selected study features, and then Meta-analyze each subset separately. The meta-analysis should estimate true treatment effects under conditions typical of those represented in the studies and predict treatment effects under conditions determined by the reviewer. Accordingly, this technique requires substantial information from individual studies for accurate correction of effect sizes, and this may not be available in all cases.

MODELLING THE EXAM DATA

All modeling involves a compromise in which simplicity is traded off against reality. A model is an abstraction. It is to be judged not by how exactly it corresponds to reality but by how successfully it captures

⁵⁸ Hunter, J.E., & F.L. Schmidt (1990), *Methods of meta-analysis*, Newbury Park, CA: Sage.

the essence of the process being handled. The concept of a model might be a means of establishing a common ground. Here we are trying to establish that common ground through discussion and analysis of the official data of the secondary school examination results for the years 1993—2001. This set of data differs from other educational data in that it is complete and consistent. So it can be dependable as a source data to analyze and model, examine the degree of accountability of the educational process and as a means of uncovering any weaknesses that may be inherent in the educational attainment of the students.

By experience we know that the most problematic areas in the student's educational achievement are related to the areas of Mathematics and English. Observation of our students at the university level shows that they are facing a real difficulty in grasping or even dealing with the basic principles of those two subjects in particular though they have attained marks that are more or less similar to their attainment in other subjects.

This observation gives us an indication to hypothesize that the examination results for those two subjects do not reflect the actual state of student attainment. We hypothesize that the actual student marks were manipulated by some criterion to alleviate the low turnout.

The data

Our sets of data are composed of the secondary level final examination results for Mathematics and English for the years 1993-2001 with the exception of 2000 which is characterized by certain data deficiencies. This data is summarized into 5-mark class intervals for this classification will be sufficiently suitable for multivariate statistical analysis.

Statistical Analysis

The aim of this analysis is to test the accountability of these results by finding out if there any inherent abnormal variations in them between and among the years for those two subjects. We need to find out whether the examination results are an accumulation of the actual marks attained by the candidates, or whether those actual results have been manipulated by a certain criterion of adjustment. We do not mean to intrude into the naturally accepted right of the examiners to look into

the marginal cases, but we are behind the macro-level adjustments, if there are, which are meant to change the whole nature of the results to cover-up the dark areas of student attainment.

Our strategy here is begin our analysis by entertaining the power of the curve and simple descriptive statistics to help us read what the data indicates before moving forward towards more advanced statistical analysis. The outputs hereunder underline a noticeable variation between results of the secondary-school examination in mathematics of the years 1993 & 1994 and the rest of the years 1995-2001. For the years 1993 & 1994 those results are almost identical with some minor variation. That can be due to the increased numbers of candidates in 1994. When numbers are replaced with percentages, then the theme is identical. Output 7 shows a comparison of 1994 and 1995. Though for those two consecutive years there is no reason to justify a sharp increase in student attainment, the striking shift of the results of 1995 as compared to 1994 is rather curious. Looking to output 5, which compares the math-exam results for the years 1993, 1995, 1998 & 2001, we notice that the last three years differ strikingly from 1993. While the curves for 1995, 1998 and 2001 indicate that the greater portion of candidates gets high marks of 58 and over, that of 1993 shows a fairer distribution. There is no reason to imply that the students have suddenly gotten better as of 1995. A more acceptable explanation is that there is a change of policy that was introduced and applied as of the year 1995 and over. The data indicates that a re-evaluation of the results is undertaken through use of a certain adjustment criterion to alleviate weaknesses and upgrade the actual results. A similar analysis of the secondary-school examination results of English would lead us to similar results and implicate the same conclusions, that is; the two years of 1993 & 1994 resemble a block that is completely different from the rest of the following years as a second block. There is an apparent homogeneity within each of these two blocks.

Modeling the exam data

The former analysis has been only concerned with data series on only one variable. We have only discussed data on Mathematics only and then moved to analysis of the data on English. However, when we possess information on two or more related or concomitant variables, it becomes natural to seek a way of expressing the form of the functional

relationship. In addition, it is desirable to know the strength of the relationship. That is, not only do we seek a mathematical function which tells us how the variables are interrelated, but we also wish to know how precisely the value of one variable can be predicted if we know the values of the associated variables. The techniques used to accomplish these two objectives are the regression methods and correlation methods. Regression methods are those used to determine the best functional relationship among the variables, while correlation methods are used to measure the degree to which the different variables are associated.

We are going to begin our analysis of the diversity between the examination results of the two subjects of Mathematics and English and the diversity within each of them through an examination of their consecutive standard deviations. It can be seen from the tables below that the standard deviation for mathematics was as large as 3360.78 for 1993, it went on increasing at a reasonable pace till 1999 (5448.02), i.e. an increase amounting to 162%. However, it almost doubled in 2001 (10208.63) compared to 1999. For English the std is almost double that of mathematics throughout the years under study, and it is increasing at a higher rate to be in 2001 double the std of 1993; where- as it amounted to about 180% in 1999. As known, the standard deviation is a measure of the degree of homogeneity, and here it measures the state of homogeneity of attainment of the student population who are sitting for those national exams in those years. As years pass by students attainment are apparently, getting more and more heterogeneous. One wonders if this increasing heterogeneity is an indicator of worsening student attainment.

For English the standard deviation is almost continually two times that of Mathematics throughout the years, and is increasing at a higher rate. This shouting degree of heterogeneity indicates that student attainment in English is, by far, worse than their already bad showing in Mathematics.

Correlations

Measures of correlation are the techniques developed to measure the degree of association between different variables. Because of the nature of the concept of correlation, it is clearly related to the concept of regression. In fact, for a given regression equation, it seems reasonable to expect that a correlation coefficient will measure how well the regression

equation fits the data or, how closely the data points hug the regression curve. Thus, a correlation coefficient will undoubtedly be related to the standard error of estimate which measures the dispersion of the points about the regression line.

So a correlation coefficient (R) is a function of the standard error of estimate, and it has some characteristics:

- It should be large when the degree of association is high and small when the degree of association is low.
- It should be independent of the units in which the variables are measured.
- $-1 \leq R \leq 1$, where -1 represents perfect negative linear association and $+1$ represents perfect positive linear association of the variables. A value of 0 is interpreted to mean that no linear association exists between variables
- $R^2 =$ the coefficient of determination which represents the percentage of the corrected sum of squares, or total variation, that is explained by, or removed by, the simple or multiple linear regression line. If this percentage is not large enough to the researcher's satisfaction, then he should look for a better fitting regression equation.

In the analysis of the secondary examination data for the two subjects of Mathematics and English with relation to the calculated correlation coefficients, we notice the following:

- The correlation coefficients for the years 1993 and 1994 for both subjects are reasonably high. R for Mathematics for those two years is $.886$, while for English it amounts to $.942$. Meantime, the correlations between those two years and the examination results of the rest of the series are low.
- The correlation coefficients for English for 1993 and 1994 are negative and strikingly lower than those for Mathematics. In the mean time we notice a positively high correlation that is almost perfect, between English results for the years 1995 and over. Those, in particular, are noticeably higher than correlations of those years for Mathematics.
- This will take us back to the findings we encountered when we plotted the data at the beginning. But, here the more

advanced analysis gives more strategic information about the marking, or marking adjustments, criteria:

- There is a sharp difference, in general, between 1993-1994 as a block and the rest of the years as a second block.
- The correlation coefficients for English for the first block are insignificant. The calculated t-test supports the null hypothesis of no association between the two blocks of years.
- Reexamining those findings will prove that there is a remarkable difference in the reevaluation criteria for both subjects. The results for English for the first block are completely uncorrelated with the exam results of the second block. From this we arrive to a conclusion that those results, being strikingly bad, are being adjusted by an adjustment factor that is higher than the factor used to adjust the results of Mathematics, i.e., candidates, though weak in Mathematics, their attainment in English is by far weaker.

The Regression Models

The regression analysis for math results for all of the eight years shows a highly significant model with an R of .880 and coefficient of determination amounting to .774. The model is:

$$Y=80.106-.0074y_{93}+.004y_{94}-.014y_{95}-.0046y_{96}-.002y_{97}+.0048y_{98}+.0017y_{99}+.0047y_{2001}$$

With the exception of the intercept, all other coefficients are very small. The t-tests shows insignificant results for most Bs, thus indicating the absence of a linear relationship between the years, i.e. accepting the null hypothesis that: $B = 0$. Thus our model wouldn't help us predict the future as models are expected to do.

Now let us consider the case of two models:

A full model containing all parameters (years), and⁵⁹ a restricted or reduced model, that places some restrictions on the values of some of these parameters. The effects of these restrictions are measured by the decrease in the effectiveness of the of the restricted model in describing

⁵⁹ Statistical Methods(1997), Rudolf J.Freund &William J. Wilson, Academic Press

our set of data. In regression analyses, the decrease in effectiveness is measured by the increase in the error sum of squares.

The most common inference is to test the null hypothesis that one or more of the coefficients is restricted to a value of zero. This is equivalent to saying that the corresponding independent variables are not used in the restricted model. The measure of the reduction in effectiveness of the restricted model is the increase in the error sum of squares (or equivalently, the decrease in model sum of squares) due to imposing the restriction, that is, due to removing those variables from the model.

The testing procedure is implemented as follows:

- Divide the coefficients in \mathbf{B} into two sets represented by matrices B_1 and B_2 , that is:

$$\mathbf{B} = \begin{vmatrix} B_1 \\ B_2 \end{vmatrix}$$

so as to test the hypotheses

$$H_0: B_2 = 0,$$

$$H_1: \text{at least one element of } B_2 \neq 0$$

Let us denote the number of coefficients in B_1 by q and those in B_2 by p . Thus, $p+q = m+1$, where m is the number of independent variables. Since the ordering of elements in the matrix of coefficients is arbitrary, B_2 may contain any desired subset of the entire subset of coefficients.

- Perform the regression using all coefficients, that is, using the full model $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$. The error sum of squares for the full model is $\mathbf{SSE}(\mathbf{B})$, with $n-m-1$ degrees of freedom.
- Perform the regression using only the coefficients in B_1 , that is, using the reduced model $\mathbf{Y} = \mathbf{X}_1\mathbf{B}_1 + \mathbf{E}$, which the model specified by H_0 . The error sum of squares for the reduced model is $\mathbf{SSE}(B_1)$, with $n-q$ degrees of freedom.

The difference, $\mathbf{SSE}(B_1) - \mathbf{SSE}(\mathbf{B})$ is the increase in the error sum of squares due to the deletion from the model of the coefficients in B_2 . This is defined as the partial contribution of the coefficients in B_2 . Since there are p coefficients in B_2 , this sum of squares has p degrees of freedom, which is the difference between the number of parameters in the full and reduced models. For any model $\mathbf{TSS} = \mathbf{SSR} + \mathbf{SSE}$, hence this difference can also be described as the decrease in the regression (or model) sum of squares due to the deletion of the coefficients in B_2 .