

# eVALUAtion Matters

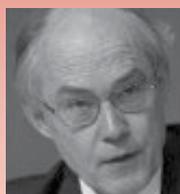
A Quarterly Knowledge Publication on Development Evaluation

Second Quarter 2015

## IMPACT EVALUATION: Insights from Practitioners



*Emmanuel Jimenez*



*Richard Manning*



*Annette N. Brown*



*Lycia Lima*



*André Portela Souza*



*Dalila Figueiredo*



*Elsa de Moraes Sarmiento*



*Christian Kingombe*



*Laura Camfield*



*Maren Duvendack*



*Richard Palmer-Jones*



*Jean-Louis Arcand*



*Jyotsna Puri*



*Cameron Breslin*



*Alexandra Orsola-Vidal*



*Temina Madon*



*Jacob Oduor*



*Laurence Lannes*



*Fabrizio Felloni*



*Simona Somma*



*Nompumelelo Mohohlwane*



*Stephen Taylor*



*Ricardo Paes Mamede*



*Daniela Cruz*



*Teresa Farinha Fernandes*



*Jeremy D. Foltz*



*Kweku Opoku-Agyemang*



*Peter Huisman*



*Rik Linsen*



*Anne Oudes*



AFRICAN DEVELOPMENT BANK GROUP  
GROUPE DE LA BANQUE AFRICAINE  
DE DEVELOPPEMENT



IDEV

Independent Development Evaluation  
African Development Bank

[idev.afdb.org](http://idev.afdb.org)



*Nompumelelo Mohohwane and Stephen Taylor*

# Using Impact Evaluation for Education Policy Innovations: *the Case of Early Grade Literacy in South Africa*

20 March 2015

### ***Background: Development through quality education***

South African and international development planning is increasingly centered around education. Agreements such as the Millennium Development Goals (MDGs) and the Education For All (EFA) drive are testament to the international recognition of the pivotal role of education. In South Africa, the National Development Plan, which is arguably the country's most prominent planning document, affords a central role to education, while the Presidency officially regards improved quality of basic education as the country's number 1 priority (NPC, 2012).

Despite substantial progress in expanding access to schooling in developing countries over the past few decades, there is now a growing recognition that in many countries the learning outcomes achieved by those attending school are often dismally poor. Spaul and Taylor (2015), for example, demonstrate that despite improved access to schooling in many Southern and East African countries there are large proportions of children who reach grade 6 without having acquired basic literacy and numeracy skills. This is important since there is clear evidence that the quality of skills achieved (over and above the quantity of schooling attained) has a significant impact on economic growth and on the labour market prospects of individuals (Hanushek & Woessmann, 2007). While there are numerous social, psychological and other benefits of education, the quality of learning outcomes should be seriously considered when analyzing education system performance.

### ***The education quality challenge in South Africa***

Local and international assessments of learner performance consistently indicate poor performance by South African learners in Mathematics, Science, and Languages (Spaul, 2014). The low performance across the board is concerning but for the purposes of this paper the focus will be on literacy and reading.

The Pre-PIRLS 2011 results indicated that 29 per cent of South African Grade 4 learners did not have the rudimentary reading skills required at a Grade 2 level. The situation was most severe for those learning in an African language. For example, 57 per cent of learners that took tests in Sepedi or Tshivenda did not reach this level (University of Pretoria, 2012). The 2006 PIRLS study, which tested grade 5 South African learners on a somewhat more advanced reading test, showed that approximately 80 per cent of children had not learned to read for meaning by grade 5, but at best could only extract basic factual details from a text. If children have not learned to read fluently by this time, it stands to common sense that they will not be able to cope with the requirements of the curriculum at higher grades. Weak reading foundations are therefore at the heart of the education quality challenge in South Africa.

South Africa's underperformance, even relative to many poorer countries, prompts an enquiry into the causes behind the limited ability of the education system to convert inputs into outcomes. South Africa's per pupil expenditure in schools in purchasing power parity (PPP) terms exceeded that of all the other thirteen countries in SACMEQ with the exception of Seychelles. An important part of the contextual background is the inequality resulting from the political history of South Africa. The changes in the education system following the end of Apartheid and the establishment of a new democratic state in 1994 have had

limited success in changing the nature of schooling. Several authors, such as Fleisch (2008) and Spaul (2014), argue that South African schooling has still effectively got two systems in one. The first and largest part of the system comprises the historically disadvantaged schools and is characterized by inefficiency including poor school management, continuous underperformance, high and indiscriminate grade repetition and dropout. The second system refers to historically white and Indian schools where learner performance is at a higher standard, parents make substantial fee contributions, organizational and instructional processes are more efficient and schools are well endowed with infrastructure.

The South African government is well aware of these challenges and continues to allocate the largest share of government expenditure to education. Since the early 1990s education spending has become increasingly well targeted to poor schools (Gustafsson and Patel, 2008). Specific initiatives and policies implemented by the South African government to address equity challenges in education include the introduction of no-fee schooling which is implemented in about 77 per cent of public schools, and the provision of daily meals through the National School Nutrition Program to approximately 70 per cent of schools focusing on the poorest schools (Department of Basic Education, 2014).

Despite these considerable efforts, however, learning outcomes remain low in South Africa and little is known about the effectiveness of particular policies and programs designed to improve learning. Where evidence is available it is often self-reported, focused on inputs, anecdotal or part of a larger initiative where the effect of specific efforts is difficult to isolate. This motivates for an agenda of impact evaluation to inform policy-making going forward, as will be argued below.

### *The importance of early literacy learning in South Africa*

Literature on the evidence of early learning emphasizes the importance of mastering certain learning foundations for the sake all further learning. The

literature refers to ‘self-productivity’, explaining that skills acquired during one period generally persist into the next period and may make the acquisition of other skills in another dimension easier (Girdwood, 2013).

In addition to the argument for the cognitive benefits of the development of good educational foundations and their lasting effects, James Heckman (2007), amongst others, contends that intervening earlier rather than later is more cost-effective. The costs of providing curriculum support for areas of learner deficits identified early, such as in the Foundation Phase are expectantly lower than mediating learning later in schooling where the gap between curriculum expectations and learner knowledge may be excessively large in a multitude of subjects, as Pritchett and Beatty (2015) have shown. The costs accrued at later stages include high rates of grade repetition and dropping out of the education system.

One critical learning foundation that needs to be acquired during the early grades of primary schooling is reading. A large theoretical literature points to the benefits of learning to read in the home (or first) language. One of the expected benefits is that second language acquisition should be easier once a firm grasp of the nature of reading and literacy has been attained in one language. A paper by Taylor and Coetzee (2013) provides some empirical evidence from South Africa that home language instruction in grades 1 to 3 caused improved English literacy in grades 4 to 6 compared with children who were taught in English as the language of instruction. This finding substantiates the argument that all learning builds on prior learning; as such mastery of a second language is enabled by the mastery of the first language. This points to the strategic value of finding ways to improve home language reading acquisition in the Foundation Phase. Yet, the reality is that the majority of children will experience a transition to English as the language of instruction in the fourth grade. Finding ways to strengthen English vocabulary and manage this transition most effectively will therefore also be important.

## *Education Policy Development and the Evaluation Process*

A detailed review of South African education policy development is beyond the scope of this paper. What is clear, however, is that despite many policy changes and new programs, little is known about the ultimate impact of these initiatives on learning outcomes. The lack of a focus on impact evaluation is not unique to South Africa, as the following quote illustrates:

“Development programs and policies are typically designed to change outcomes, for example, to raise incomes, to improve learning, or to reduce illness. Whether or not these changes are actually achieved is a crucial public policy question but one that is not often examined. More commonly, program managers and policy makers focus on controlling and measuring the inputs and immediate outputs of a program—how much money is spent, how many textbooks are distributed—rather than on assessing whether programs have achieved their intended goals of improving well-being” (World Bank, 2010).

In cases where extensive research is done – at least in South Africa – it is typically focused on diagnosing areas requiring attention rather than evaluating possible solutions. Where interventions are evaluated it is often through conducting case studies or piloting in a small number of schools. The shortcoming of this approach is that the implementation model often used in case studies or small-scale pilots is often resource intensive and may be difficult to replicate at a larger scale.

A focus on evaluation is now emerging within the South African government through the introduction of the National Evaluation Policy Framework in 2011. This policy framework includes a National Evaluation Plan (NEP) which commissions independent evaluations of priority government programs in a partnership between the custodian department and the Department of Planning, Monitoring and Evaluation (DPME, 2014). Several Department of Basic Education (DBE) programs have been evaluated through the NEP, namely the Grade R

program, the Funza Lushaka Bursary Program and the National School Nutrition Program.

The evaluations referred to above are all retrospective evaluations, assessing how well programs were implemented or if the intended program goals were attained. Prospective impact evaluations, where programs are evaluated prior to being taken to scale, remain extremely rare. One exception to this is the impact evaluation of a new set of study guides introduced by the DBE in 2012 (Department of Basic Education, 2013).

## *Using Randomised Control Trials in Education*

The major challenge in impact evaluation is the need to identify a counterfactual – what would have happened to program recipients in the absence of the intervention? Since one can never actually observe a counterfactual to reality, one needs to use a “control group” or “comparison group” to provide a valid estimate of the counterfactual. Simply comparing recipients with non-recipients or pre- and post-outcomes amongst recipients is usually not likely to provide a valid estimate of the counterfactual since recipients are usually systematically different to non-recipients and outcomes would change over time in any event.

While various quantitative impact evaluation methods are available, the cleanest method for identifying an internally valid estimate of the counterfactual is obtained through conducting a Randomised Controlled Trial (RCT). Through using a lottery to allocate participants to intervention and control groups, an RCT constructs a credible “counterfactual” scenario – what would have happened to those who received an intervention had they not received that intervention.

Prospective impact evaluations also have the advantage for research of uncovering knowledge of the binding constraints in the school system. In complex environments, such as education, there are multiple factors influencing outcomes and it is not always clear which factors to address first. For example,

high quality teaching requires both competent and motivated teachers, though it is not clear which of these is the more binding constraint in South Africa. Teacher knowledge in South Africa is weak: Carnoy et al (2011) found that grade 6 teachers recorded an average score of around 40 per cent on a test designed to assess their mathematics knowledge for that grade. Furthermore, studies show low teacher motivation in poor schools, manifested in high absentee rates and low teaching activity (Reddy et al, 2010). Yet, it is unclear whether to address teacher capacity or teacher motivation first. The lack of rigorous evaluations to establish which of these challenges to address first is a shortcoming of conventional policy and program development.

### **Practical considerations when implementing an RCT**

Statistical expertise is required in the design of an RCT. This involves calculating the required sample sizes in each intervention and control group and conducting the random assignment. For practical reasons, when conducting an education RCT it is often necessary to assign schools as a whole to intervention or control groups, as opposed to assigning individuals to the different experimental groups. This leads to rather large required samples, which has cost implications. The need to raise funds takes time and requires significant stakeholder engagement and government support to convince donors to be involved.

There are two main components to an education RCT – there is the *implementation* of the new interventions and there is the *evaluation* of their impact. The evaluation side of the RCT involves the collection of outcomes data as well as contextual data for the sake of measuring changes in intermediate outcomes and identifying factors that mediate the impact of the intervention. Both the implementation and the evaluation components require financing and should be conducted by separate organizations. In some cases, an NGO or a government department may fund and implement the interventions, thus reducing the need for additional fund raising. If reliable outcomes data already exist, through for

instance a nationally standardized examination, then one might be able to significantly reduce costs associated with the evaluation side of the project.

Both academic researchers and implementing agencies face various perverse incentives when considering or conducting evaluations. A publication bias exists in academia where it is more likely to see studies with positive results published than studies showing no impact (Duflo, Glennerster and Kremer, 2006). RCTs are less prone to this bias, since the large investment of time and resources together with the high reliability of the results mean that even evaluations showing zero impact are likely to be published. Government departments and NGOs may resist evaluations due to the risk of negative findings. Therefore, prospective evaluations of alternative programs or variations of programs under consideration may be more amenable to policy makers and program managers who will then not feel that their entire work for several years is being judged.

For these reasons, the success of any RCT is dependent on extensive stakeholder consultation and support. This ensures that funds and other resources including personnel are available; the integrity of the research design is upheld; the implementation of the interventions is conducted properly, and the findings are considered for program or policy scale-up or redesign.

### ***Experiences from two new reading evaluations in SA***

Through a developing partnership between education researchers, government and donors, two studies are being undertaken to evaluate possible ways to improve reading acquisition in South African schools.

### **Evaluation of a remedial reading program in Grade 4**

The Gauteng Primary Language and Mathematics Strategy (GPLMS) implemented between 2011 and 2014, included various new interventions focused on the early grades. As part of this, a Reading Catch-Up Program (RCUP) was developed to strengthen the

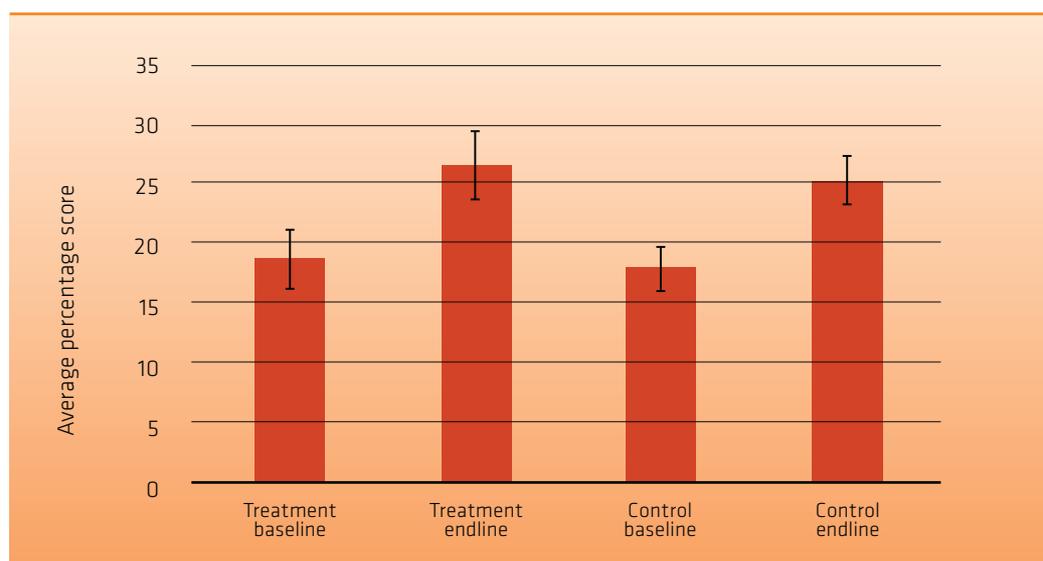
English skills of children in grades 4 and 5 whose first language was not English but who are required to learn using English as the language of instruction in those grades. A preliminary evaluation of the program indicated large gains in the language skills of program recipients over time (Hellman, 2012). However, there was no control group. Therefore, some strong assumptions had to be made about how much learning would have taken place over the period had there been no intervention.

These initially promising, though inconclusive, results prompted an RCT of the RCUP to be conducted in the district of Pinetown in the KwaZulu-Natal province in 2014 (Fleisch, Taylor, Schöer, and Mabogoane, 2015). The intervention lasted for 11 weeks and consisted of on-site teacher support by reading coaches and the provision of scripted lesson plans and additional graded reading books. The RCUP targeted Grade 4 learners in schools that transition to English as the language of instruction after using the home language in the Foundation Phase. The hypothesis underlying the program was that the learning gaps in learner mastery of English at the end of the second term in Grade 4 may be caught-up through the provision of a well-designed relatively short intervention. The program was implemented

in 40 intervention schools with a control group of 60 schools. Assignment to intervention and control group was done through a computerized lottery. Different organizations were contracted to conduct the intervention and the data collection for evaluation, and the evaluation agent was kept blind to which schools were in the intervention group versus the control group.

The most notable finding of the study was that although learners in intervention (“treatment”) schools improved their test scores between the baseline and the endline assessment, the learners in control schools improved by a similar margin, as depicted in Figure 1. This illustrates the importance of obtaining an estimate of the counterfactual: in the absence of a randomly selected control group a false positive result would have been obtained. The main finding, then, is that the RCUP intervention had no statistically significant impact on the overall reading achievement of learners. However, treatment schools improved more than control schools in the spelling and grammar subcomponents of the test. The program impact was larger for learners who initially had a basic minimum of English skills and for those whose teachers participated actively in the program.

**Figure 1** Average Pre- and Post-Scores for Intervention and Control Schools



Note: 95% confidence intervals are indicated

Source: Fleisch, Taylor, Schöer, and Mabogoane, 2015

The findings from the RCUP RCT yield several important policy lessons and raise several questions for further research. Firstly, it is evident that ostensibly well-designed programs may not have as large an effect as one might expect in the absence of a rigorous evaluation. Apart from the design quality of a program and the integrity of implementation, there may be contextual factors pertaining to learners, schools and communities which either preclude or are conducive to the effectiveness of an intervention.

Secondly, the findings indicate that the RCUP cannot yet be implemented on a wider scale with any confidence that it will have a significant impact on learning outcomes. This does not necessarily mean that reading remedial programs in general, or even the RCUP specifically, should not be further explored. It does, however, mean that a revised version should first be experimented with and shown to work before government should consider a larger scale implementation.

Thirdly, the independently administered tests indicated that the learning deficits existing by the end of the Foundation Phase are apparently much larger than expected. The finding that initially better-performing learners gained more from the intervention may imply that the program would have been more appropriate at the grade 5 level, even though it was covering topics that should have been covered in the Foundation Phase. Another possibility is that an 11-week intervention is simply too short a time to deeply influence classroom practice and learning.

### **The Early Grade Reading Study (EGRS)**

The second RCT which is being conducted by the Department of Basic Education (DBE) is the Early Grade Reading Study (EGRS). This RCT has recently (February, 2015) commenced in 230 schools in the North West province and has been crafted to evaluate three competing interventions all aimed at improving home language (Setswana) reading acquisition in grades 1 and 2.

The sample of schools includes non-fee paying schools in the Dr Kenneth Kaunda and Ngaka

Modiri Molema districts of the North West province. All schools selected use Setswana as the language of instruction in the Foundation Phase. All three interventions will occur over a two-year period working with the cohort of children entering grade 1 in 2015. The RCT will evaluate the causal impacts of three interventions: (i) a teacher training course focused specifically on the teaching of Setswana reading and literacy, accompanied by scripted lesson plans and graded reading materials; (ii) an on-site support program to teachers from reading coaches, accompanied by scripted lesson plans and graded reading materials; (iii) and a package designed to improve parent involvement in – and monitoring of – learning to read. Each intervention will be implemented in 50 schools within the sample. A further 80 schools have been selected as the comparison group. As was the case in the RCUP study, separate organizations have been contracted to undertake the implementation of interventions and the data collection for evaluation, with the evaluation agent blind to which group schools fit into.

This study is expected to shed light on several research and policy questions. Firstly, it will show which of three alternative interventions is most cost-effective. Although each intervention has a different unit cost, the improvement in test scores per Rand spent for each intervention will be calculated. The evaluation will also investigate whether the impacts of interventions are different for various sub-groups of learners or schools. This will inform the most appropriate targeting of interventions if scaled up. The study is also designed to look at long-term effects and spillover benefits of faster reading acquisition. Do the impacts of the interventions persist, dissipate or compound over time? If one succeeds in improving the acquisition of home language reading in the early grades, are there spillover benefits into other learning areas such as Numeracy and First Additional Language? This will be measured using results of the Annual National Assessments (ANA) in subsequent years.

### **Conclusion**

This paper has demonstrated that improving the acquisition of reading in the early grades is central to the education quality challenge in developing countries, and especially in South Africa. In the absence of solid evidence of effective policies and programs to address this challenge, the use of prospective impact evaluations is recommended. The paper has pointed out the necessity for innovation and rigor to establish such evidence and understanding the binding constraints in the complex South African education system.

These points are substantiated through a description of two recent RCTs focusing on early grade reading in South Africa. The RCT of the Reading Catch-up Program, implemented in Pinetown, Kwa-Zulu Natal has provided important lessons. The findings have highlighted the need for a valid counter-factual in measuring impact, which is a strength of the RCT methodology. The second RCT discussed, the Early Grade Reading Study (EGRS), which is being implemented in 230 schools in the North West province provides an exciting opportunity for further learning. It is anticipated that the findings will address some of the questions emerging from the RCUP study as well as provide substantive information on the binding constraints in the teaching of language in South African schools.

## References

- Boiseiere, M. (2004). *“Determinants of Primary Education Outcomes in Developing Countries”*. Washington: The World Bank.
- Carnoy, M. Chisholm, L. Addy, N. Arends, F. Baloyi, H. Irving, M. Raab, E. et al. (2011). *The process of learning in South Africa: the quality of mathematics teaching in North West Province*. Technical Report. Stanford University School of Education and HSRC.
- Department of Basic Education. (2014). *Annual Report 2013/2014*. Pretoria. Retrieved 15 May 2013 from [www.education.gov.za](http://www.education.gov.za).
- Department of Basic Education (2013). Technical Report. *Developing and Evaluating the First Phase of the Grade 12 Mind the Gap Study Guide Series*. Pretoria.
- Department of Basic Education. (2012). *Summary Report: Evidence Based Policy Review Workshop*. Pretoria.
- Department of Basic Education. (2014). Terms of Reference. *Implementation of Reading Interventions for the research project: “Early Grade Reading Study*. Pretoria.
- Department of Planning, Monitoring and Evaluation. (2014). *Concept for the National Evaluation Plan 2015/16*. Pretoria.
- Duflo, E., Glennerster, R., & Kremer, M. (2007). *Using Randomization in Development Economics Research: A Toolkit*. Center for Economic Policy Research.
- Fleisch, B. (2008). *Primary Education in Crisis: Why South African schoolchildren underachieve in reading and mathematics*. Cape Town: Juta.
- Fleisch, B. Taylor, S. Schöer, V. & Mabogoane, T. (2015). *Assessing the impact of the RCUP: The findings of the impact evaluation of the Reading Catch-Up Program*. Paper Presented at the Wits School of Education Research Seminar.
- Girdwood, E. (2013). *Exploring the Evidence on Early Learning: A Literature Review*. Impact Evaluation of the Introduction of Grade R on Learning Outcomes. Research on Socio-Economic Policy (RESEP). Department of Economics, University Of Stellenbosch.
- Gustafsson, M. & Patel, F. (2008). *Managing the teacher pay system*. What the local and international data are telling us. Stellenbosch Economic Working Papers. 26(06).
- Hanushek, E. A & Woessmann, L. (2007). *Education quality and economic Growth*. The World Bank: Washington.

- Heckman, J. (2007). *The economics, technology, and neuroscience of human capability formation*. NBER Working paper 13195.
- Hellman, L. (2012). GPLMS Intersen Catch-up Program: Analysis of Results. Memo.
- National Planning Commission. (2012). *National Development Plan: Vision for 2030*. Retrieved January 18, 2013, from <http://www.npconline.co.za/>.
- Pritchett, L. & Beatty, A. (2015). Slow down, you're going too fast: Matching curricula to student skill levels. *International Journal of Educational Development*. 40 (1), 276–288.
- Reddy, V., Prinsloo, C., Netshitingani, T., Moletsane, R., Juna, A., & Janse van Rensburg, D. (2010). *An investigation into educator leave in the South African ordinary public schooling system*.
- Spaull, N. (2014) *Education Quality In South Africa And Sub-saharan Africa: An Economic Approach* (Unpublished doctoral thesis, University of Stellenbosch, Stellenbosch).
- Spaull, N., & Taylor, S. (2015). Access to What? Creating a Composite Measure of Educational Quantity and Educational Quality for 11 African Countries. *Comparative Education Review*, 59 (1), 133-165.
- Taylor, S. & Coetzee, M. (2013). *Estimating the impact of language of instruction in South African primary schools: A fixed effects approach*. Stellenbosch Economic Working Papers: 21/13.
- University of Pretoria. (2012). *PIRLS 2011: Media Release*. 11 December 2012. Retrieved 12 December 2012. from [www.up.ac.za](http://www.up.ac.za).
- World Bank. (2010) *Impact Evaluations in Practice*. Retrieved 10 October 2014 from <http://www.worldbank.org/ieinpractice>.

## PROFILE OF THE AUTHORS

Nompumelelo Mohohlwane is an Assistant Director in the Research Coordination, Monitoring and Evaluation Directorate at the South African Department of Basic Education. The unit is responsible for system monitoring, supporting performance information management, and conducting research and evaluation of education interventions. She is currently completing her Masters in Education at the University of Witwatersrand.

Affiliations: Department of Basic Education

Stephen Taylor is an Advisor in the South African Department of Basic Education. His work includes impact evaluation of education interventions, measuring educational performance and equity in educational outcomes. In 2010 he completed a PhD in economics at the University of Stellenbosch, analysing educational outcomes of poor South African children.

Affiliations: Department of Basic Education Research Associate of the Department of Economics, University of Stellenbosch