Imhotep Mathematical Proceedings
Volume 1, Numéro 1, (2014).

\[ \beta = \frac{v}{c} = \tanh \chi \]

\[ \int_{-\infty}^{+\infty} \frac{\sin t}{t} \, dt = \frac{\pi}{2} \]

\[ \int_{-\infty}^{+\infty} e^{-x^2} \, dx = \sqrt{\pi} \]

Proceedings of the 2nd Strathmore International Mathematics Conference (SIMC 2013), 12 - 16 August 2013, Strathmore University, Nairobi, Kenya.
Aims and Scope: IMHOTEP is an Electronic African Journal of Pure and Applied Mathematics which publish original and significant research articles in Pure and Applied Mathematics. It also publishes Proceedings of Colloquia written in English or in French, Master and PhD Thesis which have been already defended and Lectures notes. A paper is accepted for publication after a blind refereeing procedure. The referee may be an editor or a solicited expert.

Where does the name IMHOTEP come from?

IMHOTEP was a great well-read scientist and architect in antic Egypt, whose first striking achievements go back to the year 2778 A.D. He was an adviser of Pharaon Djoser, for whom he built the Saqqarah complex. He initiated the building of the first pyramids.

EDITORIAL TEAM

EDITOR IN CHIEF
David Békollé, University of Ngaoundéré, Cameroon.

EDITORIAL ASSISTANT
Edgar Tchoundja, University of Yaoundé I, Cameroon.

EDITORIAL BOARD
Augustin Banyaga, Pennsylvania State University, United States.
Jean-Marc Bardet, Université Paris I, France.
Shiferaw Berhanu, Temple University, United States.
Charles Chidume, African University of Science and Technology, Nigeria.
Marcel Dossa, University of Yaoundé I, Cameroon.
Michael Hortmann, University of Bremen, Germany.
Aderemi KuKu, Grambling State University, United States.
Takaaki Nomura, Kyushu University, Japan.
Brett Wick, Georgia Institute of Technology, United States.

STEERING COMMITTEE
Bitjong Ndombol, University of Yaoundé I, Cameroon.
Khalil Ezzinbi, Université de Cadi Ayyad, Morocco.
Ibrahim Fofana, Université de Cocody, Côte d’Ivoire.
Guy-Martial Nkiet, Université des Sciences et Techniques de Masuku, Gabon.
Eugene Okassa, Université Marien Ngouabi, Congo.
Mary Teuw Niane, Université Gaston Berger de Saint-Louis, Sénégal.
Hamidou Touré, Université de Ouagadougou, Burkina Faso.

Print. ISSN: 1608-9324 - Indexed by MathSciNet and Zentralblatt.
Lang’o Odondi
MRC ConDuCT Hub, School of Social and Community Medicine, University of Bristol, UK
Lango.Odondi@bristol.ac.uk

Abstract

Understanding the causal relationship between intervention and outcome is at the heart of most research in the health sciences, and a variety of statistical methods have been developed to address causality. However, noncompliance with treatment assignment is a key source of complication for causal inference. Estimation of causal effects is likely to be compounded by the presence of noncompliance in both treatment arms of clinical trials where the intention-to-treat (ITT) analysis produces a biased estimator for the true causal estimate even under homogeneous treatment effects assumption. Principal stratification method has been developed to address such posttreatment complications by stratifying the population into partially latent classes (principal strata) based on potential values observed after randomization (e.g. noncompliance) under each of the levels of randomized intervention. The present work combines the two strategies of model selection and principal stratification with a novel application to a real data from a trial conducted to ascertain whether or not unopposed oestrogen (hormone replacement therapy - HRT) reduced the risk of further cardiac events in postmenopausal women who survive a first myocardial infarction. The causal model links the resulting two marginal prediction models with a user-defined sensitivity parameter which is a function of the correlation between the two compliance behaviours. The method’s key assumption of conditional prediction is verified for our data via sensitivity analysis comparing results of causal estimates using different sets of predictors of compliance. We adjust for noncompliance in both treatment arms under a Bayesian framework to produce causal risk ratio estimates for each principal stratum. The results suggested better efficacy for HRT among those who would comply with it compared to those who would comply with either HRT or placebo: compliance with HRT treatment only and with either treatment allocation would reduce the risk for death (reinfarction) by 47%(25%) and 13%(60%) respectively.
Optimal compliance prediction models for estimating causal effects

Lang’o Odondi

Abstract. Understanding the causal relationship between intervention and outcome is at the heart of most research in the health sciences, and a variety of statistical methods have been developed to address causality. However, noncompliance with treatment assignment is a key source of complication for causal inference. Estimation of causal effects is likely to be compounded by the presence of noncompliance in both treatment arms of clinical trials where the intention-to-treat (ITT) analysis produces a biased estimator for the true causal estimate even under homogeneous treatment effects assumption. Principal stratification method has been developed to address such posttreatment complications by stratifying the population into partially latent classes (principal strata) based on potential values observed after randomization (e.g. noncompliance) under each of the levels of randomized intervention. The present work combines the two strategies of model selection and principal stratification with a novel application to a real data from a trial conducted to ascertain whether or not unopposed oestrogen (hormone replacement therapy - HRT) reduced the risk of further cardiac events in postmenopausal women who survive a first myocardial infarction. The causal model links the resulting two marginal prediction models with a user-defined sensitivity parameter which is a function of the correlation between the two compliance behaviours. The method’s key assumption of conditional prediction is verified for our data via sensitivity analysis comparing results of causal estimates using different sets of predictors of compliance. We adjust for noncompliance in both treatment arms under a Bayesian framework to produce causal risk ratio estimates for each principal stratum. The results suggested better efficacy for HRT among those who would comply with it compared to those who would comply with either HRT or placebo: compliance with HRT treatment only and with either treatment allocation would reduce the risk for death (reinfarction) by 47%(25%) and 13%(60%) respectively.

Paper presented at the 2nd Strathmore International Mathematics Conference (SIMC 2013), 12 - 16 August 2013, Strathmore University, Nairobi, Kenya.
I. Introduction

Valid causal inference is a central motivation in the analysis of data from randomised controlled trials when comparing two or more interventions. Effective randomization of subjects between the treatment groups plays a key role in permitting such statistical comparison [22, 19]. However, the presence of intermediate variables in the intervention-outcome causal pathway is likely to complicate estimation of causal effects by introducing selection bias since they often manifest themselves as non-random phenomena [41, 24]. Noncompliance with treatment assignment one such phenomenon which may often manifests itself as treatment discontinuation, switching or subject dropout from the study. Intention-to-treat (ITT) is the standard analysis for estimating causal effects under perfect compliance with treatment assignment. By comparing treatment groups as assigned, the ITT analysis preserves the baseline comparability between treatment groups. However, using ITT results in presence of treatment noncompliance are likely to underestimate the treatment efficacy by mixing the effects of treatment compliers and non-compliers [43]. Since noncompliance mostly manifests itself as a non-random phenomenon, it is a challenge adjusting for its corresponding informative characteristics.

Applying standard regression methods to adjust for intermediate variables produce estimates which lack causal interpretation [29]. One reason for failure in causal analysis is because such methods make the very strong but often implausible assumption of unmeasured confounders between the intermediate variable and outcome. In the presence of noncompliance, per-protocol and as-treated analyses are commonly used to supplement ITT in evaluating efficacy [42, 20]. But these post-hoc analyses lack the benefits of randomization, for example, selection bias in these methods may be evident in different compliance behaviour in different treatment arms. Efron and Feldman [9] in their seminal work used compliance as a covariate in a regression adjustment for a placebo-controlled clinical trial evaluating the efficacy of Cholestyramine in lowering serum cholesterol levels towards reducing the risk of coronary heart disease. However, their method has been criticized for the implicit strong assumption of comparability in compliance between the active treatment and placebo arms [1, 3]. In the presence of selectivity effects, many methods have been developed to account for noncompliance in more than one treatment arm.

Frangakis and Rubin [11] developed the principal stratification as a general framework to adjust for intermediate variables observed post-randomization. The method basically stratifies the population into partially latent classes (principal strata) based on potential values of posttreatment variable, like a noncompliance status. Principal strata comprise units having the same values of the intermediate potential outcomes and are not affected by treatment assignment hence retains the tenets of randomization and so provide valid and well-defined causal effect estimates for selected subgroup/strata.
Principal stratification is a flexible method for causal modelling which may be extended to adjusting for noncompliance in more than one treatment arm. But owing to the latent nature of principal strata, causal inference using the method often requires making structural assumptions to allow parameter identification. According to Cole and Frangakis [6], causal estimates are generally identifiable under three sufficient assumptions: exchangeability (no unmeasured confounders), positivity (existence of a non-zero probability to receive treatment) and consistency (relating observed data to counterfactual data). In the presence of more than one active treatment, a joint analysis may provide additional analytical insights than pairwise efficacy comparisons [4]. In general, comparing more than one active treatment compounds the challenge of identification of causal estimates due to possible multiple forms/degrees of noncompliance with treatment assignment [30, 21].

Crucial to parameter identification under principal stratification method is selection of good baseline covariates which are predictive of the intermediate status (e.g. noncompliance). The implicit challenge of model selection is not only a statistical problem [17] but may be compounded when applied to intermediate variables occurring on the causal pathway and observed post-randomisation. An efficient selection of plausible predictors of intermediate variables can be used to effectively address identification problem of causal estimands by reducing bias in addition to relaxing implicit causal assumptions [20]. From a clinical perspective, adequate knowledge about predictors of treatment compliance may be a valuable tool to inform treatment decisions. Shrinkage principle is a common strategy of reducing regression coefficients to improve the quality of predictions through bias-variance tradeoff so as to produce stable models in the presence of many predictors [34]. For example, Tibshirani [38] introduced the least absolute shrinkage and selection operator (Lasso) as an efficient model selection method which performs the twin tasks of variable selection and regression coefficient estimation simultaneously.

To adjust for noncompliance with treatment assignment, Roy et al. [30] proposed a principal stratification framework for trials to compare two active treatments using baseline covariates to address identification problem. Long et al. [21] also proposed a likelihood estimation method to provide point causal estimands for a three-armed trial by using Bayesian methods to model the arm-specific compliances directly while treating the principal compliance status as missing. And by using Bayesian methods in a potential outcome framework, Zigler and Belin [44] recently used a key covariate predictive of compliance for causal effect estimation in an active-control trial. On the other hand, Fischer et al. [10] proposed a structural mean modelling approach using baseline covariates predictive of compliance in each arm to obtain compliance-adjusted efficacy in a randomized controlled trial comparing two-active treatments.

Central to the application of Roy et al. [30] model is the conditional prediction assumption that potential outcomes are statistically independent of the set of covariates predictive of
compliance given a stratum and treatment assignment. In the presence of many recorded baseline covariates and given this defining assumption, this underscores the crucial role of selecting efficient and meaningful predictors of compliance with treatment assignment for each trial arm. Using a Bayesian approach, the present work modifies and extends the principal stratification method of Roy et al. [30] to integrate optimal model selection procedures using plausible separate predictors of compliance in each arm and apply it to analyze survival data in terms of causal risk ratio estimates for each principal stratum of the Esprit study.

The rest of the paper is organized as follows: Section 2 describes the motivating data from the Esprit study. In Section 3, we describe the relevant causal modeling assumptions. Section 4 provide a brief outline of the methods of analysis by first presenting model selection predicting arm-specific compliance followed by the causal model framework linking the marginal compliance models and the resulting Bayesian inference. Section 5 present an application and results from analysis of the Esprit data. Finally, Section 6 presents a broad discussion.

II. Motivating data: The Esprit study

The onset of menopause is often characterized by diminishing production of oestrogen hormones due to a decline in ovarian function whose unpleasant symptoms (e.g. vasomotor, insomnia, fatigue and depression) can impact negatively on the body leading to low quality of life among such women for the better part of the last third of their lives [27]. Hormone replacement therapy (HRT) is a treatment for oestrogen-deficiency symptoms which is mainly administered in two broad forms depending on whether a woman has her uterus intact or not: unopposed oestrogen (oestrogen taken by itself) for those who have had hysterectomy (removal of the uterus) or oestrogen with progestin for the non-hysterectomized. The addition of progestin is meant to counteract the effects of estrogen on the uterus like endometrial cancer. Although observational studies conducted in the last quarter of last Century showed benefits of HRT in lowering rates of coronary heart disease [12], follow-up clinical trials failed to confirm such beneficial effects among postmenopausal women.

The oEStrogen in the Prevention of ReInfarction Trial (Esprit) study was one of the trials whose ITT results revealed no HRT benefit among postmenopausal women [5]. Esprit was a two-armed placebo-controlled double blind clinical trial conducted to ascertain whether or not unopposed oestrogen reduces the risk of further cardiac events in postmenopausal women aged between 50 – 69 years who survived a first myocardial infarction in England and Wales. The study comprised a total of 1017 subjects: 513 and 504 women were randomized to HRT treatment and placebo arms respectively and monitored over 24 months period. The primary outcomes were cardiac deaths and all-cause mortality or reinfarction. Although ITT analysis
of the data has been previously published, however the analysis took no account of compliance data which we use in this paper to estimate true causal effects.

III. Notation and assumptions

We consider a parallel two-armed clinical trial set up. Let \( Z \in \{0, 1\} \) denote a randomization indicator: \( Z = 1 \) indicate randomization to the new treatment and \( Z = 0 \) indicates randomization to control. In our application, 1 \((Z = 1)\) and 0 \((Z = 0)\) will represent randomization to HRT tablets and placebo treatment respectively. We define \( Y \in \{0, 1\} \) to be the outcome of interest (e.g. death). Also let \( A \in \{0, 1\} \) denote compliance with assigned treatment. For the Esprit study we define compliance as actual taking of assigned treatment up to a day before experiencing event of interest (death/reinfarction) or end of study, whichever occurred first. Although this all-or-nothing compliance definition may appear restrictive, it was considered adequate and plausible since any potential treatment switches are assumed to occur soon after randomization and HRT tablets are presumed to have no carryover effects, i.e. assuming no residual effect of treatment once a subject is classified a non-complier. We note that each subject has two potential compliance levels \( A_0 \) and \( A_1 \) (compliance with placebo and HRT treatment respectively) and two potential outcomes \( Y_0 \) and \( Y_1 \) (outcome under placebo and HRT treatment respectively). But the observed compliance and outcomes are respectively given by

\[
A = ZA_1 + (1 - Z)A_0 \quad \text{and} \quad Y = ZY_1 + (1 - Z)Y_0.
\]

Analysis under the principal stratification formulation utilizes baseline covariates \( X \) to modify the standard assumptions (a)-(e) for causal modelling [18, 2] together with a new assumption (f):

(a) Randomization: \( Z \perp \{Y_0, Y_1, A_0, A_1, X\} \), i.e. ignorable treatment assignment assumption.

(b) Stable unit-treatment value assumption (SUTVA), i.e. no interference between treatment units.

(c) Treatment access restriction: which posits no treatment switches among subjects.

(d) Exclusion restriction: \( \Pr(Y_1|A_Z, X) = \Pr(Y_0|A_Z, X) \), i.e. treatment assignment has no effect on outcome except only through treatment received (knowledge of treatment assignment alone has no effect on outcome).

(e) Monotonicity: \( \Pr(A_1 = 1|A_0 = 1, X) \geq \Pr(A_1 = 1|A_0 = 0, X) \), i.e. the probability of compliance with treatment assigned by \( Z = 1 \) is higher among those who would comply with treatment assigned by \( Z = 0 \), compared to those who would not.

(f) Conditional prediction: \( Y \perp X|S, Z \), i.e potential outcome is statistically independent (ignorable) of the set of covariates predictive of compliance for a given principal stratum and treatment assignment.
Any switching of treatment is assumed to have occurred soon after randomization such that a subject is assumed to have completely taken HRT or placebo treatment. Assuming no other form of treatment interruptions among patients and no carryover effects of treatment, the all-or-nothing compliance definition above may be considered plausible with respect to the exclusion restriction assumption. The monotonicity assumption as applied here helps tighten the bounds of causal effects, i.e. ensures compliance type is observable when $Z \neq A$ [4, 30, 37].

In addition to the basic monotonicity assumption (no treatment defiers), the ‘extended’ monotonicity assumption posits similar compliance behaviour for both treatment arms. Plausibility of the ‘extended’ version of monotonicity assumption for the Esprit study may be discerned from the fact that there was no preference for one treatment over the other, i.e. compliance with HRT treatment would be more prevalent among those who would comply with placebo. In our application, this assumption is reflected through a user-defined positive correlation (sensitivity parameter $\phi$) between $A_0$ and $A_1$.

The conditional prediction assumption (f) is crucial for parameter identification in the Roy et al. [30] model. The assumption underscores an integral component of the method which involves selecting suitable predictors of compliance. The first part of this paper will address this challenge through a comprehensive model selection of the Esprit study to obtain optimal arm-specific predictors compliance. We will use penalized maximum likelihood (shrinkage) procedures to select plausible separate predictors of compliance for HRT treatment and placebo arms. Although this is an untestable assumption, we will compare results from different sets of predictors as a form of sensitivity analysis.

Each subject is assumed to belong to one of four basic principal strata defined by unique combinations of $(A_0, A_1)$ where the principal strata comprise the set $S = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$. The causal inference of interest (Section IV.2) will be to seek the joint distributions $[(Y_0, Y_1)|S = s] \forall s \in S$ which provides principal effects of interest for each stratum.

**IV. Methods for analysis**

**IV.1. Compliance prediction models and validation**

We use the logistic models to predict compliance to treatment allocation for each arm separately given a selected set of predictors of compliance $x_0 = 1$ and $x_1, \ldots, x_n$:

$$\text{logit} \left[ \mu_j(x) \right] = \sum_{i=0}^{n} \gamma_{ji}x_i, \; j = 0, 1,$$

where $\mu_j(x)$ is the probability of compliance with allocated treatment $j$ given a set of covariates $X$: the estimated probabilities of complying with arm-specific treatment allocation may then
be obtained using
\[
\hat{\mu}_j(x) = \left[ 1 + \exp \left( - \sum_{i=0}^{n} \hat{\gamma}_{ji} x_i \right) \right]^{-1}, \quad j = 0, 1,
\] (2)
where \( \gamma \) represent the log odds ratio estimates of compliance.

How the two compliance behaviours are correlated is a crucial issue. Following Roy et al. [30], we define a non-negative sensitivity parameter \( \phi \) as a function of the correlation \( \rho \) between compliances to treatment allocation (0/1): if \( \hat{\mu}_0(x) > \hat{\mu}_1(x) \) then
\[
\phi = \rho \sqrt{\frac{\hat{\mu}_0(x)[1 - \hat{\mu}_1(x)]}{\hat{\mu}_1(x)[1 - \hat{\mu}_0(x)]}}.
\] (3)

The compliance models for each treatment arm provided by Equations (1) and (2) may be obtained through comprehensive model selection for predictive covariates, i.e. selecting both clinically sensible and statistically concise predictors of compliance with treatment assignment for each trial arm. With many covariates, the classical stepwise model selection procedures are likely to produce suboptimal prediction models [39, 13]. In comparison, penalized maximum likelihood methods have been shown to perform relatively better in selecting optimal predictors [15]. But a selected statistical prediction model also needs validation to evaluate its predictive ability, for example, external validation enables assessment of the performance of prediction in new data [34, 32]. But in the absence of further data, bootstrap validation provides reliable results by allowing calculation of predicted probabilities from a model which can be compared with the actually observed outcomes [8].

Percentage of optimism and both calibration and discrimination indices are among the most effective and commonly used measures of validation performance [14]. While calibration is a reliability measure of how well the model predictions compare with the observed outcomes, discrimination refers to the ability of the model to distinguish between subjects with positive or negative outcomes (e.g. the ability of a model to distinguish compliers with treatment allocation from non-compliers). Calibration is often quantified in practice by the calibration slope [7] obtainable from the validation plot which is a plot of observed probabilities against the predicted probabilities. On the other hand, discrimination is commonly measured using the concordance \( c \)-statistic as widely expressed in terms of the Somers rank correlation [33]:
\[
D_{xy} = 2(c - 0.5).
\]
This is a measure of the difference between concordance and discordance probabilities [13], such that \( c = 0.5 \) (1) implies random predictions (perfect discriminations). We can readily discern that larger values of calibration and discrimination (concordance) indicate better prediction and such models should indicate lower percentage of optimism.

**IV.2. Causal model joining the marginal compliance models**

The main causal inference of interest is obtainable from the joint distributions \( [(Y_0, Y_1)|S = s] \). Following Roy et al. [30], the joint probability distribution of compliance to the standard
treatment 0 and compliance to new treatment 1 is a function of the arm-specific marginal compliance probabilities and \( \phi \) and can be estimated for a given value of \( \phi \). We however note that \( \phi \) is unknown, in general. Specifically if \( U(x) = \min\{1, \frac{\hat{\mu}_1(x)}{\hat{\mu}_0(x)}\} \) then Roy et al. [30] showed that the joint probabilities are given by

\[
\hat{\mu}_{11}(x) = Pr(A_0 = 1|X)P(A_1 = 1|A_0 = 1, X) = \hat{\mu}_0(x)\hat{\mu}_1(x) + \phi\hat{\mu}_0(x)[U(x) - \hat{\mu}_1(x)],
\]
\[
\hat{\mu}_{01}(x) = Pr(A_0 = 0|X)P(A_1 = 1|A_0 = 0, X) = \hat{\mu}_1(x) - \hat{\mu}_0(x)\hat{\mu}_1(x) - \phi\hat{\mu}_0(x)[U(x) - \hat{\mu}_1(x)],
\]
\[
\hat{\mu}_{10}(x) = Pr(A_0 = 1|X)P(A_1 = 0|A_0 = 1, X) = \hat{\mu}_0(x) - \hat{\mu}_0(x)\hat{\mu}_1(x) - \phi\hat{\mu}_0(x)[U(x) - \hat{\mu}_1(x)],
\]
\[
\hat{\mu}_{00}(x) = Pr(A_0 = 0|X)P(A_1 = 0|A_0 = 0, X) = 1 - \hat{\mu}_0(x) - \hat{\mu}_1(x) + \hat{\mu}_0(x)\hat{\mu}_1(x) + \phi\hat{\mu}_0(x)[U(x) - \hat{\mu}_1(x)],
\]

where \( X \) is the set of covariates predictive of compliance in both arms and \( A_1(A_0) \) is an indicator of compliance to HRT treatment (placebo) and \( \hat{\mu}_{ij}(x) \) denote the probability of being in the compliance subgroup \( ij \) (i.e. estimated proportion of compliance per stratum).

Following the principal stratification framework developed by Frangakis and Rubin [11], the possible values of \( A_0 \) and \( A_1 \) define a stratification factor \( S \) for the population of patients. For a defined outcome variable \( Y \) (mortality/reinfarction for the Esprit study), let \( Y_0 \) and \( Y_1 \) refer to potential outcomes under placebo and HRT treatments respectively. There are four possible realizations of \( (Y_0, Y_1) \) at each level of \( S \) (for example \( s_{11} = Pr[y_0 = 1, y_1 = 1] \)). The joint distribution of potential outcomes \( (Y_0, Y_1) \) for each stratum \( S, f(Y_0, Y_1|S, \varpi) \) may be assumed to be multinomial with probabilities \( \varpi(S) = Pr(Y_0 = y_0, Y_1 = y_1|S = s) \). And by the exclusion restriction, the expression for stratum \( S = (0, 0) \) differs from the others: \( \varpi_{10}(0, 0) = \varpi_{01}(0, 0) = 0 \).

After reparameterizing in terms of \( \pi \) (probability of experiencing event, e.g. death or myocardial reinfarction) and \( \beta = f(\gamma, \phi) \) (log odds ratio of compliance for specified sensitivity), Roy et al. showed that the observed-data likelihood is

\[
L(\pi, \beta|Y, A, Z, X) = \sum_{s=0}^{3} [\pi_Z^{S=1}]^Y [1 - \pi_Z^{S=1}]^{1-Y} Pr(S = s|X, \beta)G(s, A, Z),
\]

where \( \pi_Z^{S=1} \) is the probability that observed \( Y = 1 \), given \( S = s \) and allocation to arm \( Z \), and

\[
G(s, A, Z) = I(s=0){1-A} + I(s=1){A(1-Z) + (1-A)Z}
+ I(s=2){AZ + (1-A)(1-Z)} + I(s=3)A.
\]

Now \( Y \) and \( A \) are observed values; e.g. \( A = A_1(A_0) \) if allocated to active treatment (placebo). We can decompose the expression (5) for each stratum:
We then obtain the stratum-specific relative risks for experiencing an event (death/reinfarction)

\[
\tau_{11} = \frac{\pi_{1}^{s(1,1)}}{\pi_{0}^{s(1,1)}} = \frac{\hat{\pi}_{2}}{\hat{\pi}_{0}}, \quad \tau_{01} = \frac{\pi_{1}^{s(0,1)}}{\pi_{0}^{s(0,1)}} = \frac{\hat{\pi}_{1}}{\hat{\pi}_{7}}, \quad \tau_{10} = \frac{\pi_{1}^{s(1,0)}}{\pi_{0}^{s(1,0)}} = \frac{\hat{\pi}_{4}}{\hat{\pi}_{5}}.
\]

By the exclusion restriction \(\pi_{1}^{s(0,0)} = \pi_{0}^{s(0,0)}\), i.e. the risk of experiencing event of interest (e.g. death) is independent of the arm of allocation among the people who would comply with neither allocation. Writing

\[
\pi_{1} = \pi_{1}^{s(0,1)}, \pi_{2} = \pi_{1}^{s(1,1)}, \pi_{3} = \pi_{1}^{s(0,0)}, \pi_{4} = \pi_{1}^{s(1,0)},
\]

\[
\pi_{5} = \pi_{0}^{s(1,0)}, \pi_{6} = \pi_{0}^{s(1,1)}, \pi_{7} = \pi_{0}^{s(0,1)},
\]

we obtain 7 parameters captured by \(\pi\) from the likelihoods above using logistic models:

\[
\Pr[Y = 1|A = 1, Z = 1] = \pi_{1}\mu_{01} + \pi_{2}\mu_{11},
\]

\[
\Pr[Y = 1|A = 0, Z = 1] = \pi_{3}\mu_{00} + \pi_{4}\mu_{10},
\]

\[
\Pr[Y = 1|A = 1, Z = 0] = \pi_{5}\mu_{10} + \pi_{6}\mu_{11},
\]

\[
\Pr[Y = 1|A = 0, Z = 0] = \pi_{3}\mu_{00} + \pi_{7}\mu_{01}.
\]

We then obtain the stratum-specific relative risks for experiencing an event (death/reinfarction)
The $\tau_{ij}$ above provides the desired principal (causal) effects in terms of causal risk ratios obtained as medians of posterior relative risks of event (death or reinfarction) for each stratum defined by compliance type:

(i) $\tau_{11}$: causal risk ratio of event due to compliance with HRT treatment relative to placebo among the subgroup of patients who would comply with either treatment allocation, i.e. $S=(1,1)$,

(ii) $\tau_{01}$: causal risk ratio of event due to compliance with HRT treatment only among women who would comply if allocated to it, i.e. $S=(0,1)$ and

(iii) $\tau_{10}$: causal risk ratio of event due to compliance with placebo treatment only among the subgroup who would comply if allocated to it, i.e. $S=(1,0)$.

The parameters above (Equation 9) can be estimated using Bayesian methods with suitable priors. Using uninformative (flat) priors $\pi \sim U(0,1)$, for example, may be satisfactory for our analyses given the likelihood of a typical trial data to dominate such priors and the fact that randomized trials are principally designed to be conclusive [16].

To extend the methods which adjust for noncompliance in one treatment arm to adjusting for noncompliance in two treatment arms, we will use a Bayesian approach to apply principal stratification using the Roy et al. [30] model reviewed above for survival data but which was originally proposed for binary data. Specifically we perform a comprehensive model selection to obtain arm-specific optimal predictors of compliance and develop a causal model linking the two marginal models from which we obtain causal effects for each stratum.

V. Application to the Esprit study

ITT analysis of the Esprit data showed no statistical difference between HRT and placebo treatment with hazard ratio results (HR=exp(\hat{\psi})=0.795, \ p-value=0.335, 95%CI : 0.498, 1.268) suggesting a beneficial effect of HRT over placebo with respect to death. However, the ITT analysis took no account of compliance data. The rate of observed compliance was higher in the placebo (63%) arm compared to the HRT (42%) treatment arm which may be attributed to noncompliance as a result of possible unpleasant symptoms like bleeding. Utilizing compliance data, we consider two outcomes (all-cause mortality and myocardial reinfarction) for causal analysis using the methods described in the previous section.

When applying the Roy et al. [30] method for survival data, we use relative risks to approximate hazard ratios. This is justifiable for our analysis given that under short follow-up times (monthly) and small event rates conditions (death and myocardial reinfarctions), relative risk has been shown to be an algebraic approximation of hazard ratio, i.e. $\exp(\hat{\psi}) \approx \hat{\tau}$ [36].
V.1. Predicting compliance in each arm

In predicting compliance for each arm, we first choose a full (saturated) model consisting of all potential predictors of compliance to treatment allocation on the basis of both clinical and statistical plausibility. In addition we consider penalized maximum likelihood estimation regression versions of this (saturated) model and also Lasso (Least Absolute Shrinkage and Selection Operator) model obtained by the method which performs the twin tasks of parameter shrinkage and model selection [38]. To evaluate the predictive performance of the selected models, we used calibration slope, percentage of optimism and calculated discrimination’s concordance $c$-statistics from the reported Somers $D_{xy}$ statistics. These validation measures will be recorded for each individual arm in five models:

(i) Original saturated model with all the 9 predictors without any selection:

$$\text{logit}(\mu_j) = \gamma_0 + \gamma_1 \text{Hysterectomy} + \gamma_2 \text{Smoking status} + \gamma_3 \text{Social-class} + \gamma_4 \text{Age} + \gamma_5 \text{CVD Risks} + \gamma_6 \text{Diabetes} + \gamma_7 \text{Fracture} + \gamma_8 \text{Alcohol} + \gamma_9 \text{HRT},$$

where $\mu$ is the probability of compliance with treatment (placebo/HRT) allocation and histories of hysterectomy, cerebrovascular disease (CVD) risks, diabetes, fracture, alcohol, HRT use together with smoking status are taken as binary 0/1 predictors.

(ii) Reduced model obtained from (i) above by stepwise backward elimination procedures using Akaike information criterion (AIC) stopping rule and 0.10 significance level for a variable to be retained in a model.

(iii) Model fitted with the retained predictors in reduced model (ii) above but the predictors assumed pre-specified (following suggestion by Harrell et al. [14]).

(iv) Intermediate model composed of 6 variables constructed using penalized maximum likelihood estimation with modified AIC ($\chi^2 > 2\text{df}$).

(v) Least Absolute Shrinkage and Selection Operator (Lasso) model selection from the original (i) above.

V.2. Validation: evaluating performance of selected models

We used enhanced bootstrap on all aspects of models development (selection and estimation procedures) to revalidate on samples taken with replacement from the whole sample and apply on the five models specified above. The reduced model was obtained from the original model using stepwise backward elimination procedures using AIC stopping rule and 0.10 significance level for retaining a predictor in a model. The variables selected for the reduced model were consistently (90%) selected across bootstrap resamples. These were the same predictors deemed important by the backward elimination algorithm.

Table 1 provide results showing the performance of the 5 prediction models in terms of validation indices outlined earlier (see Section IV.1). The saturated (original) model consisting
Table 1. Validation performance of 5 models in terms of calibration, concordance and optimism (Section IV.1)

<table>
<thead>
<tr>
<th>Model</th>
<th>Selected predictors</th>
<th>Calibration slope</th>
<th>Optimism (%)</th>
<th>Concordance c-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Original</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRT</td>
<td></td>
<td>0.818</td>
<td>6.1</td>
<td>0.639</td>
</tr>
<tr>
<td>Placebo</td>
<td></td>
<td>0.671</td>
<td>8.6</td>
<td>0.573</td>
</tr>
<tr>
<td>(ii) Reduced</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRT</td>
<td>(hyst+smk+CVD)</td>
<td>0.827</td>
<td>5.8</td>
<td>0.620</td>
</tr>
<tr>
<td>Placebo</td>
<td>(hyst+smk+alc)</td>
<td>0.667</td>
<td>8.1</td>
<td>0.566</td>
</tr>
<tr>
<td>(iii) Reduced†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRT</td>
<td>(hyst+smk+CVD)</td>
<td>0.961</td>
<td>1.4</td>
<td>0.642</td>
</tr>
<tr>
<td>Placebo</td>
<td>(hyst+smk+alc)</td>
<td>0.950</td>
<td>2.0</td>
<td>0.597</td>
</tr>
<tr>
<td>(iv) Intermediate (6 predictors)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRT</td>
<td></td>
<td>0.879</td>
<td>4.1</td>
<td>0.636</td>
</tr>
<tr>
<td>Placebo</td>
<td></td>
<td>0.766</td>
<td>6.0</td>
<td>0.580</td>
</tr>
<tr>
<td>(v) Lasso</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRT</td>
<td>(hyst+smk+age+CVD)</td>
<td>0.935</td>
<td>2.3</td>
<td>0.647</td>
</tr>
<tr>
<td>Placebo</td>
<td>(hyst+smk+alc)</td>
<td>0.925</td>
<td>2.3</td>
<td>0.595</td>
</tr>
</tbody>
</table>

hyst≡hysterectomy; smk≡smoking status; alc≡alcohol; CVD≡cerebrovascular disease; †model assumed pre-specified

of 9 predictors produced better predictions of compliance to HRT than placebo. Here predicting compliance to HRT and placebo, the original models would be overfitted by 18% and 33% respectively. Also these models would be optimistic by 6% and 9% respectively in predicting compliance to HRT and placebo. We note that although the observed rates of compliance for placebo were higher than for HRT, the relatively ‘poor’ performance of the former compared to the later predictive models may be attributed to poor quality of compliance data owing to the common practice to monitor compliance with active treatment more accurately than placebo treatment.

Compared to saturated model, the reduced model would perform relatively better in predicting compliance to HRT compared to predicting placebo: specifically the reduced model
predicting compliance to HRT would perform better at distinguishing compliers from non-compliers (concordance \( c = 0.620 \)) than the reduced model predicting compliance to placebo \( (c = 0.566) \). Reduced models predicting compliance to placebo would be more optimistic (8%) than those predicting compliance to HRT (6%). Predictions for compliance to HRT using the reduced model would be equally well calibrated \((\text{slope}=0.83)\) compared to predictions from the original full model \((\text{slope}=0.82)\). As expected, the model with 3 predictors if assumed pre-specified performed ‘best’ in terms of both calibration and discrimination among the 5 models considered in predicting compliance to both HRT and placebo. These models also produced least optimistic fits for predicting compliance to both arms. Specifically predictions of compliance to both HRT and placebo using the 3 predictors assumed pre-specified were almost perfectly calibrated (0.96 and 0.95) and least optimistic (1% and 2%).

Validation of the model with 6 predictors showed adequate performance with intermediate measures between the saturated models composed of 9 predictors and the Lasso models. We observe that predictions of compliance to both HRT and placebo using the intermediate model performed relatively ‘better’ than both predictions of compliance using the reduced model. For example, predictions of compliance to placebo using the intermediate model is now equally optimistic (6%) as predictions of compliance to HRT using the reduced model, a result which may make the assumption of ‘no preference to one treatment over the other’ (extended monotonicity) plausible for the Esprit study.

Besides the reduced model fitted with predictors if assumed pre-specified, Lasso models produced the best calibrated and discriminative models predictive of compliance to both HRT and placebo (Table 1, lower panel). Predictions of compliance to both HRT and placebo using the Lasso models were the least optimistic (2%) and almost perfectly calibrated \((\text{slope}=0.93)\). Although the Lasso prediction models performed ‘best’ compared to predictions from the other four models, we note that the method uses the same tuning parameter for all coefficients. The drawback of shrinking all coefficients by a constant, even for those non-zero coefficients, may result in suboptimal choice of covariates with the potential to exclude potential predictors, i.e. data wastage. Moreover, Lasso is known to fail efficient model selection in the presence of correlated variable where it tends to select one variable from a group and ignore the others [45]. Overall, the intermediate models provided substantially improved predictions of compliance to both HRT and placebo in terms of calibration and optimism without affecting the capability to discriminate between compliers and non-compliers.

Table 2 provides estimated median compliance proportion for each of the four strata for both all-cause mortality and myocardial reinfarction outcomes. On average, the estimated median probabilities of compliance was higher among those patients allocated to placebo for both all-cause mortality (myocardial reinfarction) outcomes \((\hat{\mu}_0(x) = 0.567 \ (0.565))\) compared to
Table 2. Median compliance proportion per principal stratum for different values of $\phi$

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Outcome: All-cause mortality</th>
<th>Outcome: Reinfarction</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{11}(x)$</td>
<td>$\phi$</td>
<td>0.264 0.296 0.353 0.406 0.460</td>
</tr>
<tr>
<td>$\mu_{01}(x)$</td>
<td>0.197 0.165 0.105 0.047 0</td>
<td>0.200 0.164 0.111 0.059 0.023</td>
</tr>
<tr>
<td>$\mu_{10}(x)$</td>
<td>0.303 0.262 0.211 0.155 0.105</td>
<td>0.300 0.263 0.206 0.142 0.102</td>
</tr>
<tr>
<td>$\mu_{00}(x)$</td>
<td>0.236 0.277 0.330 0.391 0.435</td>
<td>0.235 0.270 0.325 0.382 0.419</td>
</tr>
</tbody>
</table>

those on HRT tablets ($\bar{\mu}_1(x) = 0.461 (0.470)$), i.e. the ratio $U(x) = \min\{1, \frac{\bar{\mu}_1(x)}{\bar{\mu}_0(x)}\} = 0.795 (0.810)$. We note a likelihood that a higher prevalence (proportion) of placebo compliance compared to HRT may be a limitation of the model to effectively evaluate active HRT efficacy.

For all the four strata at mild value of sensitivity parameter ($\phi = 0.2$), the group with the highest prevalence was patients who would comply with either treatment ($\bar{\mu}_{11} = 0.296 (0.303)$) and the group with the lowest prevalence is those who would only comply with HRT tablets ($\bar{\mu}_{01} = 0.165 (0.164)$). The median proportion of compliance among those patients who would comply with placebo only and those who would not comply with either treatment allocation were $\bar{\mu}_{10} = 0.262 (0.263)$ and $\bar{\mu}_{00} = 0.277 (0.270)$ respectively. Overall the estimated rates of potential compliance were generally similar in each stratum for a given value of $\phi$ (except for perfect correlation $\phi = 1$). The apparent independence between the potential compliance rates and outcome may be an indication of the plausibility of conditional compliance assumption with respect to the Esprit data.

### V.3. Causal risk ratio inference

We estimated the causal risk ratio parameters $\tau_{ij}$ (Eq. 9) in a Bayesian setting using non-informative priors for all log odds ratio parameters $\gamma_j$ for potential predictors of compliance. We specified uniform $(0, 1)$ priors for the $\pi_2^{z_{S}} (\pi_i, i = 1, \ldots, 7)$ parameters, $z = 0, 1, S=(0,0),(1,0),(0,1),(1,1)$ and set the sensitivity parameter $\phi = 0, 0.2, 0.5$ and 0.8. The choice of $\phi$ were motivated by the need to explore all possible compliance behaviours including conditional independence ($\phi = 0$) and almost-perfect correlation ($\phi = 0.8$). We ran three chains: null starting values for chain one, mean and median values from a trial run for chains two and three respectively. For convergence assessment, we ran simulation for 101,000 iterations for each of the three chains and excluded the first 1,000 as burn-in. Posterior median relative risks provided Bayesian point estimates for each stratum.
Table 3. Causal risk ratio estimates (means of median posterior relative risks) for mortality and reinfarction (mean median 95% CI) for each stratum for different values of φ: (a) All-cause mortality (b) Myocardial reinfarction.

<table>
<thead>
<tr>
<th>Compliance with both HRT and placebo</th>
<th>Compliance with HRT only</th>
<th>Compliance with placebo only</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\pi}_2(\pi_1^{e(1,1)}))</td>
<td>(\hat{\pi}_0(\pi_0^{e(1,1)}))</td>
<td></td>
</tr>
<tr>
<td>0.011</td>
<td>0.12</td>
<td>0.053</td>
</tr>
<tr>
<td>(\tau_{11} = 0.941)</td>
<td>(0.026, 34.349)</td>
<td>(0.101, 1.347)</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\pi}_1(\pi_1^{e(0,1)}))</td>
<td>(\hat{\pi}_7(\pi_0^{e(0,1)}))</td>
<td></td>
</tr>
<tr>
<td>0.011</td>
<td></td>
<td>0.099</td>
</tr>
<tr>
<td>(\tau_{11} = 0.867)</td>
<td>(0.023, 31.489)</td>
<td>(0.083, 1.522)</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\pi}_1(\pi_1^{e(0,1)}))</td>
<td>(\hat{\pi}_7(\pi_0^{e(0,1)}))</td>
<td></td>
</tr>
<tr>
<td>0.011</td>
<td></td>
<td>0.118</td>
</tr>
<tr>
<td>(\tau_{11} = 0.878)</td>
<td>(0.025, 30.539)</td>
<td>(0.068, 2.310)</td>
</tr>
<tr>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\pi}_1(\pi_1^{e(0,1)}))</td>
<td>(\hat{\pi}_7(\pi_0^{e(0,1)}))</td>
<td></td>
</tr>
<tr>
<td>0.012</td>
<td></td>
<td>0.243</td>
</tr>
<tr>
<td>(\tau_{11} = 0.974)</td>
<td>(0.032, 30.698)</td>
<td>(0.065, 10.399)</td>
</tr>
</tbody>
</table>

(b)

<table>
<thead>
<tr>
<th>Compliance with both HRT and placebo</th>
<th>Compliance with HRT only</th>
<th>Compliance with placebo only</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\pi}_2(\pi_1^{e(1,1)}))</td>
<td>(\hat{\pi}_0(\pi_0^{e(1,1)}))</td>
<td></td>
</tr>
<tr>
<td>0.012</td>
<td>0.021</td>
<td>0.106</td>
</tr>
<tr>
<td>(\tau_{11} = 0.561)</td>
<td>(0.061, 19.989)</td>
<td>(0.293, 1.406)</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\pi}_1(\pi_1^{e(0,1)}))</td>
<td>(\hat{\pi}_7(\pi_0^{e(0,1)}))</td>
<td></td>
</tr>
<tr>
<td>0.010</td>
<td></td>
<td>0.143</td>
</tr>
<tr>
<td>(\tau_{11} = 0.397)</td>
<td>(0.011, 13.839)</td>
<td>(0.301, 1.561)</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\pi}_1(\pi_1^{e(0,1)}))</td>
<td>(\hat{\pi}_7(\pi_0^{e(0,1)}))</td>
<td></td>
</tr>
<tr>
<td>0.008</td>
<td></td>
<td>0.244</td>
</tr>
<tr>
<td>(\tau_{11} = 0.204)</td>
<td>(0.006, 5.479)</td>
<td>(0.432, 2.881)</td>
</tr>
<tr>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\pi}_1(\pi_1^{e(0,1)}))</td>
<td>(\hat{\pi}_7(\pi_0^{e(0,1)}))</td>
<td></td>
</tr>
<tr>
<td>0.008</td>
<td></td>
<td>0.256</td>
</tr>
<tr>
<td>(\tau_{11} = 0.121)</td>
<td>(0.004, 1.186)</td>
<td>(0.465, 45.238)</td>
</tr>
</tbody>
</table>

\(^a\) All-cause mortality; \(^b\) Myocardial reinfarction

Table 3 provide causal risk ratio estimates (Bayesian principal effects) obtained from mean posterior median relative risks for each stratum and corresponding mean 95% credible intervals for different values of sensitivity parameter \(\phi\). Here a posterior relative risk \(\tau\) was obtained as the ratio of two probabilities of experiencing an event due to compliance with one treatment allocation relative to another in a stratum. Most of results (not all) show posterior median relative risks of less than one for all values of \(\phi\) which indicates lower risks for mortality.
and myocardial reinfarction for those women randomized to HRT who would be highly compli-
ant with their treatment allocation. A primary interest is the quantity $\tau_{01} = \frac{\pi_{S=0.1}}{\pi_{S=0.1} - \pi_{S=0.1}}$, i.e. the posterior (causal) relative risk for mortality/reinfarction among the subgroup who would comply with HRT treatment only. The results shows that the mean median 95% credible intervals widened with increase in $\phi$ values, an indication of less correlation between HRT treatment and placebo compliances. Overall, the results indicated that HRT tablets reduced risks for myocardial reinfarction more than the reduction in risks for all-cause mortality.

For a mild correlation value ($\phi = 0.2$), the results suggest that compliance with HRT tablets only would substantially reduce the risk of death by about 47%, i.e. causal risk ratio $\tau_{01} = 0.533$, 95% CI : 0.083,1.522. Also for this value of sensitivity parameter, the results suggest that compliance with HRT treatment compared to taking placebo among those who would comply with either treatment reduced the risk for all-cause mortality by 13%, i.e. causal risk ratio $\tau_{11} = 0.867$, 95% CI : 0.023,31.489. However, compliance with placebo treatment only would marginally reduce the risk of death by about 6%, i.e. causal risk ratio $\tau_{01} = 0.938$, 95% CI : 0.267, 4.038. In general we note that compliance with HRT treatment only consistently suggested reduction of risk for death at all sensitivity parameter values $\phi$. For example, when $\phi = 0.8$, while compliance with HRT treatment compared to taking placebo among those who would comply with either treatment would essentially have no effect ($\tau_{11} = 0.974$, 95% CI : 0.032, 30.698), compliance with placebo treatment only would reduce the risk of death by about 44%, i.e. causal risk ratio $\tau_{10} = 0.560$, 95% CI : 0.024, 5.423.

The size of causal (principal) effects varied according to the value of sensitivity parameter. Risks for myocardial reinfarction among those who would comply with placebo treatment only increased with increase in the value of sensitivity parameter $\phi$. On the other hand results show reduction in risks for myocardial reinfarction among those who would comply with either placebo or HRT treatment as the value of $\phi$ increased. As expected the risks for both death and myocardial reinfarction outcomes were higher among those who would comply with placebo only compared to those who would comply with HRT only for any chosen value of the sensitivity parameter $\phi$.

V.4. Sensitivity analysis

As outlined in assumptions (Section III), application of the Roy et al. [30] method is premised on plausibility of the crucial, but untestable, conditional compliance assumption which posits that the potential outcome is independent of the set of covariates predictive of treatment compliance given a compliance type and treatment assignment. Hence the task of selecting suitable predictors of treatment compliance constitutes an integral part of ensuring plausibility of this assumption. A sensitivity analysis using different models (i.e sets of predictors) may be used to assess how the causal estimates depend on departures from this crucial assumption [35].
Table 4. Sensitivity analysis using 3, 6 and 9 predictors of compliance: Causal risk ratio (median 95% CI)

<table>
<thead>
<tr>
<th></th>
<th>All-cause mortality</th>
<th></th>
<th>Myocardial reinfarction</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Comply HRT and placebo</td>
<td>Comply HRT only</td>
<td>Comply placebo only</td>
<td>Comply HRT and placebo</td>
</tr>
<tr>
<td>(\phi)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.126 (0.030,44.588)</td>
<td>0.546 (0.106,1.417)</td>
<td>0.872 (0.281,2.260)</td>
<td>0.539 (0.014,19.650)</td>
</tr>
<tr>
<td>0.2</td>
<td>1.169 (0.031,42.058)</td>
<td>0.543 (0.088,1.633)</td>
<td>0.833 (0.210,2.428)</td>
<td>0.419 (0.012,14.459)</td>
</tr>
<tr>
<td>0.5</td>
<td>1.247 (0.036,41.790)</td>
<td>0.590 (0.070,3.136)</td>
<td>0.716 (0.067,2.316)</td>
<td>0.250 (0.007,6.184)</td>
</tr>
<tr>
<td>0.8</td>
<td>1.460 (0.050,44.267)</td>
<td>0.854 (0.068,14.310)</td>
<td>0.315 (0.012,1.769)</td>
<td>0.175 (0.006,1.862)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>((\tau_{11}))</td>
<td>((\tau_{01}))</td>
<td>((\tau_{10}))</td>
<td>((\tau_{11}))</td>
</tr>
<tr>
<td>0</td>
<td>0.941 (0.026,34.349)</td>
<td>0.534 (0.101,1.347)</td>
<td>0.933 (0.340,3.012)</td>
<td>0.561 (0.061,19.989)</td>
</tr>
<tr>
<td>0.2</td>
<td>0.867 (0.023,31.489)</td>
<td>0.533 (0.083,1.522)</td>
<td>0.938 (0.267,4.038)</td>
<td>0.397 (0.011,13.839)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.878 (0.025,30.539)</td>
<td>0.533 (0.068,2.310)</td>
<td>0.874 (0.126,5.702)</td>
<td>0.204 (0.006,5.479)</td>
</tr>
<tr>
<td>0.8</td>
<td>0.974 (0.032,30.698)</td>
<td>0.723 (0.065,10.399)</td>
<td>0.560 (0.024,5.423)</td>
<td>0.121 (0.004,1.186)</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1.264 (0.032,47.799)</td>
<td>0.533 (0.102,1.323)</td>
<td>0.889 (0.331,2.161)</td>
<td>0.707 (0.018,26.389)</td>
</tr>
<tr>
<td>0.2</td>
<td>1.341 (0.035,47.004)</td>
<td>0.530 (0.081,1.523)</td>
<td>0.859 (0.254,2.151)</td>
<td>0.562 (0.014,21.870)</td>
</tr>
<tr>
<td>0.5</td>
<td>1.413 (0.038,49.229)</td>
<td>0.553 (0.081,1.523)</td>
<td>0.782 (0.254,2.151)</td>
<td>0.356 (0.014,21.870)</td>
</tr>
<tr>
<td>0.8</td>
<td>1.707 (0.055,49.708)</td>
<td>0.706 (0.056,9.777)</td>
<td>0.463 (0.022,1.856)</td>
<td>0.214 (0.007,3.491)</td>
</tr>
</tbody>
</table>

Model comprising A3 (Lasso); B6 (Intermediate) and C9 (Saturated) predictors of compliance
Table 4 show results in terms of causal relative risks for models predicting compliance using three sets of predictors considered earlier: using 3, 6 and 9 predictors from Lasso, intermediate and all plausible predictors respectively. For mild values of sensitivity parameters, model selection using penalized maximum procedures (6 predictors) produced ‘best’ causal risk ratio estimates showing reduction of risks for both all-cause mortality and myocardial reinfarction. Specifically HRT treatment was consistently effective (reduced risks) among those who would comply with HRT only \( (\tau_{01}) \). The efficacy corresponding to this stratum was not dependent on the chosen value of sensitivity parameter. On the other hand the causal risk ratio estimates using 3 and 9 sets of predictors were comparable for all strata considered. In general, given a set of predictors, the results show same trend in principal effects with respect to change in magnitude of the sensitivity parameter \( \phi \) for both outcomes (mortality and myocardial reinfarction). Surprisingly these results using 3 and 9 predictors of compliance now suggested harmful effects (increased risks) of HRT treatment relative to placebo among those who would comply with either treatment.

Results from the sensitivity analysis above (Table 4) may be a useful demonstration of the phenomenon that causal (principal) effects are dependent on the choice of covariates predicting compliance. This may be an indication that the advantages of classical model selection are transferable to the Roy et al. [30] method via use of optimal marginal compliance models, i.e. comprehensive model selection may be useful in providing optimal predictors of compliance to enrich principal stratification. However, we note that while selecting plausible predictors of compliance is an integral component of the method, model selection only acts as an intermediate step that provides covariates for marginal compliance prediction models which are then joined into a causal model using the crucial but unknown sensitivity parameter.

In general we observe from results in both Tables 3 and 4 that for a given stratum and set of selected covariates predicting compliance, the change in resulting causal risk ratio estimates were more pronounced for the reinfarction compared to all-cause mortality outcome. This apparent interaction of outcome with sensitivity parameter \( \phi \) may be attributed to features of the two different outcomes. A possible explanation may be the fact that the choice of optimal predictors of compliance to make the conditional compliance assumption \( f \) valid might depend on outcome (death/reinfarction). Such an association may make conditional prediction assumption questionable for the Esprit data especially with regard to history of hysterectomy and cerebrovascular risks which are likely to be associated with better treatment compliance and subsequently favourable outcome.
VI. Discussion

By using optimal predictors of treatment compliance at mild values of the sensitivity parameter $\phi$, compliance with HRT tablets showed reduction in risks for both all-cause mortality and myocardial reinfarction. Compliance with HRT treatment compared to taking placebo among those who would comply with either treatment also indicated beneficial effects in reducing risks for both outcomes. Compliance with HRT treatment suggested beneficial effects compared to placebo for all other values of $\phi$ among the subgroup who would comply with HRT treatment only and those who would comply with either treatment allocation. However, causal risk ratios estimating efficacy of compliance to either treatment ($\tau_{11}$) had relatively wider mean 95% credible intervals compared to estimates for efficacy of compliance with HRT only ($\tau_{01}$) or placebo only ($\tau_{10}$). In general, the risk for myocardial reinfarction reduced with increase in the value of sensitivity parameter $\phi$ among those women who would comply with either treatment. On the other hand the risk of myocardial reinfarction increased with increase in $\phi$ among those who would comply with placebo treatment only. Overall as expected, for any chosen value of the sensitivity parameter $\phi$, the risks for both death and myocardial reinfarction were higher among those who would comply with placebo only compared to those who would comply with HRT only.

The variation in the HRT efficacy estimates from the Roy et al. [30] model may be an indication of difference in compliance behaviour between those allocated to placebo and HRT treatment. By adjusting for noncompliance in both arms, the Roy et al. method perhaps accounts for potential correlation between compliance behaviours in respective arms through the chosen value of the sensitivity parameter which implicitly makes the results depend on $\phi$. The fact that the results vary a lot with $\phi$ and yet we do not know its value suggests benefits of HRT treatment among those who comply when allocated it, i.e. strong monotonicity assumption (strong correlation in compliance behaviour between the two arms). Finally our analyses with flat priors may be considered adequate given the likelihood of a typical trial data to dominate such priors in addition to the fact that randomized trials are principally designed to be conclusive [16].

Noncompliance with treatment assignment in both arms of a clinical trial is likely to complicate efficacy estimation. Here the ITT provide a biased estimator for the true causal estimate even under homogeneous treatment effects assumption. Extending and applying the Roy et al. model to survival data (Esprit study) may be suitable by utilizing key covariates predictive of arm-specific compliance models to develop causal models linking the two marginal models. The resulting principal effects provides efficacy estimates for the different subgroups defined by compliance types. The method performed relatively ‘better’ than specialist randomization-based causal methods adjusting for noncompliance in one treatment arm only [25]. Simulation studies
indicated satisfactory performance of the method however, the results were heavily dependent on the choice of sensitivity parameters and hence may not be recommended in the presence of known heterogeneous treatment effects which produced large bias and wider corresponding 95% credible intervals. As a result, the method may only be recommended in the presence of sufficient information about compliance behaviours in respective arms.

Model selection in regression may be correctly considered as one of the most significant challenges in modern statistics [17]. Hitherto this challenge has not been extended to include prediction of compliance with treatment assignment in causal modelling. There are presently limited studies integrating model selection for compliance prediction in causal inference. While principal stratification has independently been demonstrated to provide better alternative identification strategies compared to selection model [23], integrating the two strategies may produce even more flexible models under relaxed assumptions. A record of plausible predictors of compliance can be used to effectively address identification problem of causal estimands by reducing bias and relaxing the implicit assumptions [20]. From a clinical perspective, knowledge about predictors of compliance may be a valuable tool to inform treatment decisions. As a result, there is need to adopt comprehensive model selection methods for accurate prediction (of compliance). After model selection, there is further need to use suitable validation indices (e.g. optimism, calibration and discrimination indices) to evaluate performance of selected models. With potentially many recorded baseline covariates, using penalized regression techniques may be recommended for building compliance prediction models.

The merits of standard model selection procedures are transferable to the principal stratification method adjusting for noncompliance in two treatment arms by linking the respective optimal marginal compliance models into an association model [26]. However, application of the method is premised on the plausibility of a defining assumption that potential outcome is independent of the set of covariates predicting treatment compliance for a given stratum. This assumption may be questionable for the Esprit data especially with regard to history of hysterectomy and cerebrovascular risks which potentially have a higher likelihood to be associated with treatment compliance leading to possible efficacy. For example, while the unpleasant experience of bleeding may affect treatment compliance negatively, those with history of cerebrovascular risks may comply with their treatment allocation with a hope to derive potential protective benefits. Further sensitivity analysis on the departure of conditional prediction assumption implicit in the Roy et al. model may be conducted using alternative methods which incorporate less stringent assumptions. For example, by adopting the Bayesian framework introduced by Long et al. [21] to model the principal compliances directly in multitreatment arms and for more general outcomes by treating the principal compliance status as missing data instead of joining them with a user-defined sensitivity parameter $\phi$.  

Imhotep Proc.
Although principal stratification provides a powerful framework which is extendible to analyse complex surrogate outcomes like ‘truncation by death’ where death occurs before a primary outcome of interest is recorded hence resulting in censored records [31], the method’s application and usefulness may be limited to intermediates with fewer categories (e.g. binary) [40]. Although the all-or-nothing compliance suitably applied to the Esprit data, principal stratification in general produce inconsistent causal estimate for a truly continuous stratification variable but which has been coarsened for analysis [28]. As a result, policy informed by analysis based on principal stratification should be implemented with caution owing to the fact that the principal strata themselves remain unidentified.

References


Lang’o Odondi
e-mail: Lango.Odondi@bristol.ac.uk
MRC ConDuCT Hub, School of Social and Community Medicine, University of Bristol, UK
The need for effective study skills under the 21st century: a case of USA and KENYA

Peter T. Olszewski
M.S Penn State Erie, The Behrend College
pto2@psu.edu

Dickson S. O. Owiti
Kisii University, Department of Science Education, P. O. Box 408(40200), Kisii, Kenya.
owitidick@yahoo.com

Abstract

The current world is operating in an economy that values creativity and innovation for scientific and technological development. Education gives people appropriate skills and knowledge they need to address their social problems. Mathematics and Science education is at the centre of this and needs to be at the forefront to connect the present to the future. The fact that a new generation of learners is in our classrooms requiring a paradigm shift in pedagogy is indisputable. Teaching in the same old way and emphasis on examinations, grades, certificates as well as lack of basic facilities have affected learning by generation Y students. As a result, Kenya like the United States of America faces a myriad of problems despite the fact that the youth is a reach reservoir for development. More than 50 per cent of the world's gold reserves, diamond, manganese, chromium, and cobalt are in Africa yet Africans live in the poorest situations imaginable. The United States, despite being the most powerful nation on the planet has, in general, have poor test scores in mathematics if results of international comparative studies are anything to go by. This paper argues in addition to poor teaching methods, strategies, and techniques, the assumption that students know how to study mathematics once in secondary school, college, and university and the failure to teach the same is partly to blame since year after year, students either drop out, receive poor grades, fail to attend classes and or don’t take mathematics seriously. Millennials therefore need to be taught study skills in mathematics to ensure quality mathematics learning for creativity and innovativeness in the citizens. This will ensure education empowers Kenya, Africa, and the United States for global competitiveness. In particular, this paper intends to address the following current issues in Kenyan and Unites States schools: 1. Describe the Millennial Student, 2. Ramifications for Kenya and the United States, 3. Kenyan and United States curricula, 4. How to teach effective study skills, 5. What is needed of educators, and 6. What to do in the future.
The need for effective study skills under the 21st century: a case of USA and KENYA

Peter T. Olszewski and Dickson S. O. Owiti

Abstract. The current world is operating in an economy that values creativity and innovation for scientific and technological development. Education gives people appropriate skills and knowledge they need to address their social problems. Mathematics and Science education is at the centre of this and needs to be at the forefront to connect the present to the future. The fact that a new generation of learners is in our classrooms requiring a paradigm shift in pedagogy is indisputable. Teaching in the same old way and emphasis on examinations, grades, certificates as well as lack of basic facilities have affected learning by generation Y students. As a result, Kenya like the United States of America faces a myriad of problems despite the fact that the youth is a reach reservoir for development. More than 50 per cent of the world’s gold reserves, diamond, manganese, chromium, and cobalt are in Africa yet Africans live in the poorest situations imaginable. The United States, despite being the most powerful nation on the planet has, in general, have poor test scores in mathematics if results of international comparative studies are anything to go by. This paper argues in addition to poor teaching methods, strategies, and techniques, the assumption that students know how to study mathematics once in secondary school, college, and university and the failure to teach the same is partly to blame since year after year, students either drop out, receive poor grades, fail to attend classes and or don’t take mathematics seriously. Millennials therefore need to be taught study skills in mathematics to ensure quality mathematics learning for creativity and innovativeness in the citizens. This will ensure education empowers Kenya, Africa, and the United States for global competitiveness. In particular, this paper intends to address the following current issues in Kenyan and Unites States schools: 1. Describe the Millennial Student, 2. Ramifications for Kenya and the United States, 3. Kenyan and United States curricula, 4. How to teach effective study skills, 5. What is needed of educators, and 6. What to do in the future.

Keywords. Kenya, United States, millennials, education, globalization, study skills, empowerment.

Theme: Nurturing the Future Generation of Mathematics Researchers in African Universities
Sub-theme: Teaching Mathematics Study Skills to Millennial Students.

Vaughan Monroe: “It now costs more to amuse a child than it once did to educate his father”

Objective:

Each generation of students entering college has their own set of distinguishing qualities. The purpose of this article is to introduce the characteristics associated with the millennial
I. Introduction

A new generation of learners has immersed themselves in the classroom. Our old ways of teaching are no longer suitable to serve their interest. They are infatuated with technology, impatient, and give up quite easily on tasks that appear challenging. They bring unique learning style preferences and worldviews to class (Monaco and Martin, 2007). The challenge is for today’s teacher to attract and engage this generation of learners, particularly in Mathematics and the Sciences. There is a general concern by employers, educators, and the general public that today’s students are leaving school and colleges without the necessary skills needed for survival, the world of work, and global competitiveness (Mugo, 2013; World Bank, 2013). Quality of education programs offered at schools and universities is thus questionable as graduates are unable to meet the demands of 21st century jobs (Mugo, 2013; Wanga, 2013). Global innovative and competitiveness indexes reveal Kenya and the rest of Africa trail behind comparative regions such as Southeast Asia, Latin America, and the Caribbean; an indication all is not well in our classrooms and institutions (World Bank, 2013). Furthermore, Africa trails behind other regions in its use of ICT and the gap is large in the basic building blocks for competitive economy such as infrastructure, education, governance, and institution. Despite being the most powerful nation under the sky, American students’ poor test scores in international comparative studies on mathematics and sciences (TIMSS, PISA and OECD) is of great concern (Stigler & Hiebert, 1999; Stevenson & Stigler, 1992). Why is this so? What is the overall impact of this?

From a college professor’s point of view, there is an assumption that students know how to study when first starting their college careers (Kiewra, 2009). The reality is, however that, many students don’t know how to study mathematics. Despite a number of measures and reforms such as the introduction of calculators and SMASSE (Strengthening of Mathematics and Science in Secondary Education) project in Kenya, the score card in mathematics has persistently made depressing reading (Birgen, 2004; Owiti, 2012). Application of technology in the teaching of mathematics in the United States seems not to be working any better than the talk and chalk in Japan and China (Stigler and Hiebert, 1999; Stevenson and Stigler, 1992). Besides, transitioning from high school learning to college or university learning seem a major jump and semester after semester, students either drop out, receive poor grades, don’t attend class, and or don’t take studies seriously (www.mathtutor.ac.uk). Why is this true? Though no known research in either Kenya or the US relates the poor performance in mathematics to ineffective study skills, several studies have indicated that an effective route to improving self-efficacy in mathematics is by teaching specific learning strategies (Peter and Gareth, 2010). The authors believe therefore that lack of students’ learning of effective study skills in mathematics and teachers’ view of the same as remedial and their failure to think twice about teaching it could be behind the poor performance in the subject (Kiewra, 2009).

II. Who are Millennials?

The millennial student is a learner characterized by a need to have immediate feedback, a sense of entitlement, lack of critical thinking skills, unrealistic expectations, high level of parental involvement, and an expected “how to guide to succeed” in and out of the classroom (Monaco & Malissa, 2007). They want to spend less time on tasks and reach success with little effort. They are over “babied” by parents with little free time at childhood. This has decreased opportunity
Vol. 1 (2014) The need for effective study skills 27

for independent creative thought and decision making skills and as a result, provide challenges for teachers and educators.

They are team oriented and less comfortable working independently for fear of higher risks of personal failure. Consequently, they prefer working cooperatively on projects and participate within collaborative group settings (Monaco & Malissa, 2007). They are confident and highly optimistic but less committed, have big dreams and expectations with an unclear path on how to reach the level of success they are so confident of. This confidence stems from easy attainment of success in early years of education (primary schools) with very little effort (multiple choice questions) and their ability to work with various technologies (Monaco & Malissa, 2007). They are less committed to their work and quickly become frustrated when they do not achieve an A or B grade. They have access to each other and information 24/7 through the internet, text message, face book and twitter but still can’t make use of the same technologies for academic excellence (Olszewski, 2013). They have been technology stimulated and demand this counteractively as they go through their studies. They feel pressured to constantly perform to impress on their would be judge, acquire good grades and certificates, and as such have perfected on a number of academic frauds to achieve the end (Wachira, 2012):

- Giraffing- where a student sticks out the neck to see another student’s answer sheet.
- Lateral connection- implemented through a sitting arrangement whereby the bright candidate is flanked on both sides by weaker candidates who peep on their work.
- Nothing-nothing- where empty biro pens are used to trace information on a blank white piece of paper. Seeing this on the table, one would think there is nothing on the paper, but closer observation reveals well loaded with facts about the examination.
- Mwakanya (dubbing and tattoos)- where a candidate, mainly female, writes information on thighs where she can easily adjust to read and copy.
- Missile catch- where answers are written on a piece of paper, squeezed, and thrown to a colleague during the examination.
- Body aids- such as use of slippers with answers, under-pants, handkerchiefs and use of coded sign language to communicate during examinations particularly in multiple-choice examinations.
- Table top guide- where anticipated answers are written on desks before an examination starts. Mostly, formulas, diagrams and maps are written in short form.
- Use of tokens- These are short notes on the mathematical set, razor blades, rulers, handkerchiefs and other items for referencing during examinations.
- Coms- where calculators that facilitate multiple entries are used. Quite often such special calculators are put in casings of ordinary calculators and might look ordinary and escape invigilators.
- Direct access- this is whereby an examiner helps candidates during examinations.
- Mercenary service- involves another person writing the examination for the other i.e in 2011 KCSE some university students were arrested sitting examinations for private candidates.
- Rank zeroxing- which happens when a candidate collects and copies a colleague’s answers word for word.
- Contract- This occurs when a student’s grade is influenced with the assistance of a friendly lecturer. It can occur when a lecturer is paid money or sex for marks to enhance grades.
- Time out- This is where a student pretends to be suffering from diarrhea and visits the toilet several times to read prepared answers.

Youngsters live in a world dominated by reality television and celebrities where success appears to come instantly and without any real effort. This culture of instant gratification could be making today’s school children harder to teach mathematics than ever. Overwhelming distractions preventing the millennials from deep study of mathematics, creativity, innovation,
and thinking are impacting negatively on their grades and attitude towards learning (Olszewski, 2013). In Kenya, cases of students committing suicide on failure to attain desired grades in national exams (KCPE and KCSE), teachers being beaten and frog matched by parents and students due to poor examination results are abound. Perhaps these are confirmations that things are not right in our classrooms. In particular, Kenyan millennials are lazy, focus much on grades and certificates, dream big but do not work towards achieving their dreams, are carefree, expect technology to work for them (calculators and computers in mathematics), impatient and give up quite easily. This casts a lot of doubt on the success of president Uhuru’s “laptop for every standard one pupil” unless the implementation is well done and study skills are taught.

A large group of students worldwide who never achieve in mathematics either lack good study habits or do not understand how to study for mathematics (Dawkins, 2013). A few spend hours each day studying but still don’t do well while others because of inefficient study habits or distractions simply do not spend enough time studying mathematics (Dawkins, 2013). This latter group need to be made to realise that they need to take their time if they want to pass mathematics, hence the need for mathematics study skills.

III. How did we get here?

Special education came to the forefront in the 1970s in the United States. Educators believed the best course of action was to change the curriculums and not try to change the student (Hopkins, 2010). However, what if the student needs an alternative way of learning but can still perform just as well as non-special educated student? For example, instead of having classrooms set up in traditional rows, have a diverse educational level group of three or four students assigned to circular tables. This way, students can work together to solve problems, work on homework, and exchange ideas. Teachers must encourage healthy debates instead of passing out worksheets to students where all they need to do is follow patterns to get solutions. Instead, ask questions that motivate discussion. For example, why, in general, does $(a + b)^2$ not equal $a^2 + b^2$?

In Kenya the idea of special needs education started after Kamunge’s report of 1988. Currently, there are over 1,100 units and 100 public special schools in Kenya, which include vocational and technical institutions that cater for learners with special needs and disabilities. It is estimated three quarters of Kenyan pupils with special educational needs are in special schools with only a quarter in special units located within mainstream schools. Special education in Kenya faces a number of challenges: Special education has not been fully mainstreamed across the education sector and so there are difficulties exercising the right to education because of the lack of access at all levels. Most regular schools are ill-equipped to deal with special needs learners. There is lack of specialized learning resources, materials and equipment. Teachers with specialist knowledge and relevant skills are in short supply. The Ministry of Education’s Quality Assurance and Standards Division does not regularly inspect special education learning institutions. Where quality assurance is undertaken, most officers generally lack expertise to do so. There is lack of equity because many regions are without special needs institutions particularly in ASAL and slum areas.

In the United States, special education is proving to be very difficult to comprehend. Teachers are required, by Federal Laws, to comply with Individual Education Plans (IEPs) and must make accommodations for these students. Of course, if a student needs extended testing time or large font sizes, these are reasonable accommodations. However, many students are sent to their special education teachers when other, non-special education students are taking an assessment in the classroom. A teacher would think their assessment questions would not be altered in anyway, providing the same level of difficulty for all students in the class. However, many special education teachers are writing out examples of problems on the margins of assessments and changing questions. Regular education teachers can’t do anything to prevent
this from happening, as we, as a society, must believe all students are "special." When student receives feedback, they are given a high grade. However, what have they really learned? Not the mathematics or, for that matter, any other subject. To lead students along and making them think this is acceptable in high school and even worse, later on in life, is a crime. Thinking about how many students we can work with and help them learn the same material, on the same level of difficulty as non-special educated students, should be the goal of special education teachers. To sit back and not motivate learning is hard to fathom in one of the most powerful countries in the world and a developing nation. Kenya and the United States are taking potential scholars and not providing them with a curriculum necessary for the future. With both countries, special education is robbing students of their true potential and instead of encouraging students to work hard and study, the learning and studying is being done by regurgitation and handholding to give the impression students are learning and advancing.

IV. What this means?

Mathematics and Science education being the single most powerful leverage we have to address the needs of our future is at stake. The generation Y students (millennials), with their unique characteristics that affect what and how students learn, go through education but education never goes through them. They fail to learn what they ought to have learnt by different levels of education (UWEZO, 2010, 2011). This is partly due to distractions, teachers less innovativeness in teaching practices and poor study skills resulting into exam cheatings and students with certificates but without ability to engage in problem solving (The Standard, 2011; Abagi, 2012; Ayiro, 2012; Mugo, 2012). In a survey entitled "where to be born" at the end of 2012, the Economic Intelligence Unit (EIU) ranked Kenya the second worst country under the sun for a child to be born in 2013 (Kiarie, 2013). The annual Global Peace Index (GPI) ranked Kenya in position 120 out of 158 countries while in the Transparency International (TI) corruption index of 2012, Kenya was ranked the 35th most corrupt country in the world. A look at certain global indexes reveals the following for Kenya:

<table>
<thead>
<tr>
<th>Sub indexes</th>
<th>Kenya’s Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Competitiveness Index (GCI)</td>
<td>106/144</td>
</tr>
<tr>
<td>Higher education and Training</td>
<td>101/144</td>
</tr>
<tr>
<td>Quality of Education System</td>
<td>37/144</td>
</tr>
<tr>
<td>Quality of Primary Education</td>
<td>78/144</td>
</tr>
<tr>
<td>Quality of Math and Science Education</td>
<td>76/144</td>
</tr>
<tr>
<td>Innovation</td>
<td>50/144</td>
</tr>
<tr>
<td>Human development index</td>
<td>143/187</td>
</tr>
</tbody>
</table>


From the Table 1, it is clear Kenya is not among the top ten countries when it comes to innovativeness, quality of education system, primary education, as well as mathematics and science education. Kenya is towards the bottom of the list with respect to human development, global competitiveness as well as higher education and training. In 2009, 22 of the 24 nations identified as having "Low Human Development" on the United Nations (UN) human development index were in Sub-Saharan Africa to which Kenya belongs (www.hdr.undp.org/en/statistics). This state of affairs could be attributed to the fact that millennials are not learning mathematical skills as they should be.
In the United States, high school seniors take the SAT exams, which provide an indicator for college readiness. According to the College Board, critical reading is down by four points (505 in 2000 and 501 in 2010) and mathematics is up by two points (514 in 2000 and 516 in 2010) within the last ten years. Writing has gone down each year since first tested in 2006 (497, 494, 494, 493, and 492). Fifty-seven percent of students who took the SATs did not score high enough to indicate likely success in college. Of course, the SATs are not the only indicator of college readiness as GPAs and other extra-curricular activities are weighed in a college or university decision for admission. As a comparison to the global indexes found in Kenya, the Unites States global indexes are as follows:

Table 2. Some global indexes of the United States 2011-2013

<table>
<thead>
<tr>
<th>Sub indexes</th>
<th>USA’s Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Competitiveness Index (GCI)</td>
<td>7/144</td>
</tr>
<tr>
<td>Higher education &amp; Training</td>
<td>8/144</td>
</tr>
<tr>
<td>Quality of Education System</td>
<td>28/144</td>
</tr>
<tr>
<td>Quality of Primary Education</td>
<td>34/144</td>
</tr>
<tr>
<td>Quality of Math &amp; Science Education</td>
<td>47/144</td>
</tr>
<tr>
<td>Innovation</td>
<td>7/144</td>
</tr>
<tr>
<td>Human development index</td>
<td>3/187</td>
</tr>
</tbody>
</table>


From the Table 2, the Unites States is among the top ten countries for global competitiveness, higher education, innovation, and human development. However, with being the most powerful country in the world, one would expect to have very high indexes for education. Comparing the quality of mathematics and science education with the United States and Kenya, there is only a difference of 29, in primary education, a difference of 44 in higher education, and in the educational system as a whole, a difference of 9. The latter difference is most striking given the higher rankings of the United States in terms of innovation, human development, the economy, and financial stability of each country. Why the small difference? The answer, we believe is, no matter the country, we live in a world that is constantly changing, much more than in previous generations. The way we as humans interact and have access to technology has changed our outlook for the future. Students are no longer wanting to spend hours reading books and studying on their own. They want the answers now, put as little effort as possible, and study only when they have time around their busy virtual worlds.

V. United States Curriculums

As the United States moved from generation to generation, from the Industrial Age to the new modern era of the Information Age, schools have been teaching students the same way. Classrooms have been set up in rows and the teacher did all the lecturing. Until recently, schools have been under the influence of TTWWADI, which stands for That’s The Way We’ve Always Done It. The world is constantly changing and often times, teachers cannot keep up with this change. As McMain, Jukes, and Kelly point out: “We cannot carry on preparing students for the farms and factories of yesterday while the world jumps to light speed with biotechnology, nanotechnology, neurotechnology, global high speed wired and wireless networks, and incredibly powerful personal portable devices. We strongly believe that schools must prepare kids for the world of tomorrow - the world they will spend the rest of their lives.” Lessons and the curriculum must be designed with the use of technology that will not further inhibit students to learn.

When we think about our college years, we always remember going to class, listening through lecturers, performing labs, studying, and taking exams. This is how college operates
because of TTWADI. However, what we don’t think of is someone teaching us how to study when we were freshmen. It is assumed upon graduation from high school, one knows how to study and is ready for college studies. However, for many millennial students, this is not the case. There are big differences between high school level learning and college level learning and most graduating high school seniors have no comprehension of how to study. As Kiewra and Jairam point out, "Many believe that college students should be expert learners; after all, they have practiced learning for a dozen or more years. In reality, 73% of college students report difficulties preparing for exams, and this percentage of reported study problems is consistent across college years. Research also confirms that college students employ weak strategies in the classroom and while studying. Those weak strategies include poor note taking, organizing ideas linearly, learning in a piecemeal fashion, and employing redundant strategies." There is a gap between how professors expect students to learn and how students actually learn. To close the gap, students have to first realize the differences and then take control of their own learning. Second, students have to be equipped with effective learning strategies.

As part of the No Child Left Behind Law (NCLB) signed by President Bush, high schools across the United States, are seeing a growth of students enrolled in a non college preparatory mathematics plan of study. However, students receive their high school diploma and are accepted into a community college, college, or a university. However, with their high school education, they are not prepared for the challenges ahead. Currently, a typical non-college preparatory mathematics curriculum involves Algebra I, Liberal Arts Mathematics designed to get students ready for Geometry, and Geometry. In a college preparatory curriculum, students take Algebra I-II, Geometry, Pre-Calculus, and either Calculus or Statistics. Views are mixed if the No Child Left Behind Law is working. Some states are seeing improvements in mathematics, while others aren’t. Below are two charts of two American states with a comparative view before and after NCLB according to Center on Education Policy (CEP).

**Table 3.** Average yearly gains by subject and grade for states with greater overall gains after NCLB than before.

<table>
<thead>
<tr>
<th>Subject / Grade Level</th>
<th>Pre-NCLB Percentage Point Gain</th>
<th>Post-NCLB Percentage Point Gain</th>
<th>Average Yearly Effect Size Gain Pre-NCLB</th>
<th>Average Yearly Effect Size Gain Post-NCLB</th>
<th>Comparisons of Pre vs. Post-NCLB Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math 5</td>
<td>0.1</td>
<td>3.5</td>
<td>0.03</td>
<td>0.11</td>
<td>Post-NCLB gains exceed Pre-NCLB gains on 8 of 12 comparisons</td>
</tr>
<tr>
<td>Math 8</td>
<td>0.7</td>
<td>2.0</td>
<td>0.03</td>
<td>0.06</td>
<td>Post-NCLB gains exceed Pre-NCLB gains on 8 of 12 comparisons</td>
</tr>
<tr>
<td>Math 11</td>
<td>1.7</td>
<td>0.0</td>
<td>0.03</td>
<td>0.02</td>
<td>Post-NCLB gains exceed Pre-NCLB gains on 8 of 12 comparisons</td>
</tr>
</tbody>
</table>

**Table 4.** Average yearly gains by subject and grade for states with greater overall gains before NCLB than after.

<table>
<thead>
<tr>
<th>Subject / Grade Level</th>
<th>Pre-NCLB Percentage Point Gain</th>
<th>Post-NCLB Percentage Point Gain</th>
<th>Average Yearly Effect Size Gain Pre-NCLB</th>
<th>Average Yearly Effect Size Gain Post-NCLB</th>
<th>Comparisons of Pre vs. Post-NCLB Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math 4/5</td>
<td>3.9</td>
<td>3.2</td>
<td>0.10</td>
<td>0.08</td>
<td>Post-NCLB gains exceed Pre-NCLB gains on 8 of 12 comparisons</td>
</tr>
<tr>
<td>Math 8</td>
<td>4.1</td>
<td>3.6</td>
<td>0.10</td>
<td>0.05</td>
<td>Post-NCLB gains exceed Pre-NCLB gains on 8 of 12 comparisons</td>
</tr>
<tr>
<td>Math 10</td>
<td>4.2</td>
<td>3.0</td>
<td>0.11</td>
<td>0.08</td>
<td>Post-NCLB gains exceed Pre-NCLB gains on 8 of 12 comparisons</td>
</tr>
</tbody>
</table>

Source: Center on Education Policy.
As we can see, Pennsylvania did see an improvement in grades 5 and 8. On the other hand, Delaware has not seen improvements. The next logical question is, have curriculum's changed along with NCLB? Are we teaching with the same level of rigor before NCLB? This is also a mixed answer. Some schools have created new mathematics courses that don’t motivate students to study higher-level mathematics. For example, many United States high schools have different sequences of mathematics courses for students to take. According to the Center for the Study of Mathematics Curriculum, 24 States require three years of mathematics courses. However, none of those 24 States require courses past Geometry. Only Arkansas and Michigan require mathematics course through Algebra 2 (see figures below).

**Table 5. Number of years of high school mathematics course/credits required for graduation in the United States.**

<table>
<thead>
<tr>
<th>Specified at Local Level</th>
<th>1 year</th>
<th>2 years</th>
<th>3 years</th>
<th>4 years</th>
<th>Varies by Diploma</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO, IA, ME, MA, WI, NE</td>
<td>AK, AZ, CA, ID, MT, ND, WI</td>
<td>CT, DC, DoDEA, HI, IL, KS, KY, LA, MD, MN, MO, NH, NM, NJ, NY, OH, OK, OR, PA, TN, UT, VT, WI</td>
<td>AL, AR, DE, FL, MI, MS, RI, SC, TX, WA, WV</td>
<td>IN (2-4 yrs) GA (3-4 yrs) NC (3-4 yrs) SD (3-4 yrs) VA (3-4 yrs)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>7</td>
<td>24</td>
<td>11</td>
<td>5</td>
</tr>
</tbody>
</table>

**Table 6. Courses required for United States high school graduation/diploma.**

<table>
<thead>
<tr>
<th>Course</th>
<th>States Requiring Course</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algebra I</td>
<td>AL, AR, CA, DoDEA, DC, FL, GA, IL, KY, MD, MI, MS, ND, NH, NM, OK, SD, TX, UT</td>
<td>19</td>
</tr>
<tr>
<td>Algebra I or Integrated Mathematics I</td>
<td>IN, LA, NC, TN</td>
<td>4</td>
</tr>
<tr>
<td>Geometry</td>
<td>AL, AR, DoDEA, IL, KY, MD, MI, TX, UT</td>
<td>9</td>
</tr>
<tr>
<td>Geometry or Integrated Mathematics II</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Algebra II</td>
<td>AR, MI</td>
<td>2</td>
</tr>
<tr>
<td>Algebra II or Integrated Mathematics II</td>
<td>DE</td>
<td>1</td>
</tr>
<tr>
<td>Algebra I, Geometry, Algebra II or Integrated Mathematics I-III</td>
<td>LA, TN, VA</td>
<td>3</td>
</tr>
</tbody>
</table>

Source: Center for the Study of Mathematics Curriculum.

As a supplement for high school graduation, students may register in a developmental course, a senior seminar, or a capstone course. Regardless of the fact that some students maybe going for a Liberal Arts degree after high school, will still need to take some form of mathematics courses in college. However, with the lack of students the non-college preparatory track in high school, students will feel at a loss. There is simply not enough in the curriculum to motivate critical thinking skills. Many students across the United States face the reality they must first take a remedial mathematics course when starting their college career in order to be placed in a College Algebra II course needed for most programs. However, College Algebra II is very similar to Algebra II taught in high schools, which is not taught in the non-college track. We are not doing students any favors in our high schools by letting students enroll in a curriculum that doesn’t prepare them for future studies. We should be encouraging students and parents to help make our students the best they can be and giving them all the knowledge possible for success in college and for the future.

**VI. Kenya’s curriculum**

From independence, except in the recent past when electronic calculators were introduced in secondary school mathematics and now president Uhuru’s 53 billion laptop project for class one pupils earmarked for January 2014, Kenyan generation after generation have been taught in the same old way devoid of ICT. Classrooms have been set up in straight military rows and the teacher did all the lecturing. Given the calculator innovation has not been very successful in enhancing performance in mathematics, alternatives A and B curriculums were introduced at secondary schools in 2010 to cushion weak students. Kenyans are yet to see the implementation

---

1Percentage proficient data are for grade 5, and effect size data are for grade 4.

Imhotep Proc.
of the laptop project and whether it will enhance mathematics learning or not. The Dakar World Education Forum (2000) declaration of EFA saw the introduction of free primary and day secondary education in 2003, a move that resulted to an increased enrollment of up to 10 million pupils at primary schools and a transition rate of 72% from primary to secondary by 2010. Despite this achievement, mathematics has remained a much feared and poorly performed subject in Kenya (Eshiwani, 1987, 1993; Owiti, 2008). Learning assessment by UWEZO (2010, 2011) found that pupils are leaving school without the ability to perform the most basics sums. For instance, 3 out of every 10 children at primary schools can’t perform basic arithmetic, 70% of children in class 3 can’t do class 2 work while 60% of standard eight pupils can’t solve problems tackled in lower classes (Abagi, 2012). Furthermore, Kenyan 13 year olds are behind their counterparts in Mauritius in mathematics (SACMEQ, 2010). Effective pupil learning and achievement in mathematics is hampered by weaknesses in teachers’ pedagogical classroom practice (Pontefract and Hardman, 2005). Dominant is individual seatwork teaching activity in mathematics while recitation is dominant in English (APHRC, 2010-2011). Interaction discourse comprises of teacher explanation and question and answer sequences that are ritualistic exchanges (Pontefract and Hardman, 2005). The lessons are reiterative rather than developmental in nature to ensure progression in learning. Focus is on revision of a topic previously taught rather than the development of a concept or introduction of new ideas. Most of the learning tasks put strong emphasis on factual, propositional knowledge (knowing that) rather than procedural knowledge (knowing how). Teachers’ questions (99%) are closed and are factual narrow requiring recall and single word responses (Pontefract and Hardman, 2005). Majority of teachers of mathematics are not conversant with the syllabus they are teaching and learning is too teacher knowledge (knowing that) rather than procedural knowledge (knowing how). Teachers’ questions (99%) are closed and are factual narrow requiring recall and single word responses (Pontefract and Hardman, 2005). Majority of teachers of mathematics are not conversant with the syllabus they are teaching and learning is too teacher knowledge (knowing that) rather than procedural knowledge (knowing how).

At Secondary School level, 10% of students like mathematics and perform well in it. The rest 90% dislike the subject, perform poorly in it (average 25%) and have to be forced to study it (KNEC, 1999; Ayodo, 2009). A baseline survey by SMASSE in 1998 in nine districts (Kajiado, Gucha, Kakamega, Lugari, Butere-Mumias, Kisii, Maragua, and Makueni) revealed problems to do with attitudes towards science and mathematics, inappropriate teaching methodology; content mastery, inadequate assignments, few or no interaction for teachers, infrequent professional guidance and missing links between primary and secondary school levels. Analysis of Kenya Certificate of Secondary Education (KCSE) mathematics results over the years by the Kenya National Examinations Council (KNEC) reveal that Kenya is producing a generation that is mathematically incompetent (Daily Nation, 1996). Over 80% of the candidates, majority of who are girls, are failing in the subject (Daily Nation, 1996). For instance, between 2001 and 2006 the mean score in mathematics was 19% (Walikenga, 2003; 2nd International handbook of mathematics education part one). The introduction of the new mathematics curriculum (alternative A and B) in 2010 for students without ambition of pursuing careers requiring advance mathematics and science is an indication that all is not well at secondary schools Kenya (Otieno, 2009). There is very little or no use of teaching aids and or technology (ICT). Use of calculator, an innovation in the curriculum, has not yet succeeded partly because teachers lack the pedagogical competent in making the correct use of the aid thanks to the mathematics education curriculum teachers go through in college. It has remained traditional and conservative and so teachers leave colleges without knowledge of incorporating ICT in their teaching. Fingers point at the school-based programs (TSC) at the universities for producing incompetent teachers who can’t work with students to produce results (Ayiro, 2012).

At Colleges and Universities, shortage of qualified and experienced staff is forcing colleges and universities to resort to unqualified lecturers, many of whom are still pursuing masters courses (Mugo, 2012). Kenya’s former higher education minister Prof. Kamar admitted that quality of university education is dwindling and calls for deliberate measures to end the same (Oduor, 2012). Lecturers and Professors are less innovative in their teaching and are driven by economic gain (commercial mentality) at the expense of meaningful teaching often resulting into mediocre teachers of mathematics (Mugo, 2012). Teacher education curriculum doesn’t adequately address the issues pertinent to secondary school mathematics teaching (CEMASTEA, 2010-2011). Theories in the curriculum are often outdated and inapplicable in the today’s classroom (CEMASTEA, 2010-2011). There are no tutorials and emphasis is on theory with haphazardly organised school-based practices lasting just a few months (Owiti, 2012). Teaching practice (TP) assessment results are never analysed to inform subsequent mathematics teacher education resulting into a continuous supply of teachers who are unable to embrace and apply technology in their teaching and are uncommitted to the profession (Alianga, 2012). Furthermore, the best students never opt for B.Ed programmes. The majority of those who join school of education have had to "bridge" meaning they are weak in mathematics, the subject they are expected to teach after graduation. The end result is a Kenyan child who is less innovative.

Thinking about Kenya and the United States education a few years back, no one gave the now famous motivational talks or taught students how to study. Perhaps there were fewer distractions comparatively and students would spend time studying mathematics. For many millennial students, this is not the case. Students are at a loss and constantly need guidance. Many students across Kenya and the United States face the reality that they must first take a remedial mathematics course or bridge in mathematics in order to be placed in college programs requiring mathematics. To close the gap, students have to be equipped with effective mathematics learning strategies. Most teaching and learning interventions are not designed with millennial needs in mind.
Mathematics is a cumulative subject. Students must recall past concepts in order to be successful on future exams and classes. In Algebra I, students are first introduced to simplifying complex fractions. This skill has many practical applications in higher mathematics courses. For example, if a student is asked to prove \( x \) to get \( 2 \times 6x = 16 \) must be written as \( x^2 - 6x - 16 = 0 \), before applying the formula will make it difficult to handle the problem (Paul, 2013). Besides, Mathematics is cumulative. For instance one cannot do logarithms without first taking indices e.g., solve for \( x \) in the equation \( 2^x = 55 \) giving your answer correct to 2 decimal places. Again a candidate failing to give the answer as required in this case may not score all the credit. In addition, millennial needs to be taught General tips for studying mathematics, taking notes, getting help, homework, problem solving, studying for exams and taking a mathematics exam.

Memorization of simple formulas does not work in mathematics as one has to in addition understand how to use the formula. For instance, simply memorizing the quadratic formulae:

\[
x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}
\]

without understanding that one has to first express a quadratic in standard form i.e., \( x^2 - 6x = 16 \) must be written as \( x^2 - 6x - 16 = 0 \), before applying the formula will make it difficult to handle the problem (Paul, 2013). Besides, Mathematics is cumulative. For instance one cannot do logarithms without first taking indices e.g., solve for \( x \) in the equation \( 2^x = 55 \) giving your answer correct to 2 decimal places. Again a candidate failing to give the answer as required in this case may not score all the credit. In addition, millennial needs to be taught General tips for studying mathematics, taking notes, getting help, homework, problem solving, studying for exams and taking a mathematics exam.

Mathematics is a cumulative subject. Students must recall past concepts in order to be successful on future exams and classes. In Algebra I, students are first introduced to simplifying complex fractions. This skill has many practical applications in higher mathematics courses. For example, if a student is asked to prove

\[
\frac{\sin^2 \theta - \tan \theta}{\cos^2 \theta - \cot \theta} = \tan^2 \theta,
\]

the natural starting point is to write

\[
\frac{\sin^2 \theta - \tan \theta}{\cos^2 \theta - \cot \theta} = \frac{\sin^2 \theta}{\cos^2 \theta} - \frac{\sin \theta}{\cos \theta}.
\]

However, many students make the classical mistake of cancelling out

\[
\frac{\sin \theta}{\cos \theta} - \frac{\sin \theta}{\cos \theta}
\]

to get \( \frac{\sin^2 \theta}{\cos^2 \theta} = \tan^2 \theta \). Clearly, there is a misunderstanding of the basic concepts of complex fractions in a PreCalculus class. In addition, students may not have the proper exposure to higher level thinking skills or not enough exposure for them to be successful in college.

PowerPoint lectures are becoming more commonplace for classes. However, students should still take notes about the PowerPoint lectures as many PowerPoint’s only have pictures or key ideas of topics. If students totally put faith in printing out the PowerPoint’s once an instructor has completed talking about them, they may be faced with leaving out vital information. This is especially true in an art history class where the PowerPoint lectures may only be pictures of art while the details about the works are pouring out of the professors’ lips. While the typical millennial studies, many times, students are memorizing isolate facts. Whether it is due to over highlighting, underlining, making note cards, if one can’t make connections and draw on topics, one will never be successful in class. Think about it, are the vast majority of exams in college going to ask students isolated and, quite possibly, boring facts about when certain events took place? Or when someone did something special? Highly doubtful. Why? Professors want to know what the student has learned by collecting the

Imhotep Proc.
notes, making connections, and applying new knowledge to a different and often times, a harder situation. By studying the same random facts without drawing on the topics is not what professors want to hear, because, quite frankly, it’s tedious. Professors know the facts and instead, what you, as the student, to show off your knowledge mussels. Splitting up ideas and memorizing facts is good to know but will serve no benefit for the student in the larger scheme of things. As Kiewra points out: "There are two types of associations you can create when learning: internal and external. Internal associations are those made within the material being learned. External associations are made between the new material being learned and past knowledge already stored in memory." If one can make connections to things already learned, one is already well on their way to making connections. This is where real learning takes place.

An excellent technique to note taking are the four BEs - BE there, BE on time, BE up front, and BE on edge - before lectures (Kiewra, 2009). This is even more the case in a mathematics course. Seeing a professor work out a problem is beneficial at first, but working out problems with the professor and fellow classmates is critical. In other words, mathematics is a not a subject one can lean back in your chair and only listen. One must be active and participate in the discussion. If a student misses class, they are missing out on the human interaction they would receive. In mathematics, a student may miss an important problem solving strategy that may not be mentioned in the text. For example, if a student is learning how to long-divide polynomials and is asked to divide \( x^3 - 1 \) by \( x - 1 \). Here, the quotient turns out to be \( x^2 + x + 1 \). However, the student may wonder how the author of the text obtained \( x^2 \) in the solution. This is where going to class is vital. Of course, the key idea is to ask how many times does \( x \) go into \( x^3 \) or \( \frac{1}{x} = x^2 \). All professors should stress to their students to come to each class. If students are not in class, they miss out on lecture and a chance to ask questions that motivate dialog between the professor and fellow classmates. The price of missing class is the opportunity for human interaction, a skill needed for school and everyday life. We as professors need to be pro-active with tracking student’s progress, especially freshmen in college. Being on time to review previous class notes and reading is important for the students to have a habit of doing on a daily basis. Inform students that when they are late, they miss important information, which maybe questions from other students in the class on previous materials, overview of last class with a transition of the next class by the professor, and any announcements. If possible, don’t allow students to sit in the back rows of the class as they are most likely to space out and not pay attention. Lastly, being on the edge of your seat and being physically ready for the lecture is important for successful note taking. Students will not have any benefit for sitting back in their seats, whether up front or in the back of the class, listening to you talk. Emphasize to the student to have their minds open to discussion and to activity participate. Having the concentration during the lecture requires students to be fully rested and have no distractions.

During the lecture, students should follow the four GETS - GET it all, GET it fast, GET is now, and GET it again (Kiewra, 2009). Many students only write down main ideas of the lecture thinking they will get the details later. This is never the case. In mathematics, when students are given mathematical definitions, it is not enough to simply write down the definition. The definition must be understood followed by simple examples in the notebook for quick reference (Dawkins, 2006). For example, for the definition of a logarithmic function, the student should note that \( y = \log_b x \) and \( x = b^y \) are equivalent and the first is in logarithmic form and the second in exponential form. Next, short examples are needed for quick memory recall as in \( \log_2 8 = 3 \) since \( 2^3 = 8 \). When taking notes, encourage students to obtain all the notes needed with examples for use on exams. Another effective skill is to use abbreviations for memory recall. Encourage students to ask you, the professor, to repeat parts of the lecture missed. After the lecture, students can complete their notes with filling in any missing information. This can be done by re-writing notes, looking up terms in the textbook, going over the lecture with the professor or replaying the lecture through a tape recorder. Stress to students to check their solutions. For example, when solving \( \sqrt{2x} = -10 \), once the solution of \( x = 50 \) is reached, students assume this is the final answer. But further investigation through check the solution back into the original equation leads to \( 2 \cdot 50 = \sqrt{100} = 10 \neq -10 \). In short, motivate students not to simply sit back and listen to you lecture. This is what they were taught to do in high school as all the notes were written on the board for them. Encourage them to be proactive and take charge of their education.

Highlighting has proved to be a non-effective study skill for millennial students. There are three problems with highlighting: 1. Students tend to overly highlight. 2. Highlighting alone is not effective as they do not remember key ideas better than students who read the same text and don’t highlight, and 3. Students who highlighted don’t recall non-highlighted areas of the text (Kiewra, 2009). Students need to be taught they must paraphrase their notes and not simply copy the lecture or readings word for word. In addition, right after a lecture, students should be spending time rewriting their notes and making connections across topics. The sooner the rewriting of notes and making connections takes place, the better the chances of student retaining the information in the long run occurs. For students who are visual learners, the idea of concept maps or Cmaps, proves to be an effective study skill. Here, the student creates a map of various terms and how they are related. Connections can be drawn and written up for the notebook. Details can be organized, which in turn, will make notes have a constant flow of ideas.
A major concern with millennial students is cramming for exams and other assessments. Cramming is when students wait until the last possible day or evening, to study for an exam or write a research paper. While some students can get away with cramming the night before an exam, there are no lasting benefits as the information a student crammed the night before will be forgotten for the next exam. A research paper takes weeks and even months to write properly. In most cases, some college courses have students write term papers, which are meant to be written over the course of the semester. They are designed to encourage students to draw option topics covered during the whole semester and make conclusions. If a student crams the paper in the last possible evening, they are tossing out any meaningful knowledge they may have been learned during the semester.

Taking a step back after re-writing notes, students should ask themselves, what could the professor ask as an assessment question? Motivate students to create their own test questions and have them answer them. This can be done in study groups, individually, or with the professor. Students should not be afraid to talk with their professors about how to go about studying or having a review session before an exam. Students need encouragement to take charge of their education and not to let education pass them by. This is their time to be in the spotlight and to prove to their professors, and, more importantly, to themselves, they can master any subject. Ask students to make a weekly schedule of their activities. By making a weekly schedule, students should be able to see how much time they spend on studying, being in the classroom, and time for pleasure. By seeing times spent, it should be a greater motivation to cut out certain activities, which take away for their studies. If handy, show students an example of several students’ schedules for past years of different grade levels. The example of the A student schedule should involve more hours of study than a D level student.

In recent years, multitasking has been a topic of conversation. Due to our fast-paced society, we are all receiving information at a rapid rate from various sources. Millennials have grown up with this speed of information. Today’s student may look at one screen, which maybe the computer for email (which may not be fast enough for them these days), then turn to their cell phone for texts, then facebook, then their Ipod, then Ipad, the list goes on. In short, they are wired to this rapid and quick paced environment. They believe they can perform multitasking however most aren’t good at it. Our brains aren’t designed to handle multiple tasks at once. Often times, when we attempt to multitask, we make mistakes. When we believe we are multitasking, the smallest of disruptions can cause us to lose our attention. Even when we believe we are note multitasking, we most likely are. For example, if you are sitting down studying for an exam and listening to music, you are multitasking. The reason being your brain is taking in information about your exam and taking in the musical notes and lyrics of the song being played. You are multitasking and don’t even know you are multitasking. During the schooling years and while in college, millennial students are faced with many more distractions than previous generations had been faced with. They rely on the Internet to get information fast, don’t write letters but emails instead, don’t use the phone but text instead, and play video games instead of being outdoors in the fresh air. Sitting in front of monitors and screen for long periods of time is the way of life for many of us these days. However, for the millennials, born in the late 1990s onward, they have been exposed to this from childhood.

The Internet moves quickly. It gives us answers at the blink of an eye. Millennial students expect, like the Internet, to get answers fast and do as little work as possible to achieve their goals. However, learning takes time and patience. Writing papers takes several drafts to perfect. Education and learning should be a pleasurable process, which should never be rushed because there are deadlines to meet or because one wants to finish something early to go out and party. This is how details are overlooked and one might miss on something important. Students’ ability to focus and fight through academic challenges is suffering an exponential decline. With all these new technologies, are we as teachers, supposed to alter our lessons? Are we to design lessons using phones, iPads, and video games just to keep student’s attention? Some teachers are incorporating these technologies in their classes. However, is this beneficial in the long run? Does this further isolate the student? Are the critical thinking skills harmed by this technology incorporated in the classroom? Most importantly, are we setting our expectations as high as we did before all the new modern technology came to the forefront? We believe students are as smart as they were before the millennial era. However, with the rapid advancement of technology, as good and beneficial as it has been, also has contributed and hurt students’ attention spans.

To engage the impatient generation Y, something more is needed. More inclusive assessment and evaluation criteria that could reduce the urge to cheat in examinations are needed. Children need to be encouraged to use the skills they had developed to do more independent learning. They need to learn better. They have to move from dependent learning to independent learning so that they are prepared for a society and an economy in which they will be expected to be self-directed learners, able and motivated to keep learning over a lifetime. It would be necessary to have an education system in which students are taught effective study skills that can help them deep study, get good grades, be creative, and innovative if the current trend of events in Kenya, Sub-Saharan Africa and the United States of America is to change. The Mathematics and Science curricula should be drafted around producing skilled individuals in the ICT; technical and vocational field to produce people who are more confident and employable.

Imhotep Proc.
VIII. What is needed of mathematics educators?

Mathematics educators need to understand the millennials and work in collaboration with them using a variety of instructional delivery methods to engage them. As a result of being used to being hand held in education, millennials need assistance in developing independent thinking and decision making skills. Effective mathematics study skills therefore become paramount. Teachers should use styles of communication that effectively reaches the students. Teaching skills that work well involve methods that develop millennial characteristics. Problem solving and cooperative learning are among some of the methods that could work well. General tips for studying mathematics, taking notes, getting help, homework, problem solving, studying for exams and taking a mathematics exam need to be taught to the students. Like all students, millennials learn more effectively when taught in accordance with their learning style preferences and when their worldviews are acknowledged. Prominent among their preferences are visual and kinaesthetic learning styles. Our Mathematics and Science teaching methods should thus incorporate the characteristics of this group of students if 21st century learning and globalisation is to take place.

IX. The way forward

The millennial student is no longer the type of learner our education system was designed to teach. As Monroe puts it: "It now costs more to amuse a child than it once did to educate his father" (JohnD.Cook@JohnDCook). The student constructs their knowledge and an educator should thus guide them with evidence-based practices in how to search for more knowledge. Uses of collaborative learning styles are encouraged. Thus, the educator is no longer using text recitation and lectures as modes of delivery but must provide an arena for engagement and discovery as well as being content expert and mentor. The job of the teacher is immensely harder than it was ten years ago. Designing relevant lessons that will inspire this generation of student is critical for attention to be on studies rather than on little screens. Unlike the common practice in Kenya and the United States, where most mathematics lessons consist of a review of the previous lesson, which can involve a warm-up problem, directed teaching on new material and then giving group work for students to try is simply not enough to spark enough attention of generation Y students (Owiti, 2012). Lessons must be with the times of the 21st century and must use technology and other means to motivate student learning. For example, applications involving financial mathematics should use Microsoft Excel where needed to give students a solid understanding of both the mathematics and the technology used in the field of finance. We should also encourage students to learn as much as they can on their own through self-discovery. Some American Universities are already adopting innovative practices through Mathematics Emporiums (Virginia Tech) and Mathematics MALLS (University of Central Florida). In conclusion, both Kenya and the United States, while being two totally different countries with different economic, social, financial, and demographic problems, are faced with the same crisis of the millennial student. Both countries must do something in line with their cultures to help this generation. Changes in curriculum and our ways of teaching need to be addressed to help our students be ready for life, all its challenges it brings, and helps them be the best they can be. Above all, effective study skills must be taught to the millennials.

References


S. Peter and W. Gareth, *Developing Students, self- efficacy through mathematical study skills resources*, MSOR Connections, 10 (2), May - July (2010).


SACMEQ *The SACMEQ II project, a study of the conditions of schooling and the quality of education*, Harare (2010).


www.un.org, retrieved on 2011-10-31
www.weforum.org/pdf/AFCR09Rankingsinfull.pdf

http://www.nctm.org/resources/content.aspx?id=12620

http://www.nctm.org/resources/content.aspx?id=12628

http://www.nctm.org/resources/content.aspx?id=12638

http://www.nctm.org/resources/content.aspx?id=12636

http://www.nctm.org/resources/content.aspx?id=12622

http://www.nctm.org/resources/content.aspx?id=12624

http://www.nctm.org/resources/content.aspx?id=126234

[45] Has Student Achievement Increased Since No Child Left Behind?, Center of Education Policy


Peter T. Olszewski
M.S Penn State Erie, The Behrend College
e-mail: pto2@psu.edu

Dickson S. O. Owiti
Kisii University, Department of Science Education, P. O. Box 408(40200) , Kisii , Kenya.
e-mail: owitidick@yahoo.com
Groups of units of commutative completely primary finite rings

Chiteng’a J. Chikunji
Department of Basic Sciences,
Botswana College of Agriculture,
Private Bag 0027, Gaborone,
BOTSWANA.
jchikunj@bca.bw

Abstract

It is well known that the group of units of a completely primary finite ring $R$ of order $p^{nr}$ is a semi-direct product of a cyclic group of order $p^r - 1$ by a $p$-group of order $p^{(n-1)r}$. The structure of the $p$-subgroup is not completely determined. In this presentation, we investigate and determine the structure of the $p$-subgroup of the group of units of a commutative completely primary finite ring $R$ of order $p^{nr}$ with unique maximal ideal $J$ such that $J^3 = (0)$, $J^2 \neq (0)$, and with characteristic $p^2$, for any prime number $p$ and positive integers $n$ and $r$. 
Groups of units of commutative completely primary finite rings

Chiteng’a J. Chikunji

Abstract. It is well known that the group of units of a completely primary finite ring $R$ of order $p^{nr}$ is a semi-direct product of a cyclic group of order $p^r - 1$ by a $p$-group of order $p^{(n-1)r}$. The structure of the $p$-subgroup is not completely determined. In this presentation, we investigate and determine the structure of the $p$-subgroup of the group of units of a commutative completely primary finite ring $R$ of order $p^{nr}$ with unique maximal ideal $J$ such that $J^3 = (0)$, $J^2 \neq (0)$, and with characteristic $p^2$, for any prime number $p$ and positive integers $n$ and $r$.

Mathematics Subject Classification (2000). Primary 16P10, 13M05; Secondary 20K01, 20K25.

Keywords. Finite commutative rings, Galois rings, group of units, direct products, abelian $p$-groups.

I. Introduction

A finite ring $R$ is called completely primary if all its zero divisors including the zero element form the unique maximal ideal $J$. Finite completely primary rings are precisely local rings with unique maximal ideals.

All rings considered in this work are commutative with identity $1 \neq 0$ unless specified otherwise, that ring homomorphisms preserve identities, and that a ring and its subrings have the same identity. Moreover, we adopt the notation used in [1], [2] and [3], that is, $R$ will denote a finite ring, unless otherwise stated, $J$ will denote the Jacobson radical of $R$, and we will denote the Galois ring $GR(p^k, p^{kr})$ of characteristic $p^k$ and order $p^{kr}$ by $R_{p^k/r}$, for some prime integer $p$, and positive integers $k$, $r$. We denote the unit group of $R$ by $U(R)$; if $g$ is an element of $U(R)$, then $o(g)$ denotes its order, and $\langle g \rangle$ denotes the cyclic group generated by $g$. Similarly, if $f(x) \in R[x]$, we shall denote by $\langle f(x) \rangle$ the ideal generated by $f(x)$. Further, for a subset $A$ of $R$ or $U(R)$, $|A|$ will denote the number of elements in $A$. The ring of integers modulo the number $n$ will be denoted by $\mathbb{Z}_n$, and the characteristic of $R$ will be denoted by $\text{char } R$. We denote a direct product of $r$ cyclic groups $\mathbb{Z}_m$ by $\mathbb{Z}_m^r$ or by $\mathbb{Z}_m \times \ldots \times \mathbb{Z}_m$.

The rest of the paper is organized as follows. In Section 2, we state without proofs some general results on groups of units of completely primary finite rings which are relevant to our work. In section 3, we give an explicit description of the known structures of groups of units of certain completely primary finite rings $R$ of order $p^{nr}$ with maximal ideals $J$ such

Paper presented at the 2nd Strathmore International Mathematics Conference (SIMC 2013), 12 - 16 August 2013, Strathmore University, Nairobi, Kenya.
that $\mathcal{J}^3 = (0)$, $\mathcal{J}^2 \neq (0)$. Finally, in section 4, we determine the structure of the unit group $U(R)$ of $R$ and in some cases, its generators, when the characteristic of $R$ is $p^2$, $s \geq 3$ and $1 \leq \dim_{R_{t,s}}(\mathcal{J}^2) < (s+1)/2$, without considering structural matrices of isomorphic classes of these types of rings. This complements the author’s earlier solution of the problem in the case when the characteristic of $R$ is $p$, $s = 3$, $t = 1$ and $\mathcal{J}^2 \subseteq \text{ann}(\mathcal{J})$, the annihilator of $\mathcal{J}$.

II. Completely primary finite (CPF) rings

Let $R$ be a completely primary finite ring, $\mathcal{J}$ the set of all zero divisors in $R$, $p$ a prime, $k$, $n$ and $r$ be positive integers. Properties of completely primary finite (CPF) rings and those of their groups of units, with different aims and scope, appear in several articles (e.g. [6], [7]), and below we state some of the results without proofs ([6]): $|R| = p^n$, $\mathcal{J}$ is the Jacobson radical of $R$, $\mathcal{J}^n = (0)$, $|\mathcal{J}| = p^{(n-1)r}$, $R/\mathcal{J} \cong GF(p^r)$, the finite field of $p^r$ elements and $\text{char} R = p^k$, where $1 \leq k \leq n$. If $n = k$, it is known that, up to isomorphism, there is precisely one completely primary ring of order $p^k$ having characteristic $p^k$ and residue field $GF(p^r)$. It is called the Galois ring $GR(p^k, p^r)$ and a concrete model is the quotient $\mathbb{Z}_{p^n}/\langle f(x) \rangle$, where $f(x)$ is a monic polynomial of degree $r$, irreducible modulo $p$. Any such polynomial will do: the rings are all isomorphic. Trivial cases are $GR(p^k, p^r) = \mathbb{Z}_{p^n}$ and $GR(p^k, p) = \mathbb{F}_{p^n}$. In fact, $R = \mathbb{Z}_{p^n}[b]$, where $b$ is an element of $R$ of multiplicative order $p^r - 1$; $\mathcal{J} = pR$ and $\text{Aut}(R) \cong \text{Aut}(R/pR)$ (see Proposition 2 in [6]).

Let $R$ be a completely primary ring, $[R/\mathcal{J}] = p^r$ and $\text{char} R = p^k$. Then it can be deduced from [6] and [7] that $R$ has a coefficient subring $R_o$ of the form $GR(p^k, p^{kr})$ which is clearly a maximal Galois subring of $R$. Moreover, if $R_o$ is another coefficient subring of $R$ then there exists an invertible element $x$ in $R$ such that $R_o = xR_o x^{-1}$ (see Theorem 8 in [6]). Furthermore, there exist elements $m_1, ..., m_h \in \mathcal{J}$ and automorphisms $\sigma_1, ..., \sigma_h \in \text{Aut}(R_o)$ such that $R = R_o \oplus \sum_{i=1}^{h} R_o m_i$ (as $R_o$-modules), $m_i r_o = r_o^m m_i$, for all $r_o \in R_o$ and any $i = 1, ..., h$. Moreover, $\sigma_1, ..., \sigma_h$ are uniquely determined by $R$ and $R_o$. The maximal ideal of $R$ is $\mathcal{J} = pR_o \oplus \sum_{i=1}^{h} R_o m_i$. We call $\sigma_i$ the automorphism associated with $m_i$ and $\sigma_1, ..., \sigma_h$ the associated automorphisms of $R$ with respect to $R_o$.

Now, let $R_o = \mathbb{Z}_{p^n}[b]$ be a coefficient subring of $R$ of order $p^{kr}$ and characteristic $p^k$ and let $K_o = (b) \cup \{0\}$, denote the set of coset representatives of $\mathcal{J}$ in $R$. Then it is easy to show that every element of $R_o$ can be written uniquely as $\sum_{i=0}^{h-1} \lambda_i p^i$, where $\lambda_i \in K_o$.

Let $R$ be a completely primary finite ring (not necessarily commutative). The following facts are useful (e.g. see [6, §2]): The unit group $U(R)$ of $R$ contains a cyclic subgroup $\langle b \rangle$ of order $p^r - 1$ and a $p$-Sylow subgroup $1 + \mathcal{J}$ of order $p^{(n-1)r}$; hence $U(R)$ is a semi-direct product of $1 + \mathcal{J}$ by $\langle b \rangle$ and $|U(R)| = p^{(n-1)r}(p^r - 1)$; the unit group $U(R)$ is solvable; if $G$ is a subgroup of $U(R)$ of order $p^r - 1$, then $G$ is conjugate to $\langle b \rangle$ in $U(R)$; if $U(R)$ contains a normal subgroup of order $p^r - 1$, then the set $K_o = (b) \cup \{0\}$ is contained in the center of the ring $R$; and $(1 + \mathcal{J}^i)/(1 + \mathcal{J}^{i+1}) \cong \mathcal{J}^i/\mathcal{J}^{i+1}$ (the left hand side as a multiplicative group and the right hand side as an additive group).

III. Some known groups of units of CPF rings with $\mathcal{J}^3 = (0)$

Let $R$ be a commutative completely primary finite (CPF) ring with maximal ideal $\mathcal{J}$ such that $\mathcal{J}^3 = (0)$ and $\mathcal{J}^2 \neq (0)$. Then $\text{char} R = p^k$, where $1 \leq k \leq 3$ (see [1]). Let $s$, $t$, $\lambda$ be numbers in the generating sets for the $R_o$-$\text{modules}$ $U$, $V$, $W$, respectively, so that

$$R = R_o \oplus U \oplus V \oplus W$$

Imhotep Proc.
and
\[ J = pR_o \oplus U \oplus V \oplus W. \]

In [3] we have determined the group of units \( U(R) \) of the ring \( R \) and its generators when \( s = 2, t = 1, \lambda = 0 \) and characteristic of \( R \) is \( p \); and when \( t = s(s + 1)/2, \lambda = 0 \), for a fixed integer \( s \), for all the characteristics of \( R \). It was noted that \( U(R) \) and its generators depended on the structural matrices \( (a_{ij}) \) and on the parameters \( p, k, r, \) and \( s \). In [4] we obtained the structure of \( U(R) \) and its generators when \( s = 2, t = 1, \lambda = 0 \) and characteristic of \( R \) is \( p^2 \) and \( p^3 \); and the case when \( s = 2, t = 2, \lambda = 0 \) and characteristic of \( R \) is \( p \). In both papers, [3] and [4], we assumed that \( \lambda = 0 \) so that the annihilator of the maximal ideal \( J \) coincides with \( J^2 \). It was also noted that the earlier strategy (that of considering different types of symmetric matrices) was thus not viable anymore and we followed a different approach; that of considering structural matrices of isomorphic classes of these types of rings with the same invariants \( p, r, k, s, \) and \( t \).

In [5], we proved that \( 1 + J \) is a direct product of its subgroups \( 1 + pR_o \oplus U \oplus V \) and \( 1 + W \) and further determined the structure of \( 1 + W \); in general; we also determined the structure of \( U(R) \) and its generators when \( s = 3, t = 1, \lambda \geq 1 \) and \( \text{char} R = p \). We then generalized the structure of \( U(R) \) in the cases when \( s = 2, t = 1; t = s(s + 1)/2, \) for a fixed integer \( s \), and for all characteristics of \( R \); and when \( s = 2, t = 2 \) and \( \text{char} R = p \); determined in [3] and [4], to the case when \( \text{ann}(J) = J^2 + W \) so that \( \lambda \geq 1 \).

We state the following result which summarizes the structure of the group of units \( U(R) \) of the rings \( R \) determined in [3], [4] and [5].

**Theorem III.1.** The group of units \( U(R) \) of a commutative completely primary finite (CPF) ring \( R \) with maximal ideal \( J \) such that \( J^3 = (0) \) and \( J^2 \neq (0) \), and with the invariants \( p, k, r, s, t, \) and \( \lambda \geq 1 \), is a direct product of cyclic groups as follows:

i) If \( s = 2, t = 1, \lambda \geq 1 \) and \( \text{char} R = p \), then

\[
U(R) = \begin{cases} 
Z_{2^r-1} \times Z_{2^s} \times Z_{2^r} \times (Z_{2^s})^\lambda & \text{or} \\
Z_{2^r-1} \times Z_{2^s} \times Z_{2^r} \times (Z_{2^s})^\lambda & \text{if } p = 2 \\
Z_{2^r-1} \times Z_{2^s} \times Z_{2^r} \times (Z_{2^s})^\lambda & \text{if } p \neq 2
\end{cases}
\]

and if \( p = 2 \)

\[
U(R) = \begin{cases}
(Z_2 \times Z_2) \times (Z_2 \times Z_2) \times Z_2 \times (Z_2)^\lambda & \text{if } r = 1 \text{ and } 2 \in J - \text{ann}(J); \\
Z_{2^r-1} \times Z_{2^s} \times Z_{2^r} \times (Z_{2^s})^\lambda & \text{if } r > 1 \text{ and } 2 \in J - \text{ann}(J); \\
Z_{2^r-1} \times Z_{2^s} \times Z_{2^r} \times (Z_{2^s})^\lambda & \text{if } 2 \in J^2; \\
Z_{2^r-1} \times Z_{2^s} \times Z_{2^r} \times (Z_{2^s})^\lambda & \text{if } 2 \in \text{ann}(J) - J^2;
\end{cases}
\]

ii) If \( s = 2, t = 1, \lambda \geq 1 \) and \( \text{char} R = p^2 \), then

\[
U(R) = \begin{cases} 
Z_{p^r-1} \times Z_{p^s} \times Z_{p^r} \times Z_{p^s} \times (Z_{p^r})^\lambda & \text{or} \\
Z_{p^r-1} \times Z_{p^s} \times Z_{p^r} \times Z_{p^s} \times (Z_{p^r})^\lambda & \text{if } p \neq 2
\end{cases}
\]

and

iii) If \( s = 2, t = 1, \lambda \geq 1 \) and \( \text{char} R = p^3 \), then

\[
U(R) = \begin{cases} 
Z_{p^r-1} \times Z_{p^s} \times Z_{p^r} \times Z_{p^s} \times Z_{p^r} \times (Z_{p^r})^\lambda & \text{or} \\
Z_{p^r-1} \times Z_{p^s} \times Z_{p^r} \times Z_{p^s} \times Z_{p^r} \times (Z_{p^r})^\lambda & \text{if } p \neq 2
\end{cases}
\]

and

\[
U(R) = \begin{cases} 
Z_{2^r-1} \times Z_{2^s} \times Z_{2^r} \times (Z_{2^s})^\lambda & \text{or} \\
Z_{2^r-1} \times Z_{2^s} \times Z_{2^r} \times (Z_{2^s})^\lambda & \text{if } p = 2
\end{cases}
\]
iv) If \( s = 2, \ t = 2, \ \lambda \geq 1 \) and \( \text{char}R = p \), then
\[
U(R) = \begin{cases} 
Z_{p^2-1} \times Z_p^s \times Z_p^t \times Z_p^r \times (Z_p^r)^\lambda & \text{if } p \neq 2, \\
Z_{2^r-1} \times Z_p^s \times Z_p^t \times (Z_p^r)^\lambda & \text{if } p \neq 2,
\end{cases}
\]

v) If \( t = (s+1)/2, \ \lambda \geq 1, \) and
(a) \( \text{char}R = p \), then
\[
U(R) = \begin{cases} 
Z_{2^r-1} \times Z_p^s \times (Z_p^s)^\gamma \times (Z_p^t)^\lambda & \text{if } p \neq 2, \\
Z_{p^2-1} \times (Z_p^s)^\gamma \times (Z_p^t)^\lambda & \text{if } p \neq 2,
\end{cases}
\]

(b) \( \text{char}R = p^2 \), then
\[
U(R) = \begin{cases} 
Z_{2^r-1} \times Z_p^s \times Z_p^t \times Z_p^r \times (Z_p^r)^\gamma \times (Z_p^t)^\lambda & \text{if } p \neq 2, \\
Z_{p^4-1} \times (Z_p^s)^\gamma \times (Z_p^t)^\lambda & \text{if } p \neq 2;
\end{cases}
\]

c) \( \text{char}R = p^3 \), then
\[
U(R) = \begin{cases} 
Z_{2^r-1} \times Z_p^s \times Z_p^t \times Z_p^r \times (Z_p^r)^\gamma \times (Z_p^t)^\lambda & \text{if } p \neq 2, \\
Z_{p^4-1} \times (Z_p^s)^\gamma \times (Z_p^t)^\lambda & \text{if } p \neq 2;
\end{cases}
\]

where \( \gamma = (s^2 - s)/2 \).

The proof follows from Section 3.1 in [3], Propositions 2.2, 2.3, 2.4 and 2.5 in [4], Theorem 4.1 in [3] and Proposition 2.3 in [5].

The above results describe structures of groups of units of completely primary finite rings when \( m = 2 \), that is, when \( J^2 = (0) \); when \( m = k = n \), that is, \( J^m = (0) \) and \( \text{char}R = p^m \); and when \( m = 3 \), that is, \( J^3 = (0) \), for given parameters. The solution for different parameters when \( m = 3 \) and \( m \geq 4 \) is left for further consideration.

In section 4 we extend the above problem to the case when the characteristic of \( R \) is \( p^2, \ s \geq 3 \) and \( 1 \leq \dim_{R_o/pR_o}(J^2) < s(s+1)/2 \), without considering structural matrices of isomorphic classes of these types of rings.

**IV. Group of units of CPF rings of characteristic \( p^2 \)**

We now consider the structure of the group of units of completely primary finite rings with maximal ideals \( J \) such that \( J^3 = (0), J^2 \neq (0) \), and with characteristic \( p^2 \).

**IV.1. A construction of commutative rings of characteristic \( p^2 \).**

Let \( R_o \) be the Galois ring \( GR(p^2, p^{2r}) \). Let \( s, d, t \) be integers with either \( 1 \leq 1 + t \leq s(s+1)/2 \)
or \( 1 \leq d + t \leq s(s+1)/2 \). Let \( V \) be an \( R_o/pR_o \)-space, which when considered as an \( R_o \)-module, has a generating set \( \{v_1, \ldots, v_l\} \) and let \( U \) be an \( R_o \)-module with an \( R_o \)-module generating set \( \{u_1, \ldots, u_s\} \); and suppose that \( d \geq 0 \) of the \( u_i \) are such that \( pu_i \neq 0 \). Since \( R_o \) is commutative, we can think of them as both left and right \( R_o \)-modules.

Let \( \{a_{ij}^d\} \) be \( 1 + t \) or \( t + d, s \times s \) linearly independent symmetric matrices over \( R_o/pR_o \).

On the additive group \( R = R_o \oplus U \oplus V \) we define multiplication by the following relations:
\[
\begin{align*}
u_iu_j &= a_{ij}^d p + \sum_{l=1}^d a_{ij}^l p u_l + \sum_{k=1}^t a_{ij}^{d+t} v_k; \\
u_iu_k &= v_ku_i = u_iu_jv_\lambda = p v_k = v_ku_i = v_iu_k = 0; \tag{IV.1}
\end{align*}
\]

where \( \alpha, a_{ij}^d, a_{ij}^l, a_{ij}^{d+t} \in R_o/pR_o \).
By the above relations, \( R \) is a commutative completely primary finite ring of characteristic \( p^2 \) with Jacobson radical \( \mathcal{J} = \mathcal{P}_o \oplus U \oplus V \), \( \mathcal{J}^2 = \mathcal{P}_o \oplus \mathcal{V} \) or \( \mathcal{J}^2 = pU \oplus V \) and \( \mathcal{J}^3 = (0) \). We call \((a_{ij}^i)\) the structural matrices of the ring \( R \) and the numbers \( p, n, r, s, d \) and \( t \) invariants of the ring \( R \).

The following result is proved in [1, Theorem 6.1].

**Theorem IV.1.** Let \( R \) be a ring. Then \( R \) is a commutative completely primary finite ring of characteristic \( p^2 \) with maximal ideal \( \mathcal{J} \) such that \( \mathcal{J}^3 = (0), \mathcal{J}^2 \neq (0) \), the annihilator of \( \mathcal{J} \) coincides with \( \mathcal{J}^2 \) if and only if \( R \) is isomorphic to one of the rings given by the relations in (IV.1).

**Remark IV.2.** We know that \( R = \mathcal{P}_o \oplus \mathcal{P}_o m_1 \oplus \cdots \oplus \mathcal{P}_o m_h \), where the elements \( m_i \in \mathcal{J} \); and that \( \mathcal{J} = \mathcal{P}_o \oplus \mathcal{P}_o m_1 \oplus \cdots \oplus \mathcal{P}_o m_h \). Since \( \mathcal{J}^3 = (0) \) and \( \mathcal{J}^2 = \text{ann}(\mathcal{J}) \), with \( \mathcal{J}^2 \neq (0) \), we can write

\[
\{m_1, ..., m_h\} = \{u_1, ..., u_s, v_1, ..., v_t\}
\]

where, \( u_1, ..., u_s \in \mathcal{J} = \mathcal{J}^2 \) and \( v_1, ..., v_t \in \mathcal{J}^2 \), so that \( s + t = h \).

In view of the above considerations and by 1.8 of [1], the non-zero elements of

\[
\{1, p, u_1, ..., u_s, pv_1, ..., puv_s, v_1, ..., v_t\}
\]

form a “basis” for \( R \) over \( K_o = \mathcal{P}_o / \mathcal{P}_o R \).

Since \( pm = 0 \), for all \( m \in \mathcal{J}^2 \), it is easy to check that if \( \text{char} \ R = p^2 \), then either

(i) \( p \in \mathcal{J}^2 \), or

(ii) \( p \notin \mathcal{J} \).  

These two cases do not overlap, and for clarity of our work, we consider them in turn.

**Remark IV.3.** Suppose that \( \text{char} R = p^2 \) and \( p \) lies in \( \mathcal{J}^2 \). In this case, (IV.2) becomes

\[
\{1, p, u_1, ..., u_s, v_1, ..., v_t\};
\]

and by Proposition 3.2 of [1], \( 1 \leq 1 + t \leq s(s + 1)/2 \). Hence, every element of \( R \) may be written uniquely as

\[
a_o + a_1 p + \sum_{i=1}^s b_i u_i + \sum_{k=1}^t c_k v_k; \quad a_o, a_1, b_i, c_k \in K_o;
\]

and therefore,

\[
u_i u_j = a_{ij}^o p + \sum_{k=1}^t a_{ij}^k v_k,
\]

where \( a_{ij}^o, a_{ij}^k \in \mathcal{P}_o / \mathcal{P}_o R \). Clearly, \( \dim_{\mathcal{P}_o / \mathcal{P}_o R}(\mathcal{J}^2) = 1 + t \).

**Remark IV.4.** Suppose that \( d \geq 0 \) is the number of the elements \( pu_i \) in (IV.2) which are not zero. Suppose, without loss of generality, that \( pu_1, ..., pu_d \) are the \( d \) non-zero elements. Then, (IV.2) becomes

\[
\{1, p, u_1, ..., u_s, pu_1, ..., pu_d, v_1, ..., v_t\};
\]

and by Proposition 3.2 of [1], we have \( 1 \leq d + t \leq s(s + 1)/2 \) and hence, every element of \( R \) may be written uniquely as

\[
\lambda_o + \lambda_1 p + \sum_{i=1}^s \alpha_i u_i + \sum_{l=1}^d \beta_l pu_l + \sum_{k=1}^t \gamma_k v_k; \quad \lambda_o, \lambda_1, \alpha_i, \beta_l, \gamma_k \in K_o.
\]

Clearly, the products \( u_i u_j \in \mathcal{J}^2 \). Hence,

\[
u_i u_j = \sum_{l=1}^d a_{ij}^l pu_l + \sum_{k=1}^t a_{ij}^{k+d} v_k, \quad \text{where} \quad a_{ij}^l, a_{ij}^{k+d} \in \mathcal{P}_o / \mathcal{P}_o R.
\]

Imhotep Proc.
and \( \dim_{R_o/pR_o}(\mathcal{J}^2) = d + t \).

Now, since \( pu_i, v_k \in \mathcal{J}^2 \) (\( l = 1, \ldots, d; k = 1, \ldots, t \)), we can write them as sums of products of elements of \( \mathcal{J} \). In particular, \( pu_i, v_k \) can be written as linear combinations of \( pu_i \) and \( u_iu_j \) with coefficients in \( R_o/pR_o \). Hence, since \( pu_i, v_k \) (\( l = 1, \ldots, d; k = 1, \ldots, t \)) is a basis for \( \mathcal{J}^2 \) over \( R_o/pR_o \), we conclude that \( pu_i \) and \( u_iu_j \) (\( i, j = 1, \ldots, s \)) generate \( \mathcal{J}^2 \).

Clearly, \(|R| = p^{2r} \cdot p^{sr} \cdot p^r = p^{(2+s+d+t)r} \) and \(|\mathcal{J}| = p^{(1+s+d+t)r} \).

**IV.2. The group of units.**

Notice that since \( R \) is of order \( p^{kr} \) and \( U(R) = R - \mathcal{J} \), it is easy to see that \(|U(R)| = p^{(k-1)r} (p^r - 1)\) and \(|1 + \mathcal{J}| = p^{(k-1)r} \), so that \( 1 + \mathcal{J} \) is an abelian \( p \)-group. Thus, since \( R \) is commutative,

\[
U(R) = \langle b \rangle \cdot (1 + \mathcal{J}) \cong \langle b \rangle \times (1 + \mathcal{J});
\]

a direct product of the \( p \)-group \( 1 + \mathcal{J} \) by the cyclic subgroup \( \langle b \rangle \). Thus, it suffices to determine the structure of the subgroup \( 1 + \mathcal{J} \) of the group \( U(R) \).

Notice that

\[
1 + \mathcal{J} = 1 + pR_o \oplus \sum_{i=1}^{s} R_o u_i \oplus \sum_{k=1}^{t} R_o v_k.
\]

**Proposition IV.5.** ([3], Proposition 3.4) If \( \text{char} R = p^2 \), then \( 1 + \mathcal{J} \) contains \( 1 + pR_o \) as its subgroup.

The structure of \( 1 + pR_o \) is completely determined by Raghavendran in [6]. For convenience of the reader, we state here the following result. For details, refer to [6, Theorem 9].

We take \( r \) elements \( \varepsilon_1, \ldots, \varepsilon_r \) in \( R_o \) with \( \varepsilon_1 = 1 \) such that \( \{ \overline{\varepsilon_1}, \ldots, \overline{\varepsilon_r} \} \) is a basis of the quotient ring \( R_o/pR_o \) regarded as a vector space over its prime subfield \( GF(p) \). Then we have the following.

**Proposition IV.6.** ([6, Theorem 9]) If \( \text{char} R_o = p^2 \), then \( 1 + pR_o \) is a direct product of \( r \) cyclic groups \( \langle \varepsilon_j \rangle \) (\( j = 1, \ldots, r \)), each of order \( p \) for any prime number \( p \).

**IV.2.1. Group of units of rings of characteristic \( p^2 \) in which \( p \in \mathcal{J}^2 \).** Let \( R \) be a commutative completely primary finite ring of characteristic \( p^2 \) in which \( p \in \mathcal{J}^2 \). Then \( \dim_{R_o/pR_o}(\mathcal{J}^2) = 1+t \). The following results determine the structure of the subgroup \( 1 + \mathcal{J} \) of the group of units \( U(R) \) of the ring \( R \).

**Proposition IV.7.** Let \( \text{char} R = p^2 \), \( s \geq 3, 1 + t < s(s + 1)/2 \), and suppose that \( p \in \mathcal{J}^2 \). If \( p \) is odd, then

\[
1 + \mathcal{J} \cong Z_{p^2} \times \cdots \times Z_{p^2}\]

a direct product of \( (1 + s + t)r \) cyclic groups of order \( p \).

**Proof.** If \( p \in \mathcal{J}^2 \), let \( a = 1 + x \) be an element of \( 1 + \mathcal{J} \) with the highest possible order and assume that \( x \in \mathcal{J} - \mathcal{J}^2 \). Then \( o(a) = p \). This is true because

\[
(1 + x)^p = 1 + px + \frac{p(p-1)}{2} x^2 \quad \text{(since } x^3 = 0) \]

\[
= 1 + \frac{p(p-1)}{2} x^2 \quad \text{(since } p \in \mathcal{J}^2 \text{ and } px = 0) \]

\[
= 1 \quad \text{(since } p - 1 \text{ is even and } px^2 = 0). \]

Now let \( \varepsilon_1, \ldots, \varepsilon_r \in R_o \) with \( \varepsilon_1 = 1 \) such that \( \varepsilon_1, \ldots, \varepsilon_r \in R_o/pR_o \cong GF(p^r) \) form a basis for \( GF(p^r) \) over its prime subfield \( GF(p) \), for any prime \( p \) and positive integer \( r \).

Imhotep Proc.
We first note the following results: For each $i = 1, \ldots, r$; $j = 1, \ldots, s$; and $k = 1, \ldots, t$, $(1 + \varepsilon_ip)^p = 1$, $(1 + \varepsilon_iu_j)^p = 1$, $(1 + \varepsilon_iv_k)^p = 1$, and $g^p = 1$ for all $g \in 1 + \mathcal{J}$. For integers $l_i$, $m_i$, $n_i \leq p$, we assert that

$$
\prod_{i=1}^{r}(1 + \varepsilon_ip)^{l_i} \cdot \prod_{j=1}^{s}(1 + \varepsilon_iu_j)^{m_i} \cdot \prod_{k=1}^{t}(1 + \varepsilon_iv_k)^{n_i} = 1,
$$

will imply $l_i$, $m_i$, $n_i = p$, for all $i = 1, \ldots, r$.

If we set $D_i = \{(1 + \varepsilon_ip)^l : l = 1, \ldots, p\}$, $E_{i,j} = \{(1 + \varepsilon_iu_j)^m : m = 1, \ldots, p\}$, $(j = 1, \ldots, s)$; and $F_{i,k} = \{(1 + \varepsilon_iv_k)^n : n = 1, \ldots, p\}$, $(k = 1, \ldots, t)$, for all $i = 1, \ldots, r$; we see that $D_i$, $E_{i,j}$, $F_{i,k}$ are all subgroups of the group $1 + \mathcal{J}$ and these are all of order $p$ as indicated in their definition. The argument above will show that the product of the $(1 + s + t)r$ subgroups $D_i$, $E_{i,j}$, $F_{i,k}$ is direct. So, their product will exhaust $1 + \mathcal{J}$. This completes the proof.

\[\text{Proposition IV.8.}\]

Let $\text{char } R = p^2$, $s \geq 3$, $1 + t < s(s+1)/2$, and suppose that $p \in \mathbb{J}^2$. If $p = 2$ and $u_j^2 = 0$, for every $j = 1, \ldots, s$; then

$$
1 + \mathcal{J} \cong \bigoplus_{i=1}^{1+s+t} \mathbb{Z}_2^r \times \cdots \times \mathbb{Z}_2^r,
$$

a direct product of $(1 + s + t)r$ cyclic groups of order 2.

**Proof.** If $u_j^2 = 0$, for every $j = 1, \ldots, s$; then the highest possible order of any element in $1 + \mathcal{J}$ is 2. The proof follows a similar argument to that of the case when $p$ is odd, and may be deduced from the previous proposition.

\[\text{Proposition IV.9.}\]

Let $\text{char } R = p^2$, $s \geq 3$, $1 + t < s(s+1)/2$, and suppose that $p \in \mathbb{J}^2$. If $p = 2$ and suppose that $l \leq s$ is the number of the $u_j$ such that $u_j^2 \neq 0$. Then

$$
1 + \mathcal{J} \cong \bigoplus_{i=1}^{l} \mathbb{Z}_2^r \times \cdots \bigoplus_{i=1}^{m} \mathbb{Z}_2^r \times \cdots \bigoplus_{i=1}^{t} \mathbb{Z}_2^r,
$$

a direct product of $lr$ cyclic groups of order 4 and $mr$ cyclic groups of order 2, where $l + m = 1 + s + t$.

**Proof.** We first observe that if, without lose of generality, $u_j^2 \neq 0$ while $u_j^2 = 0$, for every $j = 2, \ldots, s$, then $(1 + \varepsilon_iu_j)^4 = 1$ and the elements $1 + \varepsilon_iu_j$, $1 + \varepsilon_iv_k$, and $1 + \varepsilon_ip$ are all of order 2; and if $u_j^2 \neq 0$, $u_j^2 \neq 0$ while $u_j^2 = 0$ ($j = 3, \ldots, s$), then $(1 + \varepsilon_iu_j)^4 = 1$, $(1 + \varepsilon_iu_j)^4 = 1$ and the elements $1 + \varepsilon_iu_j$, $1 + \varepsilon_iv_k$, and $1 + \varepsilon_ip$, are all of order 2. Continuing the argument so that every $u_j$ has a non-zero square, we see that $(1 + \varepsilon_iu_j)^4 = 1$ and the elements $1 + \varepsilon_iv_k$ and $1 + \varepsilon_ip$, are all of order 2.

Since $p$, $v_k$ are linear combinations of $u_ju_j$ with coefficients in $\mathbb{R}_n/p\mathbb{R}_n$, the products of elements $1 + \varepsilon_iu_j$ generate the elements $1 + \varepsilon_iv_k$ and $1 + \varepsilon_ip$. By induction, we obtain the desired result.

As an example to illustrate Proposition IV.9, suppose that $s = 4$ and $t = 2$. Then $|1 + \mathcal{J}| = p^{(4+2+1)r}$ and

\[
1 + \mathcal{J} = \left\{ \begin{array}{l}
\mathbb{Z}_2^4 \times \mathbb{Z}_2^4 \times \mathbb{Z}_2^4 \times \mathbb{Z}_2^4 \times \mathbb{Z}_2^4 ; \\
\mathbb{Z}_2^r \times \mathbb{Z}_2^r \times \mathbb{Z}_2^r \times \mathbb{Z}_2^r ; \\
\mathbb{Z}_2^4 \times \mathbb{Z}_2^4 \times \mathbb{Z}_2^4 \times \mathbb{Z}_2^4 ; \text{ or}
\end{array} \right.
\]

Imhotep Proc.
IV.2.2. Group of units of rings of characteristic $p^2$ in which $p \in \mathcal{J} - \mathcal{J}^2$. In this case, 
$\dim_{R_0/pR_0}(\mathcal{J}^2) = d + t$. The following result determines the structure of $1 + \mathcal{J}$ in which $d = s$.

**Proposition IV.10.** Let $\text{char} R = p^2$, $s \geq 3$, $d + t < s(s + 1)/2$, and suppose that $p \in \mathcal{J} - \mathcal{J}^2$. Suppose further that $pu_i \neq 0$, for every $i = 1, \ldots, s$. Then

$$1 + \mathcal{J} \cong \bigoplus_{i=1}^{s} \mathbb{Z}_{p}^{2} \times \mathbb{Z}_{p^{3}}^{2} \times \cdots \times \mathbb{Z}_{p^{r}}^{2} \times \cdots \times \mathbb{Z}_{p^{r}}^{2}$$

**Proof.** If $p \in \mathcal{J} - \mathcal{J}^2$, let $a = 1 + x$ be an element of $1 + \mathcal{J}$ with the highest possible order and assume that $x \in \mathcal{J} - \mathcal{J}^2$. Then $\alpha(a) = p^2$, for any prime $p$.

This is true because, for any $\varepsilon_i$ ($i = 1, \ldots, r$),

$$(1 + \varepsilon_i x)^p = 1 + p\varepsilon_i x + \frac{p(p - 1)}{2} (\varepsilon_i x)^2 \text{ (since } x^3 = 0).$$

If $p$ is odd, then $(1 + \varepsilon_i x)^p = 1 + p\varepsilon_i x, \text{ since } px^2 = 0$. Now

$$(1 + p\varepsilon_i x)^p = 1 + p^2\varepsilon_i x + \frac{p(p - 1)}{2} (p\varepsilon_i x)^2 = 1, \text{ since } \text{char} R = p^2.$$ 

Hence, $(1 + \varepsilon_i x)^p = 1$. If $p$ is even, then

$$(1 + \varepsilon_i x)^2 = 1 + 2\varepsilon_i x + (\varepsilon_i x)^2, \text{ and } (1 + \varepsilon_i x)^4 = 1.$$ 

Notice that

$$1 + \mathcal{J} = (1 + pR_0) \times (1 + \sum_{i=1}^{s} R_0 u_i + \sum_{k=1}^{t} R_0 v_k).$$

The structure of the group $1 + pR_0$ is well known; and it is a direct product of $r$ cyclic groups, each of order $p$ (see Proposition IV.6).

We now determine the structure of $1 + \sum_{i=1}^{s} R_0 u_i + \sum_{k=1}^{t} R_0 v_k$. Choose $\varepsilon_1, \ldots, \varepsilon_r \in R_0$ with $\varepsilon_1 \neq 0$ such that $\varepsilon_1, \ldots, \varepsilon_r \in R_0$ with $\varepsilon_1 = 1$ such that $\tau_i, \ldots, \tau_r \in R_0/pR_0 \cong GF(p^r)$ form a basis for $GF(p^r)$ over $GF(p)$.

For any prime $p$, since for each $i = 1, \ldots, r$, we have that $(1 + \varepsilon_i u_j)^p = 1, (j = 1, \ldots, s) (1 + \varepsilon_i v_k)^p = 1, (k = 1, \ldots, t)$. Also, intersection of $\{(1 + \varepsilon_i u_j)\}$, and $\{(1 + \varepsilon_i v_k)\}$ is trivial. Hence, the direct product of the cyclic groups $\langle (1 + \varepsilon_i u_j) \rangle$, and $\langle (1 + \varepsilon_i v_k) \rangle$ exhaust $(1 + \sum_{i=1}^{s} R_0 u_i + \sum_{k=1}^{t} R_0 v_k)$. Thus, $1 + \mathcal{J}$ is of the required form, and this completes the proof.

\[\blacksquare\]

**Remark IV.11.** Proposition IV.10 is true for the two cases when $u_i^2 = 0$ and when $u_i^2 \neq 0$, for $i = 1, \ldots, s$.

**Remark IV.12.** Suppose that $pu_i = 0$ and $u_i^2 = 0$. Then, it is easy to check that $\left|\langle (1 + \varepsilon_i u_j) \rangle \right| = p$, and this can be proved in a similar manner to Propositions IV.7 and IV.8, cases where $p \in \mathcal{J}^2$.

**Remark IV.13.** If $pu_i = 0$ and $u_i^2 = 0$. Then $\left|\langle (1 + \varepsilon_i u_j) \rangle \right| = p$, if $p$ is odd, or $\left|\langle (1 + \varepsilon_i u_j) \rangle \right| = p^2$, if $p$ is even, and this can be proved in a similar manner to Propositions IV.7 and IV.9, cases where $p \in \mathcal{J}^2$, for $p$ odd or even, respectively.

**Remark IV.14.** We remark here that the cases for which $d < s$ of $pu_1, \ldots, pu_d$ is zero have similar arguments to previous results and one may deduce the structure of $1 + \mathcal{J}$ from the preceding propositions.

**Remark IV.15.** By the above results and by equation (IV.5), the structure of $U(R)$ is now determined.

Imhotep Proc.
Rings with other invariants \( p, n, r, s, t, d \) when \( J^3 = (0) \), and the cases \( J^m = (0), J^{m-1} \neq (0) \), when \( m \geq 4 \) and \( m < k \) are left for further consideration.

References


Chiteng’a J. Chikunji
Department of Basic Sciences, Botswana College of Agriculture, Private Bag 0027, Gaborone, BOTSWANA.

E-mail: jchikunj@bca.bw

Imhotep Proc.
Abstract

This paper reports a small study of secondary school students problem-solving practices in Uganda. A sample of 225 students (109 male and 116 female), in eight government and private secondary schools were used. Students problem-solving processes and strategies were assessed using three non-routine tasks. Solutions of students from government and private schools and differences by gender were categorised and rated as exceptional, proficient, satisfactory, limited, poor and blank. The solutions mean scores were also examined using t-tests of independent samples. The results of the t-tests for independent samples indicated no statistically significant differences in students scores by gender. The results of the t-tests for independent samples indicated a statistically significant difference in students scores by school type. The results suggest that teachers in both government and private schools need to realise that when concepts and skills they teach make sense to students they learn faster, they remember better, and they are better able to use concepts and skills in subsequent problem-solving situations. This study needs replicating at different school levels and contexts using qualitative data collection techniques. Mathematics educators could benefit from knowledge of students problem-solving practices as students should be involved in solving problems rather than mastering skills and not applying them since “the ability to solve problems is at the heart of mathematics” (Cockcroft, 1982, p. 73).
Mathematical problem-solving processes of male and female secondary students in government and private schools on non-routine tasks

Charles Opolot-Okurut

Abstract. This paper reports a small study of secondary school students problem-solving practices in Uganda. A sample of 225 students (109 male and 116 female), in eight government and private secondary schools were used. Students problem-solving processes and strategies were assessed using three non-routine tasks. Solutions of students from government and private schools and differences by gender were categorised and rated as exceptional, proficient, satisfactory, limited, poor and blank. The solutions mean scores were also examined using t-tests of independent samples. The results of the t-tests for independent samples indicated no statistically significant differences in students scores by gender. The results of the t-tests for independent samples indicated a statistically significant difference in students scores by school type. The results suggest that teachers in both government and private schools need to realise that when concepts and skills they teach make sense to students they learn faster, they remember better, and they are better able to use concepts and skills in subsequent problem-solving situations. This study needs replicating at different school levels and contexts using qualitative data collection techniques. Mathematics educators could benefit from knowledge of students problem-solving practices as students should be involved in solving problems rather than mastering skills and not applying them since "the ability to solve problems is at the heart of mathematics" (Cockcroft, 1982, p. 73).

Keywords. Heuristics; Non-routine tasks; Polya; Problem-solving process; Uganda.

Introduction

In every country, education is basically intended to develop individuals who are independent critical thinkers, self-confident, motivated and multitalented, and who can comfortably fit into the various needs of adult life and society. In short, the individuals should be competent problem solvers. Problem-solving is often viewed in different ways. For example, one perspective regards problem-solving as a process where previously acquired knowledge is applied in a novel and unknown situation or a process of engaging in a task for which a prior solution method/procedure/algorithm/routine is not known before hand (Burton, 1984; National Council of Teachers of Mathematics [NCTM], 2000; Polya, 1957). Thus, involvement in problem-solving enables learners to exercise and acquire confidence in their intellectual power. Another view is that problem-solving provides a context for practising and applying concepts.
and skills that Charles (2009, p.1) calls "teaching FOR problem-solving" (emphasis in original). Teaching is itself a problem-solving activity. Furthermore, NCTM (2000, p.52) regards "problem-solving [as] an integral part of all mathematics learning." Problem-solving improves mathematical proficiency that entails five strands: “conceptual understanding, procedural fluency, strategic competence, adaptive reasoning and productive disposition” (National Research Council [NRC], 2001, p. 116) to enhance students building of helpful attitudes and beliefs about mathematics. However, Lesh (1981) has argued and cautioned that “knowing to compute does not ensure that a person will know when to compute, which operation to use in a particular situation, or how to use an answer once it is obtained” (p.235).

I agree that learners need to be confident in their mathematical knowledge and in their ability to seek, acquire and apply new knowledge. They should demonstrate persistence, resolve, flexibility and ingenuity in finding the solutions to problems. They should value and be aware of the importance of communication skills in mathematics and develop a positive attitude toward mathematics. They should appreciate the usefulness of computational competence, mathematical processes and problem-solving skills which are used in the decision making and modelling processes in society. However, the ability to perform basic computations and follow procedures may not be sufficient for learners to be successful in problem-solving. Learners need to develop and apply habits of mind such as thinking critically, thinking creatively and meta-cognition in the different phases of the problem-solving process.

A major purpose of studying mathematics is to learn to solve problems. Cockcroft (1982) argued that “the ability to solve problems is at the heart of mathematics” (p. 73). Some problems can be either routine requiring only the application of a known procedure or algorithm or non-routine requiring the development of a process or the conducting of an investigation to solve them. Several scholars have developed several heuristics and strategies for mathematical problem-solving and frameworks for analysing the problem-solving process (Burton, 1984; Heinz, 2005; Polya, 1957; Silver, Mamona-Downs, Leung & Kennedy, 1996). The most versatile and commonly used scheme of problem-solving is the one formulated by Polya (1957). Polya’s four phases of problem-solving in what he calls the LIST entails: understanding the problem, devising a plan, carrying out the plan, and looking back. A similar model was suggested by Burton (1984) comprising of entry, attack, review and extension through organising questions, procedures, and skills for entry, attack, and review-extension as one hunts for a resolution rather than a solution.

Problem-solving ability has been investigated in several ways and several heuristics for problem-solving have been suggested in the literature. For example, Collis, Romberg and Jurdak (1986) used the structure of learned outcomes (SOLO) taxonomy in which the response modes are categorised into five levels, namely: (1) pre-structural, (2) uni-structural, (3) multi-structural, (4) relational, and (5) extended abstract.

A study by McCoy (1994) investigated the problem-solving actions of second and third grade school students as they engaged in solving three unfamiliar problems. She used observation and interviewing technique. The study showed that the students used systematic solution processes to attempt the problems. However, no significant gender differences were evident. Heinz (2005) investigated differences in problem-solving strategies and processes of mathematically gifted and non-gifted elementary students and found that mathematically gifted elementary students were more capable of working systematically and quickly, verbalising and explaining their solutions than non-gifted students. Kim, S.W. and Kim, S. Y. (2003) explored and compared the problem-solving processes of two groups: good novice problem solvers and poor novice problem solvers using the thinking aloud oral method followed by post-interviews. Their findings indicate that good novice problem solvers tended to use knowledge-development strategy,
while the poor novice problem solvers have a propensity to use the means-end and random strategies. The two groups also differed in the amount of time spent on the problem and the ratio of the time spent in solving the problem to the total time spent.

In a previous research study I argued that “to be most fruitful, practice in problem-solving should not consist of repeated experiences in solving the same problems with the same techniques, but should consist of the solution of different problems by the same techniques and the application of different techniques to the same problem” (Opolot-Okurut, 1989, p. 44). Unfortunately, there has been lack of success to embed problem-solving in the curriculum despite efforts to introduce problem-solving in teacher education in Uganda (Opolot-Okurut, 1989). However, piecemeal efforts by the Ministry of Education and Sports to introduce problem-solving in primary and secondary schools exist. Whereas in the United States the NCTM (2000) in the *Principles and Standards for Mathematics teaching* recommended that instructional programmes for K-12 grades should enable student to (1) build new mathematical knowledge through problem-solving; (2) solve problems that arise in mathematics and other contexts; (3) apply and adapt a variety of appropriate strategies to solve problems; and (4) monitor and reflect on the process of mathematical problem-solving this practice is rare in mathematics classrooms of the developing world, but are worth investigating.

A brief overview of the Ugandan secondary school system, to give a context for the study is as follows. The secondary education lasts six years subdivided into Ordinary Level (four years) and Advanced level (two years). For O-level, students have to progress from senior one (S1) in the first year of lower secondary schooling at age of about 13 years to senior four (S4) in fourth of secondary school at age about 16 years. Successful students proceed to A-level. For A-level, students have to progress from senior five (S5) in the first year of higher secondary schooling at age of about 17 years to senior six (S6) in second year of higher secondary school at age about 18 years. After A-level students progress to tertiary education, after about 13 years of schooling.

The purpose of this study was to investigate the response patterns of the students from government-aided and private schools and by their gender on non-routine mathematical problems or tasks. In line with the purpose of the study the following research questions were posed:

1. Are there significant differences in achievement scores in mathematics non-routine problems between students from government-aided and private schools?
2. Are there significant differences in achievement scores in mathematics non-routine problems between male and female students?

**Methodology**

**Research design**

This exploratory study was conducted in government and private secondary schools in Central Uganda. A survey research design was followed and considered suitable because, we were interested in students problem-solving strategies and processes. The target population for the study was all students enrolled in the lower secondary education classes. Students in senior three participated.

**Sample**

The data discussed in this paper is from grade-nine secondary students. The students ages ranged from 15 to 21 years in eight secondary schools in Central Uganda. Thirty students, from one class, were randomly selected in classes whose teachers volunteered to participate in the study. Eight schools (four government-aided and four private) schools participated in the study.
Data from a total of 225 students: 109(48.4%) male and 116(51.6%) female participants were used in the analysis.

**Instrument**
The tasks used in this study were adopted from those developed by Opolot-Okurut (2004) as given in Appendix A. The instrument has two parts. Part A asked respondents for their demographic information: the gender, age and school type and Part B contained three non-routine tasks, which respondents were to provide their best answer in a test format. The students were advised to write their solutions legibly and to show clearly all working steps and strategies that were used in the space provided. These tasks did not need deep mathematical knowledge to solve but they needed the ability to apply insightful basic mathematical skills. The tasks were challenging questions and the time was unlimited. The students took between 20 and 90 minutes to complete.

**Procedure**
The writing of the test was conducted in each school under strict examination conditions. The time was unlimited to eliminate pressure of time from students. Most of the students finished within an hour. Students who could not do some of the tasks left earlier. The test administration procedure followed strict examinations regulations. The marking guide for individual task rubric followed the same pattern as outlined in Appendix B. The rubric for the three tasks consisted of a paragraph statement of the requirements of an adequate solution followed by the criteria for scoring students solutions from zero (no attempt or leaves a blank page) to five points. One point was given for a solution that was ‘inadequate;’ two or three points were given to a ‘satisfactory’ solution; four or five points were given to a solution that is ‘outstanding’. Any solution was expected to indicate a student understands mathematics and application of mathematical knowledge. For example, five marks were awarded to a student who represented houses in two columns with letters A to E and F to J and joined each letter to the other letters without double counting, and counted the number of links between two letters and obtained the correct answer. A session for the coordination of marking of the test was conducted with all the markers. During the marking each script was marked by three independent markers and an agreed mark between the markers for each task obtained after discussion. Each students solutions were scored using analytic scoring in which the marker considered whether specific points on the marking guide were addressed by the solution. The solution was then scored according to the points that it closely similar to that laid down in the scoring rubric for the particular problem.

**Data Analysis**
The collected data were analysed using descriptive statistics and the means of the achievement scores were compared using the independent samples $t$-test for the different school types and gender.

**Results**
Each script of three questions was examined and marked for evidence that satisfied the criteria outlined in the assessment rubric giving a total of 675 questions. These results are given for each task in Tables 1-3. Of the 675 solutions that the students attempted 320 solutions were categorised as blank. Of the 355 non-blank solutions 143(40.3%) were categorised as exceptional; 24(6.8%) were categorised as proficient; 66(18.6%) were categorised as satisfactory; 59(16.6%) were categorised as limited; and 73(20.6%) were categorised as poor. Table 1 shows that about 30 percent of the students wrote exceptional solutions to Task 2, which was Anthony, a school
sports prefect, has to plan a football tournament involving ten schoolhouse teams. Each house-
team has to play every other house-team once. What is the total number of games to be played
that he has to plan for? Show clearly how you worked the total out, were the overall well
attempted task even if the majority of the students left it blank. Due to space limitation the
results for only Task 2 are presented here as example of solutions.

Table 1. Overall student performance on the tasks

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Score</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>Total</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exceptional</td>
<td>5</td>
<td>41</td>
<td>68</td>
<td>34</td>
<td>143</td>
<td>40.3</td>
</tr>
<tr>
<td>Proficient</td>
<td>4</td>
<td>13</td>
<td>07</td>
<td>04</td>
<td>24</td>
<td>6.8</td>
</tr>
<tr>
<td>Satisfactory</td>
<td>3</td>
<td>38</td>
<td>13</td>
<td>15</td>
<td>66</td>
<td>18.6</td>
</tr>
<tr>
<td>Limited</td>
<td>2</td>
<td>25</td>
<td>09</td>
<td>25</td>
<td>59</td>
<td>16.6</td>
</tr>
<tr>
<td>Poor</td>
<td>1</td>
<td>25</td>
<td>03</td>
<td>35</td>
<td>73</td>
<td>20.6</td>
</tr>
<tr>
<td>Blank</td>
<td>0</td>
<td>83</td>
<td>125</td>
<td>112</td>
<td>320</td>
<td>47.4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>225</strong></td>
<td><strong>225</strong></td>
<td><strong>225</strong></td>
<td><strong>675</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
</tbody>
</table>

Q1. Are there significant differences in achievement scores in mathematics non-routine
problems between students from government-aided and private schools?

Descriptive statistics were used to capture the government and private schools students re-
sponses for each task. For this task, Table 2 shows that most of the students from government-
aided schools gave exceptional solutions to this task when compared to those from the private
schools. There were nearly 50% of the students from private schools who did not attempt the
task. Comparing the mean scores by school type the \( t \)-test of independent samples shows that
there were significant school type gender differences for this task.

Table 2. Scores and percent (in brackets) of student scores by school type for
Task 2

<table>
<thead>
<tr>
<th>School Type</th>
<th>Exceptional</th>
<th>Proficient</th>
<th>Satisfactory</th>
<th>Limited</th>
<th>Poor</th>
<th>No Attempt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Government</td>
<td>44 (36)</td>
<td>05 (04)</td>
<td>09 (07)</td>
<td>07 (12)</td>
<td>01 (06)</td>
<td>56 (46)</td>
</tr>
<tr>
<td>Private</td>
<td>24 (23)</td>
<td>02 (02)</td>
<td>04 (04)</td>
<td>02 (02)</td>
<td>02 (02)</td>
<td>69 (67)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>68</strong></td>
<td><strong>07</strong></td>
<td><strong>13</strong></td>
<td><strong>09</strong></td>
<td><strong>03</strong></td>
<td><strong>125</strong></td>
</tr>
</tbody>
</table>

Total \( N = 225 \), Government \( n = 122 \), Private \( n = 103 \); \( t(220.62) = 3.01, p < .05 \), sig.

Q.2 Are there significant differences in achievement scores in mathematics non-routine problems
between male and female students?

Descriptive statistics were used to capture the male and female students responses for each
task. For this task Table 3 shows that most of the students turned in exceptional solutions to
this task but over 50% of the students did not attempt the task. Comparing the mean scores by
gender the \( t \)-test of independent samples shows that there were no significant gender differences
for this task.
Table 3. Scores and percent (in brackets) of student scores by gender for Task 2

<table>
<thead>
<tr>
<th>Gender</th>
<th>Exceptional</th>
<th>Proficient</th>
<th>Satisfactory</th>
<th>Limited</th>
<th>Poor</th>
<th>No Attempt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>33 (30)</td>
<td>02 (02)</td>
<td>09 (08)</td>
<td>03 (03)</td>
<td>02 (02)</td>
<td>60 (55)</td>
</tr>
<tr>
<td>Female</td>
<td>35 (30)</td>
<td>05 (04)</td>
<td>04 (03)</td>
<td>06 (05)</td>
<td>01 (01)</td>
<td>65 (56)</td>
</tr>
<tr>
<td>Total</td>
<td>68</td>
<td>07</td>
<td>13</td>
<td>09</td>
<td>03</td>
<td>125</td>
</tr>
</tbody>
</table>

Total N = 225, Male n = 109, Female n = 116; t(223) = 0.04, p > .05, ns.

Discussion

This study set out to investigate lower secondary students’ achievement in non-routine mathematics tasks in government and private schools and by their gender. The paper has captured the solutions strategies and process that included for instance taking smaller number of letters to represent houses and working through to 10 houses; writing digits 1 to 10 to represent the houses and drawing loops from each digit to the other digits and counting them; and using 10 dots to represent houses and joining each dot to others and counting the number of lines, for each task using descriptive statistics. The students’ solutions were rated on six point scale ranging from blank (0) to exceptional solution (5). The mean achievement scores were compared by school type and by gender using an independent samples t-test. One key finding was, students attending both government and private schools were able to submit some exceptional solutions to given tasks. In general there are statistically significant differences in student achievement by school type. This result increases our knowledge of the problem-solving strategies used by students in lower secondary schools in Uganda. There is evidence that the common heuristics that students applied to solve the tasks were drawing diagrams and tables, attempting trial and error methods, listing and counting on strategies. This finding supports the type of strategies students in the McCoys (1994) study that showed that students closely followed Polya’s (1957) stages of problem-solving: understanding the problem, devising a plan, carrying out the plan, and looking back. This result suggests that there is need for teachers to encourage students to participate in problem-solving in school mathematics. The national examining body considers including problem-solving items in the public examinations; and text books authors think of incorporating problem-solving content via appropriate situations in their books and to include problem-solving tasks to promote problem-solving. In sum, students could solve problems using different heuristics and strategies if teachers nurture them to do so. This finding reveals that the achievement of students on non-routine tasks did vary significantly according to the type of school they were enrolled.

Another significant finding from this study was that both male and female students turned in exceptional solutions to the given tasks. This finding is not surprising because most of the schools are coeducational and students discuss among themselves some of their work. There was evidence that the heuristics commonly used were guessing and trial and error, algebra, listing counting on and underlining key words in each of the problems for all problems. However, some of the students gave up giving comments like ‘we are not told how many lessons there were before break’, students showing evidence of lack of understanding of the problem claiming ‘It is impossible for the bells to ring at once’, students introducing their own conditions and assumptions such as ‘assume four periods in each section before break’, some students used the contextual knowledge of the timing of lessons in the school timetable and guessing that ‘the bells will all ring at the same time at 1:00 pm for lunch.’ This finding suggests that the performance of students on non-routine tasks did not vary significantly according to their gender.
Conclusion

The findings and their discussion lead to the following conclusions:

1. The achievement of students on non-routine tasks did vary significantly according to the type of school they were enrolled.
2. The performance of students on non-routine tasks did not vary significantly according to their gender.

This study has investigated problem-solving behaviour of lower secondary students in Uganda. Although the study has few findings, several limitations can be cited that affect how generalisable the results of the study are to which the reader is to take the results with caution for the following reasons. First, the format of the non-routine tasks in the test maybe a factor that could influence the achievement results in this study. Second, the sample was from one area of the country close to an urban setting. Thus, these results cannot be generalised to students in other districts and schools of the country. And thirdly, the method of data collection did not include interviews and observation but limited to written solutions survey.

Implications for teaching and learning

In summary, the findings of this students problem-solving processes study have several interesting and important implications for both practice and further research. From a practical point of view, three implications are apparent from the findings. First, teachers should be made aware of the need to nurture students mathematical proficiency. For example, more opportunities must be availed to students to develop the mathematical proficiency strands (NRC, 2001). Second, administrators, especially those in the private schools, in light of the findings of this study must provide more teacher support to enable them concentrate less on only having student pass examinations. Third, systematic efforts must be made to promote problem-solving in school mathematics by focusing on teacher pre-service preparation programmes and in-service education through seminars and workshops. Fourth, mathematics teachers need to develop student mathematical understanding through a balance of solving problems and presenting examples to them in the context of a problem-focussed and question-driven classroom conversation often regarded as ‘teaching THROUGH problem-solving’. Fifth, in general, teachers need to pay more attention to the development of students problem-solving skills in their classrooms. Finally, encourage textbook authors to include problem-solving tasks in their books and introduce content in these books via appropriate problem situations. In terms of further research the following areas are suggested as needing more research. (1) Future investigations may extend the research of the current study by exploring students achievement and school-related variables in addition to those used here. (2) This study could also be replicated at different levels of education (primary, upper secondary, teacher education and tertiary levels. (3) The use of qualitative research approaches to the investigations could be considered. (4) Other factors that could be causes or associated with the student achievement should be explored.

References

APPENDICES

Appendix A: Mathematics Problem Solving Tasks

STUDENT CODE ................ GENDER: M ........ F ......... SCHOOL CODE.......
AGE......... TIME: 1 hour

INSTRUCTIONS:
(a) Attempt all problems. All reasonable solutions are acceptable.
(b) All problems carry the same number of marks.
   (i) Write legibly, and show all your working steps clearly.
   (ii) Show all strategies you use and working in the answer sheets provided.

Task 1
A school is divided into lower, middle and upper sections. The change-of-lesson bell rings after thirty, forty and forty five minutes in the lower, medium and upper sections respectively.
(a) Morning break falls when the bells ring at the same time. If the lessons in the lower and middle section start at 8.00 am at what time does break start?
(b) The break lasts 30 minutes. After break, lessons start at the same time in the lower, middle and upper sections. At what time will the three bells ring at the same time?

Task 2
Anthony, a school sports prefect, has to plan a football tournament involving ten school house teams. Each house-team has to play every other house-team once. What is the total number of games to be played that he has to plan for? Show clearly how you worked out the total.

Task 3
Children are seated around a table and pass round a packet of sixteen sweets. Ekanya takes the first sweet. Each child then takes one sweet at a time as the packet is passed around. Ekanya also receives the last sweet. Find three possible numbers of children seated on the table and how many sweets each one gets in each case.

(Source: Opolot-Okurut, 2004)

Thank you for your cooperation
Appendix B: Solutions Categories, Indicators and Scores for Non-Routine Tasks

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>INDICATORS</th>
<th>RATING</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXCEPTIONAL</td>
<td>- fully acceptable solution</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>- accurate execution of strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- appropriate choice of strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- correct interpretation of problem</td>
<td></td>
</tr>
<tr>
<td>PROFICIENT</td>
<td>- mostly acceptable solution</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>- accurate execution of strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- appropriate choice of strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- correct interpretation of problem</td>
<td></td>
</tr>
<tr>
<td>SATISFACTORY</td>
<td>- partly unacceptable solution</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>- minor errors in execution of strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- appropriate choice of strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- correct interpretation of problem</td>
<td></td>
</tr>
<tr>
<td>LIMITED</td>
<td>- unacceptable solution</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>- accurate execution of incorrect strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- inappropriate choice of strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- correct interpretation of problem</td>
<td></td>
</tr>
<tr>
<td>POOR</td>
<td>- no solution, working abandoned</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>- inaccurate execution of wrong strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- inappropriate choice of strategy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- incorrect interpretation of problem</td>
<td></td>
</tr>
<tr>
<td>BLANK</td>
<td>No attempt, empty answer script</td>
<td>0</td>
</tr>
</tbody>
</table>

A Generic Rubric for Scoring Non-Routine Tasks

<table>
<thead>
<tr>
<th>CRITERIA</th>
<th>SCORE</th>
<th>SOLUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Attempts to extend the problem; contains a full complete solution;</td>
<td>5</td>
<td>As given</td>
</tr>
<tr>
<td>correct interpretation of problem; correct strategy identified and</td>
<td></td>
<td></td>
</tr>
<tr>
<td>followed.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Starts with a correct interpretation of the problem; identifies correct</td>
<td>4</td>
<td>As given</td>
</tr>
<tr>
<td>strategies; gives a complete solution with minor errors.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Interprets the problem correctly starts with a correct strategy;</td>
<td>3</td>
<td>As given</td>
</tr>
<tr>
<td>follows some wrong steps; part correct solution.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Gives incomplete solution; shows some errors; starts with an</td>
<td>2</td>
<td>As given</td>
</tr>
<tr>
<td>appropriate strategy.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Begins with an inappropriate strategy; misunderstands the question;</td>
<td>1</td>
<td>As given</td>
</tr>
<tr>
<td>shows major errors; incomplete solution.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• No attempt or response.</td>
<td>0</td>
<td>Nil</td>
</tr>
</tbody>
</table>

Charles Opolot-Okurut
e-mail: copolok@gmail.com
College of Education and External Studies Makerere University (Uganda)
Construction of Some New Three Associate Class Partially Balanced Incomplete Block Designs in Two Replicates

E. C. Kipkemoi
kipched@gmail.com

J. k. Koske
koske4@co.uk

J. M. Mutiso
johnkasome@yahoo.com

Department of Statistics and Computing, Moi University, P O Box 3900, Eldoret, Kenya.

Abstract

Search for experimental designs which aid in research studies involving large number of treatments with minimal experimental units has been desired overtime. This paper constructs some new series of three associate Partially Balanced Incomplete Block (PBIB) designs having \( n(n - 2)/4 \) treatments with three associate classes in two replicates using the concept of triangular association scheme. The design is constructed from an even squared array of \( n \) rows and \( n \) columns (\( n \geq 8 \)) with its both diagonal entries bearing no treatment entries and that given the location of any treatment in the squared array, the other location of the same treatment in the array is predetermined. The design and association parameters for a general case of an even integer \( n \geq 8 \) are obtained with an illustrated case for \( n = 8 \). Efficiencies of the designs within the class of designs are obtained for a general case of even \( n \geq 8 \) with a listing of efficiencies of designs with blocks sizes in the interval \([8,22]\). The designs constructed have three associate classes and are irreducible to minimum number of associate classes.
Construction of Some New Three Associate Class Partially Balanced Incomplete Block Designs in Two Replicates

E. C. Kipkemoi, J. k. Koske and J. M. Mutiso

Abstract. Search for experimental designs which aid in research studies involving large number of treatments with minimal experimental units has been desired overtime. This paper constructs some new series of three associate Partially Balanced Incomplete Block (PBIB) designs having \( n(n - 2)/4 \) treatments with three associate classes in two replicates using the concept of triangular association scheme. The design is constructed from an even squared array of \( n \) rows and \( n \) columns \( (n \geq 8) \) with its both diagonal entries bearing no treatment entries and that given the location of any treatment in the squared array, the other location of the same treatment in the array is predetermined. The design and association parameters for a general case of an even integer \( n \geq 8 \) are obtained with an illustrated case for \( n = 8 \). Efficiencies of the designs within the class of designs are obtained for a general case of even \( n \geq 8 \) with a listing of efficiencies of designs with blocks sizes in the interval \([8,22]\). The designs constructed have three associate classes and are irreducible to minimum number of associate classes.

Keywords. Partially Balanced Incomplete Block (PBIB), Associate class, three associate classes.

Introduction

By changing the arrangement of treatments or omitting certain blocks and or treatments, we obtain designs that may belong to a class of new designs. Using this technique Bose and Nair (1939) introduced some PBIB designs. Atiquallah (1958) established that considering PBIB designs based on triangular association scheme with \( v = n(n - 2)/2 \), \( b = (n - 1)(n - 2)/2 \), \( r = k = n - 2 \), \( \lambda_1 = 1 \) and \( \lambda_2 = 2 \). Other methods were given by Shrikande (1960, 1965) and Chang et al (1965) based on the existence of certain BIB designs and considering the dual of a BIB design as omitting certain blocks from the BIB design. John (1966) showed that triangular association scheme can be described by representing the treatments by ordered pairs \((x,y)\) with \( 1 \leq x < y \leq q \) and then generalized (John, 1966) for the case that \( v = q(q - 1)(q - 2))/6 \), \((q > 3)\). Arya and Narain (1981) discussed a new association scheme called truncated triangular (TT) with five associate classes when \( v = p((p - 2))/2 \) with \( p \) an even positive integer \( \geq 8 \), and used to construct partial diallel crosses. Ching-Shui et al (1984) came up with a general and simple method of construction based on the relation of triangular and \( 1_2 \) type of PBIB design.

**Construction of designs**

The design is constructed from a squared array of \( n \) rows and \( n \) columns (\( n \) is even positive integer \( \geq 8 \)) with both diagonal entries in the array having no treatments allocated to as illustrated in the figure below:

\[
\begin{array}{cccccccc}
* & n_{12} & n_{13} & n_{14} & n_{15} & n_{16} & n_{17} & * \\
n_{21} & * & n_{23} & n_{24} & n_{25} & n_{26} & * & n_{28} \\
n_{31} & * & n_{32} & * & n_{34} & n_{35} & * & n_{37} \\
n_{41} & n_{42} & n_{45} & * & * & n_{46} & n_{47} & n_{48} \\
n_{51} & n_{52} & n_{53} & * & * & n_{56} & n_{57} & n_{58} \\
n_{61} & n_{62} & * & n_{64} & n_{65} & * & n_{67} & n_{68} \\
n_{71} & * & n_{73} & n_{74} & n_{75} & n_{76} & * & n_{78} \\
* & n_{82} & n_{83} & n_{84} & n_{85} & n_{86} & n_{87} & *
\end{array}
\]

\( i \) and \( j \) are integers (\( 1 \leq i, j \leq n \)) such that:

- Each row and column of the square array has \( n - 2 \) treatment entries.
- The treatment entries are allocated to as illustrated in the figure below

1. The initial set of \( v \) treatment entries are first filled on one triangle enclosed by the two diagonal and a side of the square array.
2. The second set of \( v \) treatment entries are replicated in each of the remaining triangles by simply reflecting the initial set of \( v \) treatment entries subsequently with the diagonal blank entries as mirrors in such a way that given any two entries \( n_{ij} \) and \( n_{ij} \) are allocated to treatment \( x \) if and only if the subscripts \( i + i = j + j \) or \( i + i + j + j = 2(n + 1) \).

Taking each row and column to constitute a block we obtain \( n/2 \) distinct blocks and thus a design with parameters

\[
v = \frac{n(n - 2)}{4} \quad b = n/2 \quad k = n - 2 \quad r = 2 \quad \lambda_1 = 2 \quad \lambda_2 = 1 \quad \lambda_3 = 0
\]

**Association scheme**

Two treatments are said to be:

- First associates if they both occur in the same row and column.
- Second associates if they both occur in the same row or the same column but not both.
- Third associates if they neither occur in the same row nor in the same column. Giving rise to the following association parameters.

\[
n_1 = 1 \quad n_2 = 2(n - 4) \quad n_2 = \frac{n(n-10)+24}{4}
\]

Imhotep Proc.
Construction of Some New Three Associate Class

\[ P_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2(n-4) & 0 \\ 0 & 0 & 0 & \frac{n(n-10)+24}{4} \end{pmatrix}, \quad P_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 2(n-4) & 0 \\ 0 & 0 & 0 & \frac{n(n-10)+24}{4} \end{pmatrix} \]

\[ P_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & n-4 & \frac{n-6}{4} \\ 0 & 0 & n-6 & \frac{n(n-14)+48}{4} \end{pmatrix}, \quad P_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 8 & \frac{2(n-8)}{4} \\ 1 & 1 & 2(n-8) & \frac{n(n-18)+80}{4} \end{pmatrix} \]

Illustration: taking \( n = 8 \) we obtain three associate class PBIB design with the parameters

\( v = 12 \) \( b = 4 \) \( k = 6 \) \( r = 2 \)

\( \lambda_1 = 2 \) \( \lambda_2 = 1 \) \( \lambda_3 = 0 \)

\( n_1 = 1 \) \( n_2 = 8 \) \( n_3 = 2 \)

Whose blocks are:
1. (1, 2, 3, 4, 5, 6)
2. (1, 6, 7, 8, 9, 10)
3. (2, 5, 7, 10, 11, 12)
4. (3, 4, 8, 9, 11, 12)

With the parameters of the second kind given by:

\[ P_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad P_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \]

\[ P_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 8 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \]

The efficiency factors for this class of designs are given by

\[ E_1 = \frac{n-4}{n-2} \]

\[ E_2 = -\frac{n(n-4)^2}{7n^2-6n-16-n^2} \]

\[ E_3 = -\frac{n(n-4)^2}{6n^2+2n-32-n^2} \]

and the overall efficiency factor of the design is

\[ E = \frac{1}{n(n-2)^4} \left( \frac{4(n-4)}{n-2} - \frac{8n(n-4)^3}{7n^2-6n-16-n^2} + \frac{n(n-6)(n-4)^3}{6n^2+2n-32-n^2} \right) \]

Efficiencies of three associate class PBIB designs having \((n(n-2))/4\) treatments with two replicates for \( 6 \leq k \leq 22 \) are given in the table below.
Conclusion

In this paper we have constructed some new series of three associate class PBIB designs having \((n(n - 2))/4\) treatments with two replications. The restriction of the number of replications to two helps to minimize cost. The average efficiency factors of these designs along with the three efficiencies factors \(E_1\), \(E_2\) and \(E_3\) are quite high for practical purposes.

References


Imhotep Proc.
E. C. Kipkemoi
e-mail: kipched@gmail.com
Department of Statistics and Computing, Moi University, P O Box 3900, Eldoret, Kenya.

J. k. Koske
e-mail: koske4@co.uk
Department of Statistics and Computing, Moi University, P O Box 3900, Eldoret, Kenya.

J. M. Mutiso
e-mail: johnkasome@yahoo.com
Department of Statistics and Computing, Moi University, P O Box 3900, Eldoret, Kenya.
On some sequence spaces and $A$-statistical convergence

Ekrem Savaş
Istanbul Ticaret University,
Department of Mathematics,
Sütlüce, Istanbul - TURKEY
ekremsavas@yahoo.com
esavas@ticaret.edu.tr

Abstract
In the present paper, we introduce and study some properties of the a new sequence space that is defined using the $\varphi$-function and de la Valee-Poussin mean.
On some sequence spaces and $A$-statistical convergence

Ekrem Savaş

Abstract. In the present paper, we introduce and study some properties of the a new sequence space that is defined using the $\varphi$-function and de la Valee-Poussin mean.

Mathematics Subject Classification (2000). Primary 40H05; Secondary 40C05.

Keywords. Modulus function, Invariant Mean, Almost convergence, Lacunary sequence, $\varphi$-function.

I. Introduction and Background

Let $w$ denote the set of all real and complex sequences $x = (x_k)$. By $l_\infty$ and $c$, we denote the Banach spaces of bounded and convergent sequences $x = (x_k)$ normed by $||x|| = \sup_n |x_n|$, respectively. A sequence $x \in l_\infty$ is said to be almost convergent if all of its Banach limits coincide. Let $\hat{c}$ denote the space of all almost convergent sequences. Lorentz [4] has shown that

$$\hat{c} = \{ x \in l_\infty : \lim_m t_{m,n}(x) \text{ exists uniformly in } n \}$$

where

$$t_{m,n}(x) = \frac{x_n + x_{n+1} + x_{n+2} + \cdots + x_{n+m}}{m+1}.$$ 

The space $[\hat{c}]$ of strongly almost convergent sequences was introduced by Maddox [5] as follows:

$$[\hat{c}] = \{ x \in l_\infty : \lim_m t_{m,n}(|x - L|) = 0, \text{ uniformly in } n, \text{ for some } L \}.$$

Let $\sigma$ be a mapping of the set of positive integers into itself. A continuous linear functional $\phi$ on $l_\infty$ is said to be an invariant mean or a $\sigma$-mean if and only if

1. $\phi(x) \geq 0$ when the sequence $x = (x_k)$ has $x_n \geq 0$ for all $n$;
2. $\phi(e) = 1$ where $e = (1, 1, 1, \ldots)$ and
3. $\phi(x_{\sigma(n)}) = \phi(x)$ for all $x \in l_\infty$.

Let $V_\sigma$ denote the set of bounded sequences which have unique $\sigma$-mean (see, [16]).

It was quite natural to expect that invariant mean must give rise to a new type of convergence, namely, strong invariant convergence, just as almost convergence gives rise to the concept of strong almost convergence and this concept was introduced and discussed by

Paper presented at the 2nd Strathmore International Mathematics Conference (SIMC 2013), 12 - 16 August 2013, Strathmore University, Nairobi, Kenya.
Mursaleen [8]. If $[V_σ]$ denotes the set of all strongly $σ$-convergent sequences, then Mursaleen defined,

$$[V_σ] = \{ x \in l_∞ : \lim_{m \to \infty} \frac{1}{m} \sum_{k=1}^{m} |x_{σ^k(n)} - l| = 0, \text{ uniformly in } n \}.$$  

Here $σ^k(n)$ denotes the $k$th iterate of the mapping $σ$ at $n$. Taking $σ(n) = n + 1$, we obtain $[V_σ] = [c]$ so that strong $σ$-convergence generalizes the concept of strong almost convergence.

Let $λ = (λ_i)$ be a non-decreasing sequence of positive numbers tending to $∞$ such that

$$λ_{i+1} ≤ λ_i + 1, λ_1 = 0.$$  

The generalized de la Valée-Poussin mean is defined by

$$T_i(x) = \frac{1}{λ_i} \sum_{k ∈ I_i} x_k$$  

where $I_i = [i - λ_i + 1, i]$. A sequence $x = (x_n)$ is said to be $(V,λ)$-summable to a number $L$, if

$$T_i(x) \to L \text{ as } i → ∞ \text{ (see [6]) .}$$  

Quite Recently Malkowsky and Savaš [6] have introduced the space $[V,λ]$ of $λ$-strongly convergent sequences as follows:

$$[V,λ] = \left\{ x = (x_k) : \lim_{i} \frac{1}{λ_i} \sum_{k ∈ I_i} |x_k - L| = 0, \text{ for some } L \right\}.$$  

Note that in the special case where $λ_i = i$, the space $[V,λ]$ reduces the space $w$ of strongly Cesaro summable sequences which is defined by

$$w = \left\{ x = (x_k) : \lim_{i} \frac{1}{i} \sum_{k=1}^{i} |x_k - L| = 0, \text{ for some } l \right\}.$$  

Ruckle [10] used the idea of a modulus function $f$ to construct a class of FK spaces

$$L(f) = \left\{ x = (x_k) : \sum_{k=1}^{∞} f(|x_k|) < ∞ \right\}$$  

The space $L(f)$ is closely related to the space $l_1$ which is an $L(f)$ space with $f(x) = x$ for all real $x ≥ 0$.

Recently E. Savas [13] generalized the concept of strong almost convergence by using a modulus $f$ and examined some properties of the corresponding new sequence spaces.

Following Ruckle, a modulus function $f$ is a function from $[0,∞)$ to $[0,∞)$ such that

(i) $f(x) = 0$ if and only if $x = 0$,

(ii) $f(x + y) ≤ f(x) + f(y)$ for all $x, y ≥ 0$,

(iii) $f$ increasing,

(iv) $f$ is continuous from the right at zero.

Since $|f(x) - f(y)| ≤ f(|x - y|)$, it follows from condition $(iv)$ that $f$ is continuous on $[0,∞)$. Furthermore, we have $f(nx) ≤ nf(x)$ for all $n ∈ N$, from condition $(ii)$ and so

$$f(x) = f \left( nx \frac{1}{n} \right) ≤ nf \left( \frac{x}{n} \right) \text{ hence}$$  

$$\frac{1}{n}f(x) ≤ f \left( \frac{x}{n} \right) \text{ for all } n ∈ N$$
A modulus may be bounded or unbounded. For example, \( f(x) = x^p \), for \( 0 < p \leq 1 \) is unbounded, but \( f(x) = \frac{x}{1+x} \) is bounded.

By a \( \varphi \)-function we understand a continuous non-decreasing function \( \varphi(u) \) defined for \( u \geq 0 \) and such that \( \varphi(0) = 0, \varphi(u) > 0 \) for \( u > 0 \) and \( \varphi(u) \to \infty \) as \( u \to \infty \), (see, [18]).

A \( \varphi \)-function \( \varphi \) is called non weaker than a \( \varphi \)-function \( \psi \) and we write \( \varphi \preceq \psi \) if there are constants \( c, b, k, l > 0 \) such that \( c\psi(lu) \leq b\varphi(ku) \), (for all large or small \( u \), respectively).

A \( \varphi \)-function \( \varphi \) is called non equivalent to a \( \varphi \)-function \( \psi \) and we write \( \varphi \asymp \psi \) if there are positive constants \( b_1, b_2, c, k_1, k_2, l \) such that \( b_1\varphi(k_1u) \leq c\psi(lu) \leq b_2\varphi(k_2u) \), (for all large or small \( u \), respectively), (see, [18]).

A \( \varphi \)-function \( \varphi \) is said to satisfy the condition \( (\Delta_2) \), (for all large or small \( u \), respectively) if for some constant \( k_1 > 1 \) there is satisfied the inequality \( \varphi(2u) \leq k\varphi(u) \).

In this paper, we introduce and study some properties of the following sequence space that is defined using the \( \varphi \)-function and de la Valée-Poussin mean and also some inclusion theorems are obtained.

**II. Main Results**

Let \( \Lambda = (\lambda_n) \) be same in the above, \( \varphi \) be given \( \varphi \)-function and \( f \) be given modulus function, respectively. Moreover, let an infinite matrix \( A = (a_{nk}) \) be given. Then we write,

\[
V_0^\lambda((A, \varphi, \sigma), f) = \left\{ x = (x_k) : \lim_j \frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|) \right) = 0, \text{ uniformly in } i \right\}.
\]

If \( x \in V_0^\lambda((A, \varphi, \sigma), f) \), the sequence \( x \) is said to be \( \lambda \)-strong \( (A, \varphi, \sigma) \)- convergent to zero with respect to a modulus \( f \). When \( \varphi(x) = x \) for all \( x \), we obtain,

\[
V_0^\lambda((A, \sigma), f) = \left\{ x = (x_k) : \lim_j \frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|) \right) = 0, \text{ uniformly in } i \right\}.
\]

If \( f(x) = x \), we write

\[
V_0^\lambda(A, \varphi, \sigma) = \left\{ x = (x_k) : \lim_j \frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|) \right) = 0, \text{ uniformly in } i \right\}.
\]

If we take \( A = I \) and \( \varphi(x) = x \) respectively, then we have

\[
V_0^\lambda((I, \sigma), f) = \left\{ x = (x_k) : \lim_j \frac{1}{\lambda_j} \sum_{k \in I_j} f\left( |x_{\sigma^k(i)}| \right) = 0, \text{ uniformly in } i \right\}.
\]
Theorem II.1. Let $A = (a_{nk})$ be an infinite matrix and let the $\phi$-function $\varphi(u)$ satisfies the condition ($A_2$). Then the following conditions are true.

(a) If $x = (x_k) \in V_\lambda^0((A, \varphi, \sigma), f)$ and $\alpha$ is an arbitrary number, then $\alpha x \in V_\lambda^0((A, \varphi, \sigma), f)$. 

Imhotep Proc.
(b) If \( x, y \in V^0_\lambda((A, \varphi, \sigma), f) \) where \( x = (x_k) \), \( y = (y_k) \) and \( \alpha, \beta \) are given numbers, then \( \alpha x + \beta y \in V^0_\lambda((A, \varphi, \sigma), f) \).

**Proof.** (a) Let \( x \in V^0_\lambda((A, \varphi, \sigma), f) \). First let us remark that for \( 0 < \alpha < 1 \) we write

\[
\frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \left| \sum_{k=1}^\infty a_{nk} \varphi(|\alpha x_{\sigma^k(i)}|) \right| \right) \leq \frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \left| \sum_{k=1}^\infty a_{nk} \varphi(|x_{\sigma^k(i)}|) \right| \right).
\]

Furthermore, if \( \alpha > 1 \) then we may find a positive number \( s \) such that \( \alpha < 2^s \) and we have

\[
\frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \left| \sum_{k=1}^\infty a_{nk} \varphi(|\alpha x_{\sigma^k(i)}|) \right| \right) \leq \frac{1}{\lambda_j} \sum_{n \in I_j} f\left( d^s \left| \sum_{k=1}^\infty a_{nk} \varphi(|x_{\sigma^k(i)}|) \right| \right),
\]

\[
\leq \frac{L}{\lambda_j} \sum_{n \in I_j} f\left( \left| \sum_{k=1}^\infty a_{nk} \varphi(|x_{\sigma^k(i)}|) \right| \right),
\]

where \( d \) and \( L \) are constants connected with the properties of \( \varphi \) and \( f \). Therefore we obtain the condition (a).

(b) In the following let the numbers \( \alpha, \beta \) and the elements \( x, y \in V^0_\lambda((A, \varphi, \sigma), f) \) be given. From the part (a) it follows that the following inequality is true

\[
\frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \left| \sum_{k=1}^\infty a_{nk} \varphi(|\alpha x_{\sigma^k(i)} + \beta x_{\sigma^k(i)}|) \right| \right) \leq L_1 \frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \left| \sum_{k=1}^\infty a_{nk} \varphi(|x_{\sigma^k(i)}|) \right| \right) + L_2 \frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \left| \sum_{k=1}^\infty a_{nk} \varphi(|x_{\sigma^k(i)}|) \right| \right),
\]

where the constants \( L_1 \) and \( L_2 \) are defined as in (a). Hence \( \alpha x + \beta y \in V^0_\lambda((A, \varphi, \sigma), f) \).

**Theorem II.2.** Let \( f \) be an modulus function. Then

\[
V^0_\lambda(A, \varphi, \sigma) \subset V^0_\lambda((A, \varphi, \sigma), f).
\]

**Proof.** Let \( x \in V^0_\lambda(A, \varphi, \sigma) \). For a given \( \varepsilon > 0 \) we choose \( 0 < \delta < 1 \) such that \( f(x) < \varepsilon \) for every \( x \in [0, \delta] \). We can write

\[
\frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \left| \sum_{k=1}^\infty a_{nk} \varphi(|x_{\sigma^k(i)}|) \right| \right) = S_1 + S_2 \]

where \( S_1 = \frac{1}{\lambda_j} \sum_{n \in I_j} f\left( \left| \sum_{k=1}^\infty a_{nk} \varphi(|x_{\sigma^k(i)}|) \right| \right) \) and this sum is taken over

\[
\sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|) \leq \delta
\]

Imhotep Proc.
and
\[ S_2 = \frac{1}{\lambda_j} \sum_{n \in I_j} f \left( \left\| \sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|) \right\| \right) \]
and this sum is taken over
\[ \sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|) > \delta. \]
By definition of the modulus \( f \) we have \( S_1 = \frac{1}{\lambda_j} \sum_{n \in I_j} f(\delta) = f(\delta) < \varepsilon \) and furthermore
\[ S_2 = f(1) \frac{1}{\delta} \frac{1}{\lambda_j} \sum_{n \in I_j} \sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|). \]
Hence we have \( x \in V^0_\lambda((A, \varphi, \sigma), f) \).
This completes the proof.

III. \( A^- \) statistical convergence
The idea of convergence of a real sequence had been extended to statistical convergence by Fast [1] (see also Schoenberg [17]) as follows:
A sequence \( x = (x_k) \) of real numbers is said to be statistically convergent to \( L \) if for arbitrary \( \varepsilon > 0 \), the set \( K(\varepsilon) = \{ n \in \mathbb{N} : |x_k - L| \geq \varepsilon \} \) has natural density zero, i.e.,
\[ \lim_{n} \frac{1}{n} \sum_{k=1}^{n} \chi_{K}(\varepsilon)(k) = 0, \]
where \( \chi_{K}(\varepsilon) \) denotes the characteristic function of \( K(\varepsilon) \).
Statistical convergence turned out to be one of the most active areas of research in summability theory after the work of Fridy [2] and Šalát [11]. Di Maio and Kočinac [7] introduced the concept of statistical convergence in topological spaces and statistical Cauchy condition in uniform spaces and established the topological nature of this convergence.
In another direction, a new type of convergence called \( \lambda^- \) statistical convergence was introduced in [8] as follows.
A sequence \( x = (x_k) \) of real numbers is said to be \( \lambda^- \) statistically convergent to \( L \) (or, \( S_{\lambda^-} \)-convergent to \( L \)) if for any \( \varepsilon > 0 \),
\[ \lim_{j \to \infty} \frac{1}{\lambda_j} \{ k \in I_j : |x_k - L| \geq \varepsilon \} = 0 \]
where \(|A|\) denotes the cardinality of \( A \subset \mathbb{N} \). In [8] the relation between \( \lambda^- \) statistical convergence and statistical convergence was established among other things.
Recently E. Savas [15] defined almost \( \lambda^- \)-statistical convergence by using the notion of \((V, \lambda)\)-summability to generalize the concept of statistical convergence.
Let the infinite matrix \( A = (a_{nk}) \), the sequence \( x = (x_k) \), the \( \varphi^- \) function \( \varphi(u) \) and a positive number \( \varepsilon > 0 \) be given. We write, for all \( i \)
\[ K^\lambda(\varphi, \sigma, \varepsilon) = \{ n \in I_j : \sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|) \geq \varepsilon \}. \]
The sequence $x$ is said to be $(A,\varphi,\sigma)$-statistically convergent to a number zero if for every $\varepsilon > 0$
\[
\lim_{r} \frac{1}{\lambda_j} \mu(K^j_\lambda((A,\varphi,\sigma),\varepsilon)) = 0, \text{ uniformly in } i
\]
where $\mu(K^j_\lambda((A,\varphi,\sigma),\varepsilon))$ denotes the number of elements belonging to $K^j_\lambda((A,\varphi,\sigma),\varepsilon)$. We denote by $S^0_j((A,\varphi,\sigma))$, the set of sequences $x = (x_k)$ which are uniform $(A,\varphi,\sigma)$-statistical convergent to zero. We write
\[
S^0_j((A,\varphi,\sigma)) = \left\{ x = (x_k) : \lim_{r} \frac{1}{\lambda_j} \mu(K^j_\lambda((A,\varphi,\sigma),\varepsilon)) = 0, \text{ uniformly in } i \right\}
\]
If we take $A = I$ and $\varphi(x) = x$ respectively, then $S^0_j((A,\varphi,\sigma))$ reduce to $S^0_{\sigma,\lambda}$ which was defined as follows:
\[
S^0_{\sigma,\lambda} = \left\{ x = (x_k) : \lim_{r} \frac{1}{\lambda_j} \| \{ k \in I_j : |x_{\sigma^k(i)}| \geq \varepsilon \} \| = 0, \text{ uniformly in } i \right\}
\]
Finally we conclude this paper by stating the following theorem.

**Theorem III.1.** If $\psi \prec \varphi$ then $S^0_j((A,\psi)) \subset S^0_j((A,\varphi))$.

**Proof.** By assumption we have $\psi(|x_{\sigma^k(i)}|) \leq b \varphi(c|x_{\sigma^k(i)}|)$ and we have for all $i$,
\[
\sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|) \leq b \sum_{k=1}^{\infty} a_{nk} \varphi(c|x_{\sigma^k(i)}|) \leq L \sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|)
\]
for $b, c > 0$, where the constant $L$ is connected with properties of $\varphi$. Thus, the condition $\sum_{k=1}^{\infty} a_{nk} \varphi(|x_{\sigma^k(i)}|) \geq \varepsilon$ implies the condition $\sum_{k=1}^{\infty} a_{nk} \varphi(c|x_{\sigma^k(i)}|) \geq \varepsilon$ and in consequence we get
\[
\mu(K^j_\lambda((A,\varphi,\sigma),\varepsilon)) \subset \mu(K^j_\lambda((A,\psi,\sigma),\varepsilon))
\]
and
\[
\lim_{r} \frac{1}{\lambda_j} \mu(K^j_\lambda((A,\varphi,\sigma),\varepsilon)) \leq \lim_{r} \frac{1}{\lambda_j} \mu(K^j_\lambda((A,\psi,\sigma),\varepsilon))
\]
This completes the proof. 

**References**


Imhotep Proc.

Ekrem Savaş
Istanbul Ticaret University , Department of Mathematics, Sütluce, Istanbul - TURKEY

e-mail: ekrem.savas@yahoo.com
e-mail: essavas@ticaret.edu.tr
INFORMATION TO READERS: We encourage readers to sign up for the publishing notification service for this journal. Use the Register link at the top of the homepage for the journal. This registration will result in the reader receiving the Table of Contents by email for each new issue of the journal. This list also allows the journal to claim a certain level of support or readership. See the journal’s Privacy Statement which assures readers that their name and email address will not be used for other purposes.

INFORMATION TO AUTHORS: Interested in submitting to this journal? We recommend that you review the About the Journal page for the journal’s section policies, as well as the Author Guidelines. Authors need to register with the journal prior to submitting, or if already registered can simply log in and begin the 5 step process.

Manuscripts must be submitted in pdf-format through the journal website. Authors are advised to keep a copy of their manuscripts since the publisher cannot accept responsibility for lost copy. Manuscripts received are assumed to be final. If excessive changes are requested after an article has been finalized, the author will be charged with the cost of the alterations other than the printer’s errors.

Authors are advised to study the printing style of the journal before making the final version of the manuscript. Styles files may be requested to the editorial assistant.

Authors of papers accepted for publication will be asked to supply electronic version of the paper together with a paper manuscript. The file should be in L\TeX, and it should be sent by attached mail to the editorial assistant.

INFORMATION FOR LIBRARIANS: We encourage libraries to list this open access journal among its electronic journal holdings. This publishing system is also designed to be hosted and operated by research libraries in support of the journal publishing work of its associated faculty. For more information on using Open Journal Systems to manage and publish peer-reviewed journals,

Contents

L. Odondi, Optimal compliance prediction models for estimating causal effects, ....................................................... 1 – 24.
C. J. Chikunji, Groups of units of commutative completely primary finite rings, ................................................... 40 – 48.
E. Savaş, On some sequence spaces and A- statistical convergence, 64 – 71.

http://imhotep-journal.org/index.php/imhotep/
Imhotep Mathematical Proceedings
Volume 1, Numéro 1, (2014).