MSC TRAINING IN RESEARCH METHODS SUPPORT

D. A. Stern[1], R. Coe[2,3], R. D. Stern[3], and B. M. McDermott[3]
[1]Maseno University, Kisumu, Kenya
[2]ICRAF, Nairobi, Kenya
[3]Statistical Services Centre (SSC), University of Reading, UK
volloholic@hotmail.com

INTRODUCTION

Computers have been being brought into statistics education since the 1960's (Sterling & Pollark, 1966). It is now accepted that there is a need to have a less theoretical approach to statistics education. In Kenya the need to change the way statistics is taught is recognised and accepted (Odhiambo, 2002). Internationally the GAISE (Guidelines for Assessment and Instruction in Statistics Education) reports (Aliaga, et al., 2005) (Franklin, et al., 2005) outline the problems as well as proposing solutions at the school and early college level.

Applied statistics is a key component of research but 'classical' experiments and surveys have now become a relatively small part of the research portfolio in many organizations. This is illustrated by some of the international agricultural research institutes that have replaced their biometry or statistics units with wider service, research support units.

The need for a modernised approach to mathematics education, now that computers have become ubiquitous, is being recognised and parallels the discussions about statistics education. This change in the mathematical subject as a whole broadens the training and reduces the emphasis on calculation. This implies that the changes needed are not making statistics training less mathematical; rather they are leading the way to bring mathematics and statistics training up to date.

One the key advocates for changing the Maths curriculum, Wolfram states "I am not even sure if we should brand this subject as math, but what I am sure of is that this is the mainstream subject of the future." (Wolfram, 2010) (Wolfram, The Practical Approach to Maths Education, 2010). The broadened role of statistics in research is already being rebranded and called 'research methods support'.

STATEMENT OF THE PROBLEM

Research in many application areas is limited by a lack of effective research support (statistics in a broad sense). This problem has long been recognised (Riley, 1998) and the blame is often attributed to over theoretical teaching. In general people agree on the solution which includes transforming statistics by broadening the training to cover the full set of skills needed by a modern applied statistician.

In addition, the skills needed to support good research is broader than any definition of statistics. Even if we were to train a modern applied statistician this would still not be enough to support the current research process.

THE INNOVATION,

The innovation is a modular eLearning MSc programme in Research Methods based at Maseno University in Western Kenya. The materials have drawn on the experience of research support and statistics training courses run by many organizations, including the Statistical Servics Centre (SSC), University of Reading and various institutions in Kenya. It is hoped that those used on the Maseno course can become available as open educational resources, and hence be adapted and used in training courses elsewhere, whether by e-learning, or face-to-face. The individual modules aim to be applicable for researchers and research support staff alike. The modular nature of the program allows for participants to register for a single individual module that would help them in their workplace even if they do not want to pursue the whole MSc program. The program is classified as eLearning and individual modules can be offered either as fully online or with blended teaching while fitting into the same degree. At the moment all modules, whatever their delivery mode, have an end of semester final exam so that they fit in with standard university assessment and quality control procedures.

The degree program is designed to encourage participants to acquire the skills needed to support research. As such, the five core modules cover data management and communication skills as well as the more traditional statistical background:

1. Data management within data flow
2. Exploring, describing and presenting data
3. Statistical modelling
4. Study design
5. Communicating research

Some elective courses can be more specialised to participants areas of interest and include current topics like 'climate change and climatic variability'. There is also the option of customising a module to an employer's needs and offering it as in-service training. In this case, if the customised course content is approved by the Maseno department, it can still count as a component of the MSc.

BACKGROUND AND PHILOSOPHY OF THE CURRICULUM

Kenya has been at the heart of a number of initiatives aimed at improving statistical support particularly for agriculture in the region. The University of Nairobi proposed a new degree in Biometry in the early 1990s. It received support from the Rockefeller foundation; the degree was launched in 1993, and continues to this day. One limitation in the structure of this degree was the insistence by the lecturers, that the graduates possessed the necessary (theoretical) skills to be able to proceed to a traditional PhD in statistics.

A second initiative stemmed from university research in agriculture, also supported by the Rockefeller foundation. During the 1990s this regional project found that the research projects were limited by the students and staff's limitations in statistics. In 2005 it passed responsibility for the research to an organisation called RUFORUM (Regional Universities Forum for Capacity Building in Agriculture) that is based in Kampala, Uganda. RUFORUM has responsibility for improving agricultural research for development in 25 Universities in East and Southern Africa. It has continued to recognize that better biometric skills and awareness are needed to support this improved research and funded an innovative full-time MSc programme in Research Methods, designed to produce useful 'Research Methods Professionals' or biometricians. Training began in 2009 in Jomo Kenyatta University for Science and Technology (JKUAT), and the first cohort of 29 students has now graduated.

The changing nature of research

Agricultural research is changing rapidly. These changes are very obvious to anyone working in the sector and are revealed in strategies and reports of research organisations (for example the CGIAR (CGIAR)) or syntheses (IAASTD, 2009). Trends include:

- Organizational changes, with a blurring of boundaries between research and development.
- Increasing pressure on development organizations for quantitative evidence of impact.
- Changing aims, from productivity to livelihoods, consumer choice and environmental interactions.
- Use of new research paradigms such as participation.
- Looking for integrated solutions that bring together, for example, agriculture and health, farming and conservation, or climate change and land-use.
- New tools such as crop simulation models
- New data types, such as molecular and remote sensed observations.

Scientists working in these newer ways use a huge variety of tools and methods, many of them developed for a narrow field of application. It would seem impossible to train statisticians to become familiar with all the techniques, allowing them to support the research. However, a closer look shows that the methods are built on a few basic data collection designs – surveys and experiments – and these are effective if designed according to some underlying principles. The same is true for data management, analysis and interpretation. If we can ensure that statisticians understand these principles and can apply them to novel problems, then they will be able to support scientists.

The experience among most graduate students in the region is that they are taught a lot of statistical theory, and eventually, during their research phase, concentrate on trying to fit their research problems into these "boxes" of theory. For example, agriculture students are taught ANOVA

with an emphasis on the mathematical specification of the formula and the calculation of the sums of squares and derivative results. Unfortunately, for most student, seeing the formula does not illustrate the idea of pattern and random "noise" and the emphasis on calculations does not help the students draw general concepts that can be applied to similar but not identical analysis like an unbalanced experiment. Those familiar with the plea by Wolfram to change mathematics education (Wolfram, 2010) will recognise that the problem is the same here.

McKenzie (Mckenzie, 2007) stressed that concepts are rarely emphasized in training courses, even when they can be identified. This has to change, with the emphasis moving away from statistical theory and specific techniques, to an understanding of principles and the ability to apply them. Continuing with our example, a research methods approach to teaching modelling that would include ANOVA would start with recognition of modelling as a separation of pattern from noise. Research methods. students would learn to specify many different types of patterns and noise (error) structures. The calculation of estimates of parameters would be left to the computer packages but the student would learn to relate the parameter estimates, and the estimates of their uncertainty back to the model and ultimately to what this says about that piece of the world the model is being used to describe. For the statistics student an assignment in this module would end here but the research methods student might be challenged to communicate these results to three different types of audience.

The Research Methods MSc in principle

It is helpful to think of the requirements of the job we expect graduates from the course to be able to fill. In the JKUAT MSc funded by RUFORUM, these requirements were listed as:
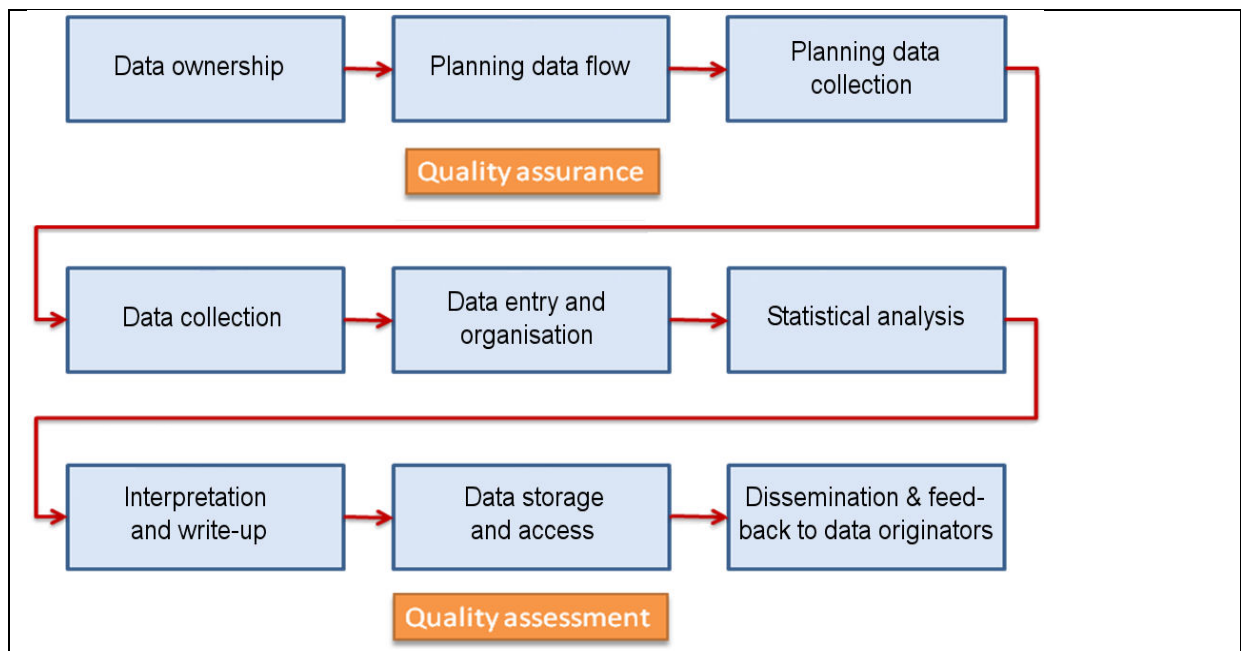
- Evidence of ability and experience in each of the following:
    - Synthesis of scientific information
    - Problem formulation
    - Application of research design principles to real problems
    - Quality control and organisation of data
    - Application of statistical analysis principles to real datasets
    - Summary and interpretation of research results
    - Critical review of proposals and reports
    - Presentation and communication of research results and methods ideas
    - Use of current tools and learning of new tools and approaches
- Ability to work independently with minimum supervision
- Ability to work with multidisciplinary teams

This list of job requirements is broader than the typical requirements of a biometrician or statistician. That is part of the reason we call these people 'Research Methods Professionals' not biometricians or statisticians, and is the reason the course is called 'MSc Research Methods'. The other reason is that scientists, particularly social scientists, often believe biometrics or statistics is too narrow to provide the help they need. 'Research Methods' has broader appeal.

The Maseno courses have been built around these skill requirements and, like the JKUAT course, aims to integrate the four areas of:

a) conception and design,
b) data collection and management,
c) data analysis and
d) presentation and interpretation.

The whole process is termed "data flow", and is summarised in the following figure (Coe, Stern, & Barahona, 2011).

Data ownership → Planning data flow → Planning data collection

Quality assurance

Data collection → Data entry and organisation → Statistical analysis

Interpretation and write-up → Data storage and access → Dissemination & feed-back to data originators

Quality assessment

One message from this diagram is that the objective of research support is to produce high-quality, defensible evidence from the data being collected.

A further element of the research methods degree is that participants should develop skills in a variety of software packages. An electronic textbook, CAST, Computer Assisted Statistics Textbooks (Stirling, 2000), is used for background reading and standard spreadsheet and database packages are used for a number of tasks. Other tasks require the use of statistical packages; sometimes participants have a free choice of package and at other times specific software such as GenStat (VSN International, 2011) and R (R Development Core Team, 2009) are required. Other specialist software have also been explored for specific topics, including Gapminder (Rosling, GapMinder, 2005) for data visualisation and storytelling and CSPro (U.S. Census Bureau, 2000) for data entry. Specialist web services are also explored and discussed, including Drop Box (Dropbox, 2007) for managing a data and document store and Dataverse (IQSS at Harvard University, 2006) for publishing Data.

PEDAGOGY AND NEW TECHNOLOGY

Until recently, statistical work in many developing countries was hampered by the lack of computing and internet access by students and researchers. This has now changed. MSc students in Kenya would expect to have their own laptop and internet access, either through the University, or via modems that use the telephone network. This, in turn, permits the use of a learning Management System throughout the modules, and Moodle 2.1 is being used on the Maseno course.

The modules aim to develop professional skills such as teamwork, communication and synthesis of scientific information, as well as emphasising self-discovery by the participants. Nothing in this approach to teaching is new (Stolovitch, 2002), but it is different from other Kenyan statistics and biometric MSc courses. The eLearning aspect has also helped in reinforcing the student centred approach to learning. The step from blended to full eLearning was taken on the back of previous positive experiences of teaching statistical concepts using eLearning in Africa (Kane, 2007) (Dale, Clark, Stern, Leidi, & Stern, 2010).

Learning management systems are still rarely used on courses in Kenya. The use of computers connected to the internet combined with learning organized on the scaffold of a learning management system revolutionizes the learning environment and tools available to the lecturer. Firstly, all the materials can be organized in such a way that the structure and materials from one offering of the course can be easily retained for the next. Even when a new faculty member takes over a module he or she can concentrate on incremental improvement rather than re-collecting and re-organizing the main course materials. Secondly the faculty member is not limited by his or her areas of expertise or lecture style. Presentations, podcasts and videos on many subjects are available on the web and their number and quality are increasing. This allows the faculty members to move away from concentrating on information delivery (the lecture) and put more time and effort into fostering

problem solving, synthesis and self-discovery by their students. Finally, computers and the web expand the means of putting across ideas. In addition to documents and lectures, the computer permits simulation and animation which can be used by, as well as for, the student.

In the Maseno RM MSc course, the modules are based on a 10 week-semester of study, consisting of 10 topics, each a week long. Topics typically have a short introduction and layout of the tasks required for completion of that topic. Resource materials include documents, podcasts and videos containing, basic information, points of view, examples and challenges. Usually there are some tasks, such as challenge questions or activities, which students are encouraged to attempt and then discuss their attempt on the topic forum. Lastly, there is one or more assessed (marked) assignment which may take the form of a practical task, answers to a number of challenging questions or a report or essay.

One of the challenges in creating a research methods course, whether e-learning or face to face, is the development of the tasks, or assignments, both marked and un-marked. These tasks need to:

- be relevant to the student,
- call on problems solving skills,
- require the understanding of the concepts central to the topic
- and yet be able to be accomplished in the few hours available to the students each week.

Doing the tasks, hopefully in concert with other students and faculty and resource persons, is when the learning takes place. In addition, the module developers must ensure that other skills such as organization, ability to write well and creativity are exercised and assessed over the course of completing a range of module tasks. The development of good module tasks takes trial and error and requires constant refinement as the research methods program matures.

Computers and internet access also help in the development of module tasks. There are materials available, such as the ILRI (International Livestock Research Institute) Biometric and research methods teaching resource case studies (ILRI, 2011), which give backgrounds and data from real research projects and can be used to create effective tasks. With computer access students don't just talk about but actually create databases, maps for a range of purposes, animated graphics and other research products.

DEVELOPMENT PROCESS

The ideas behind the research methods degrees have been being developed since 1990. Since then many things have changed but the need for courses that trains people who can support research in the region has remained. The lessons learned from the previous attempts have led to the development of the Maseno degree. The Maseno course owes much of its structure to the Modular structure of the JKUAT MSc, though this was developed as a full-time face-to-face course.

The JKUAT course took many years to come to fruition and was substantially funded, indirectly by the Gates foundation through RUFORUM. In contrast, the concept of the Maseno MSc was developed in early 2011 with no funding, was passed through the Maseno Senate in mid-2011 and was launched in September 2011.

PILOT AND IMPLEMENTATION RESULTS

In many ways the Maseno MSc has used the experience of the JKUAT MSc as a pilot study. There were a number of lessons to take from the JKUAT experience, that have fed back into the implementation of the Maseno MSc. For example, despite all students in the first year of the JKUAT program having full scholarships a number were unable to get study leave from work to attend the program full time. This was part of the motivation for creating a flexible, modular degree which caters for part time studies. The first cohort of JKUAT students also had a small eLearning component at the beginning of their program which was cited as being one of the most important parts of the degree by the students. This led in part to the decision to try and give the Maseno program largely through eLearning.

The name "research methods" for the degree was questioned by academics in both institutions. Concerns included the fact that students wouldn't apply, and also employers wouldn't recognise the qualification. So far these fears have been unfounded. An advert for this and a range of

other degrees in July 2011 produced the 96 responses for the MSc. Within a month, 34 students had been accepted to the program and paid to enrol on the course. This may partly be because it is an e-learning degree, but could also be that the participants recognise the importance of the shift in subject area. Participants vary from data managers at research organisation to a lawyer of the Kenyan High court who is also a lecturer of law.

Of concern in Maseno, was whether students would actually be able to form effective discussion groups and learn team work given that they had not met face to face and the discussions would not take place in real time. This proved to be unfounded as most of the students took an active role in the forum discussions. At some times they would be asking questions and at other times offering suggestions and relating their own, very relevant, experiences. However, there may be other levels of team work that cannot take place during asynchronous dialogue and the developers of the e-RM will have to experiment to see how this problem can be addressed.

The e-RM MSc modules require considerable time and effort from faculty. Directing a module can easily demand an hour or more a day monitoring and responding to on-line posts and there are weekly tasks/assignments to be graded. While on-line learning management systems, like Moodle, allow automatic grading of multiple choice quizzes, most of the marked tasks/assignments cannot be of this type if the objectives of the program are to be met. The online environment provides both faculty and students with more resources, allows flexible time management but superior effort on the part of both students and faculty is required to achieve the superior results sought. Much of the effectiveness of the Maseno e-RM is due to regular completion of tasks followed by timely feedback coupled with a tutorial like atmosphere on the module forums; both very demanding of faculty time. If the Maseno e-RM is to be scaled up from its present "class" size, the model will need some changes. Adopting some of the practices of universities in North America where graduate students lead tutorials and do some of the marking is one option.

DISCUSSION AND IMPLICATIONS
Few would argue against the idea that modern research needs practitioners with a broad range of skills. The more important questions are: whether it is possible to graduate research methods professionals with the competency and skill range needed to add value to research teams, and whether research organizations' institutional cultures and structures will allow RM professionals to maximize this contribution.

Judith Strother (Strother, 2002) suggests using a four level, progressive metric developed by Kirkpatrick which can be adapted to address these questions. The first level is the individual student's evaluation of the course. The second is the increase in knowledge and skills of the students after taking the course. The third and fourth levels are harder to measure and to separate from confounding factors. These are the change in behaviour or practice of the students after having taken the course and, finally, the change in the effectiveness of the institutes, organizations, communities or other groups due to the training these individuals have received.

Some student evaluation has already been gathered from the first three core modules offered in the e-RM. The best evidence of the relevance of the material included in the first three modules is the feedback from the students themselves since the majority are currently employed in research while they are enrolled in the modules. Seventeen students gave an average rating of 2.0 for relevance of module contents based on a five point scale where 1 was "very good" and 5 "very poor". Thirty seven out of thirty eight students responding agreed with the statement that they would find the module content useful even if it did not count toward a degree. Some individual comments, that cut across these evaluation levels, emphasised the value of the skills being mastered:

"For your information, I am now able to select samples using the unequal probability method in GenStat. I don't regret enrolling on this programme. I am acquiring awesome skills steadily!"

"Am appreciating Excel day by day as we progress, it is such an amazing program which can do so much.

I found this quote very applicable in this course; "I hear I forget, I see I remember, I do and I understand"....Confucius"

Addressing all but the first level of evaluation of the training of RM professionals will need some external assessment of the training, assessment and marking of tasks and exams in the e-RM program. It would be ideal to have some managers from public and private institutions determine if

the skills they are looking for in their professional staff are developed at a sufficiently high level in the e-RM modules.

This is also going to challenge e-RM faculty to work hard to stay abreast of new developments in research practice and continue to develop and improve module resources, dialogues, tasks and assessment.

At particularly the third level of evaluation, all e-RM MSc candidates are required to do a research project or thesis. Evaluation of these projects provides a good opportunity to see how spontaneously the students employ the knowledge and skills emphasized in their module work.

Students trained under the Research Methods paradigm from JKUAT are just starting to enter the job market and take up research support positions. Hopefully, there will be some follow up on their success in finding employment in the research industry and their performance once employed.

At the third and fourth levels, one part of Kenyan institutional culture which might reduce the effectiveness of RM professionals in research teams is the perceived importance of a PhD for research professionals. It may be difficult for RM professionals with an MSc to be taken seriously when partnered with a team consisting of mostly doctoral graduates. For the degree to spread and become sustainable it will be important to devise doctoral level programs that reinforce the research methods ideas.

One benefit of considering the evaluation of a programme as it is being developed is that it can influence the emphasis of the training within the programme. In many African countries there are some "low fruits to be harvested" and it is important that the students are able to contribute fully in such areas. For example the entry and organisation of the data is often a stumbling block, and this is relatively easy to change. In addition, many researchers lack the confidence to even start the analysis of their data.

These "low fruit" are covered in two of the compulsory modules. Following the dataflow module, students should be able to support efficient data entry, and also be able to re-organise data that has been entered in a way that is not yet ready for analysis. The descriptive statistics module should enable the students to support the initial analyses of the data, however complex in structure, (e.g. including data at multiple levels), in ways that correspond to the objectives of the research. Many researchers feel "liberated" when they see what is possible so simply.

Finally, the rebranding and broadening of applied statistics as research methods does seem to be working in Kenya. It has helped in two ways. Those who recognise they need to do better research, but are afraid of statistics, seem to be receptive to the statistical concepts presented as tools to aid research. It has also helped that theoretical statisticians do not see a threat from these new research methods degrees and support them as a way they think they can continue their own approach as before. The moves in mathematics by Wolfram and others, e.g. (Wolfram, 2010), indicate they may not be left in peace much longer.

SOURCES AND REFERENCES

Aliaga, m., Cobb, g., Cuff, c., Garfield, j., Gould, r., Lock, r., et al. (2005). *Guidelines for Assessment and Instruction in Statistics Education Colledge Report.* American Statistical Association.

CGIAR. (n.d.). *CGIAR Home page.* (CGIAR) Retrieved from CGIAR: http://www.cgiar.org/

Coe, R., Stern, R., & Barahona, C. (2011). *Data Flow: Organising Action on Research Methods and Data Management.* Reading UK: Statistical Services Centre.

Dale, I., Clark, C., Stern, R., Leidi, S., & Stern, D. (2010). E-learning of statistics in Africa. *Proceedings of International Conference on Teaching Statistics 8.*

Duflo, E., & Michael, K. (2003). Use of Randomization in the Evaluation of Development Effectiveness. Washington: World Bank Operations Evaluation Department.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., et al. (2005). *Guidelines for Assessment and Instruction in Statistics Education Report.* American Statistics Association.

IAASTD. (2009). *Global report: Agriculture at the cross-roads.* Washington DC: Island Press.

ILRI (International Livestock Research Institute). (2011, November). *Biometrics and Research Methods Teaching Resource.* Retrieved February 2012, from ILRI: http://ilri.org/biometrics/

Innovation in Poverty Action. (2011, 1 10). *Home page.* (Innovation in Poverty Action) Retrieved 1 10, 2011, from IPA.

Kane, A. (2007). *Is E-learning an effectivemeans of training Delivery Within Africa? A case study of Africa Meteoroligists.* University of Leicester.

Kurji, P., McDermott, B., Stern, D., & Stern, R. (2010). The growing role on computers for teaching statistics in Kenya. *ICOTS 8* .

Leadbeater, C. (2010). *Charles Leadbeater: Education innovation in the slums.* Retrieved 1 18, 2011, from TED: http://www.ted.com/talks/charles_leadbeater_on_education.html

Leadbeater, C., & Wong, A. (2010). *Learning from the Extremes.* Cisco Systems, Inc.

Mckenzie, J. J. (2007). Exercises and questions for electronic assessment of statistical concepts. *ISI 56th Session.*

Odhiambo, J. W. (2002). Teaching of Statistics in Kenya . *Proceedings of the 6th International Conference on Teaching Statistics.* Cape Town.

OpenCourseWare Consortium. (2002, October). *OpenCourseWare Consortium home page.* Retrieved October 5, 2011, from OpenCourseWare : http://www.ocwconsortium.org/

R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation.

Riley, J. (1998). *Strengthening Biometry and Statistics in Agricultural Research.* Wageningen: CTA.

Rosling, H. (2005). *GapMinder.* Retrieved 3 31, 2011, from http://www.gapminder.org/

Hillman, D. (Director). (2011). *Joy of Stats* [Motion Picture]. UK.

Sterling, T. D., & Pollark, S. V. (1966). Use of Computers to Teach Introductory Statistics. *Communications of the Association of Computing Machinery* (9), 274-276.

Stern, D., Ongati, O., Agure, j., & Ogange, B. (2010). Incremental modernisation of statistics teaching and curriculum at Maseno University, Kenya. *Proceedings of International Conference on Teaching Statistics 8.*

Stirling, D. (2005). *CAST for Africa.* Retrieved August 23, 2010, from CAST: http://cast.massey.ac.nz/collection_african.html

Stirling, D. (2010). Improving lectures with CAST applets. *ICOTS 8.* Ljubljana: International Association of Statistical Education (IASE).

Stirling, D. (2000). *Public CAST.* Retrieved August 23, 2010, from CAST: http://cast.massey.ac.nz/

Stolovitch, H. D. (2002). *Telling ain't training.* Alexandria: American Society for Training and Development.

Strother, J. (2002). An Assessment of the Effectiveness of e-learning in Corporate Training Programs. *The International Review of Research in Open and Distance Learning , 3* (1).

Tableau Software. (2010). *Tableau Public.* Retrieved October 5, 2011, from Tableau: http://www.tableausoftware.com/public

VSN International. (2011). *GenStat for Windows 14th Edition.* VSN International, Hemel Hempstead, UK.

Wolfram, C. (2010, November). *Conrad Wolfram: Teaching kids real math with computers.* Retrieved January 18, 2011, from TED Talks: http://www.ted.com/talks/conrad_wolfram_teaching_kids_real_math_with_computers.html

Wolfram, C. (2010). *The Practical Approach to Maths Education.* Retrieved March 29, 2011, from Wolfram Research: http://www.wolfram.com/solutions/highered/usformat.pdf