

## **“Beyond Gut Feeling” Report on the arepp:Theatre for Life Evaluation Project 2002-2004 - Executive Summary**

This evaluation project aimed to move beyond “gut feeling” to a more rigorous and comprehensive evaluation of the impact of arepp:Theatre for life in increasing a sense of self-efficacy in the school-going youth with whom it works. It was conducted by Nell and Shapiro, evaluation consultants. Beginning with 7 780 learners (this dropped to 3 902), the same group of learners were monitored over a period of three years (from 2002 – 2004), during their Grade 10, 11 and 12 years. This was done by means of a survey questionnaire, based on indicators of a sense of self-efficacy, administered before the initial arepp:Theatre for Life intervention, before each of two subsequent interventions, and a few months after the final intervention (there were three interventions in all). Results were cross-tabulated against certain variables in learner and school profiles.

The changes on all the indicators across the total arepp:Theatre for Life original cohort for the three year period were statistically significant at a 95% confidence level. These results were triangulated against a control group and against learner focus group input and educator interviews. Although it is impossible to isolate the attribution of impact in a complex and interactive social context, the results were encouraging and there was clear evidence that the arepp:Theatre for Life interventions do increase a sense of self-efficacy. In addition, the cross-tabulations and qualitative input provided invaluable information about context and learner differences that will enable arepp:Theatre for Life to improve its presentations and the way in which it embeds its work at the local level.

While it may never be possible for organisations working in the social sphere and dealing with the realities of human interaction and a multitude of interventions, to say conclusively: “if not  $x$  then not  $y$ ” where  $x$  is the intervention and  $y$  is the impact. The study was hoping to show from analysis of the data that “if  $x$  then  $y$ ”, where  $x$  was the arepp:Theatre for Life intervention and  $y$  was the desired impact of ‘an increased sense of self-efficacy’.

In addition, we wanted to be able to say “if  $x+a+b+c$  then better  $y$ ”, where  $a$ ,  $b$  and  $c$  were different variables or conditions in which  $x$  was offered. This would firstly enable us to confirm that arepp:Theatre for Life was not wasting its time nor the donors’ money and, secondly, it would enable arepp:Theatre for Life to adjust the product to take advantage of different facilitating circumstances, or to address the disadvantages of other hindering circumstances.

The learnings from this analysis are not only about change and impact, they are also about where arepp:Theatre for Life has most success, where the presentations are not necessarily having the desired impact and in what sort of context the organisation is most likely to achieve its goals. This makes the usefulness of the whole process far greater than it would have been had it shown only that there was a significant change on all the indicators (which it does, in fact, do).

For the purposes of the study, five statements for each of the five arepp:Theatre for Life indicators of impact were created on which learners could rate the extent of their agreement or disagreement. It is worth noting that one of the most exciting spin-offs of the evaluation project has been the further and clearer refinement of what arepp:Theatre for Life defines as impact. This has led to a necessary clarification of the indicators and a review of appropriate statements which have already informed the work of arepp:Theatre for Life and will certainly inform further efforts in evaluation.

Questions about the reliability of the statements grouped under each indicator to measure that indicator were raised by the reliability analysis done on the responses to the statements. Here we are concerned with whether the statements grouped under each indicator as a

measurement of the indicator, do, in fact, measure it. The analysis showed that there was adequate internal consistency in the responses on the indicators of “control” and “self-esteem/self-awareness”. It showed that there was some consistency on “constructive relationships” and “options and consequences”, but not sufficient to warrant confidence in terms of reliability that all the questions related to these indicators. On the indicator of “tolerance” the correlation between the responses to the five statements was weak. What this means is that the statements grouped under the indicators did not necessarily all measure the same indicator (tolerance), although the statements themselves may have remained (and largely did remain) reliable indicators of “things arepp:Theatre for Life wanted to measure”.

Overall, we think that the plays presented by arepp:Theatre for Life deal with the issues with which arepp:Theatre for Life is concerned, and that the statements remain valid indicators of what it is trying to achieve, and what it was trying to measure, even when they are not necessarily grouped correctly. We take this grouping issue into account in interpreting the analysis of the data.

### **Results across all the indicators**

The changes on all the indicators across the total arepp:Theatre for Life original cohort for the three year period were statistically significant at a 95% confidence level. The biggest change overall occurred on the indicator “constructive relationships”, followed by “self-esteem and self-awareness”, then “options and consequences,” with “control” and “tolerance” in fourth and fifth places respectively.

Although the arepp:Theatre for Life evaluation project group was stronger on all the indicators than was the control group, the difference was only significant on “tolerance”, “options and consequences” and “self-esteem/self-awareness”.

We then compared learners from the control group who had had no additional input on sexuality and related issues with learners in the arepp:Theatre for Life evaluation project group who had had no additional input, other than two arepp:Theatre for Life presentations.<sup>1</sup> In this instance, the control group learners were far weaker on all the indicators than were the learners in the arepp:Theatre for Life sample. This was one of the clearest indications we had that “if x (arepp input) then y (impact). It was also an indication that changes within the arepp:Theatre for Life sample were not simply a result of maturation.

Even when control group learners came from the same schools as the arepp:Theatre for Life sample, they tended to have weaker scores than those in the evaluation cohort on key questions.<sup>2</sup> This bore out the finding from the arepp:Theatre for Life evaluation cohort that previous contact between arepp:Theatre for Life and a school (as opposed to that specific cohort) did not make any difference to the impact of the intervention. In other words, any impact is specific to those who experience the presentation and does not go beyond them into the school as a whole.

The findings of an overall difference between the arepp:Theatre for Life cohort and the control group were confirmed to some degree by the focus groups. Those in the arepp:Theatre for Life focus groups were more likely to argue and discuss issues among themselves. So, for

---

<sup>1</sup> We were comparing the third intake of survey questionnaires (done in the first term of 2004 with the Grade 12s) with the first term Grade 12 control group results.

<sup>2</sup> For example: “When I am sexually active, it is my responsibility to protect others from diseases such as HIV/AIDS”, “HIV is mainly a sexually transmitted disease”, “I believe that women can do the same things that men can do”.

example, the issue of abortion evoked considerable and lively debate in the **arepp:Theatre for Life** groups but almost none in the control focus groups.

Educators interviewed believed that the **arepp:Theatre for Life** presentations encouraged learners to “speak openly about the issues”. They said these learners were “more able to deal with conflict”, “more self-directing in discussions around careers” and “better able to talk, even in English orals”. One of the things commented upon by learners in the **arepp:Theatre for Life** focus groups was that they felt “comfortable” with the **arepp:Theatre for Life** actors and that this enabled them to discuss things they would not formerly have been able to discuss without embarrassment:

*I remember that one guy asked if you would get AIDS if you sucked your girlfriend's vagina and they answered straight.*

Further confirmation of the impact of the presentations which we see as having been translated into the stronger quantitative scores, as articulated by learners when we asked them what difference the presentations had made to them, included:

*It gave me the confidence to say “no” and not just do what my friends do. I don't let my friends dictate to me any more.*

*It made me think about my own life – I realised you don't need to fall for pressure from your friends. And I choose my friends more carefully now.*

Within the **arepp:Theatre for Life** cohort, we then looked at the differences overall related to the variables and found the following:

- ◆ Where a school had defined itself as “religious”, this did not affect impact much.
- ◆ Change was statistically significant across all economic categories but greatest in the poorest category and least in the most affluent.
- ◆ Learners who had had “some” or “a little” input before the **arepp:Theatre for Life** presentations showed greater positive change than learners who had had “lots” of input or “none”.

These results are interesting in that they suggest that the **arepp:Theatre for Life** presentations work best in soil that has, to some extent, been prepared but not over-prepared. They also suggest that this may work differently for different indicators. So, for example, on “constructive relationships” those who had had a lot of input showed a considerable increase, far more than they showed on any other indicator. This raises interesting ideas about a continuum of self-efficacy not unlike Maslow's Hierarchy of Needs. This might mean that, for example, that development around “constructive relationships” cannot fully take place until someone has developed “self-esteem and self-awareness”.

The statistical and focus group findings confirm a number of things that **arepp:Theatre for Life** knew at a “gut level”. So, for example, it is clear that learners are at different points on a continuum of self-efficacy and the weaker they are in this regard, the greater the difference is likely to be after intervention.

While some of the differences between the 2002 and end of 2004 surveys may be a result of maturation, there is an indication that the **arepp:Theatre for Life** approach has a specific impact. The clearest indication of this is the fact that, when learners had had no previous input and then had two **arepp:Theatre for Life** presentations, they out-scored those in the control

groups who had had no previous input either but did not get the **arepp:Theatre for Life** presentations.

In other words, if x (the **arepp:Theatre for Life** presentation), then there will be y (a change in sense of self-efficacy). This does not mean that “x” could not work equally well if it were some other intervention, but it does mean that **arepp:Theatre for Life** is not wasting its time and donors’ money. In addition, the focus group comparison does indicate that the **arepp:Theatre for Life** intervention has a specific effect.

The findings also seem to confirm at least some of the **arepp:Theatre for Life** expectations about impact:

- ◆ Improved understanding about gender equality (although with some distance still to go);
- ◆ A more constructive view of relationships;
- ◆ An understanding that there are choices, that people have the right to make choices, and that some choices are more constructive than others, even when one does not necessarily make constructive choices;
- ◆ Greater tolerance around making mistakes – both for oneself and for others;
- ◆ A strong sense that one can make a difference in one’s community;
- ◆ Increased self-esteem.

It also does seem, from the focus groups, that the **arepp:Theatre for Life** approach encourages a degree of openness, plain speaking and debate. We think it is fair to say that all this adds up to “increased self-efficacy”.

There also seems to be a fairly clear indication that **arepp:Theatre for Life** is not really winning the battle (and nor is anyone else) on issues such as:

- ◆ Celibacy as a real option;
- ◆ Condom use;
- ◆ Homophobia - although those with **arepp:Theatre for Life** exposure appear to be less homophobic.

It is worth noting some findings for **arepp:Theatre for Life** writers and, particularly, actors to keep in mind:

- ◆ Girls seem to be more open and tolerant on key issues than boys;
- ◆ Co-educational schools seem to create an environment where learners are able to grow in self-esteem but less so in tolerance;
- ◆ Learners in small schools start off stronger on most indicators;
- ◆ The learners from the less affluent schools are less likely to be tolerant of homosexuality;
- ◆ Boys find it harder to accept that it is alright to make mistakes;
- ◆ Learners from schools where teaching is worse (indicated by poor matric results) are less likely to feel it is alright to make mistakes. In general, poor schooling seems to have an effect of the development of self-efficacy.

Some of the useful learnings that emerge for **arepp:Theatre for Life**, and which may influence how it does its work in the future, include:

- ◆ There is a need to check that the transfer of information has not been compromised, for whatever reason, and to reinforce it, particularly where historical events may impact negatively on it.

- ◆ It seems that, while **arepp:Theatre for Life** does impact on the cohorts that see its presentation, it does not make an impact on the school as a whole.
- ◆ The presentations are more likely to have a greater impact when there has been “some” previous input but not “lots,” although learners in schools where there has been “lots” of input show an openness to improvement in areas such as “tolerance” and “constructive relationships”.
- ◆ Position on the continuum affects the ability to “learn” – this may mean huge impact on indicators such as “self-esteem/self-awareness” for those just starting along the continuum, and almost no impact on something like “constructive relationships”, whereas those much further along may show little improvement on “self-esteem/self-awareness” but more on “constructive relationships” and “tolerance”.
- ◆ Learners in general have a sense of community and of being part of a community that is important in the development of their self-efficacy and probably to the form it takes.
- ◆ There is lots of work still to be done around issues of gender equality, tolerance and human rights in general and the de-stigmatisation of a range of issues including homosexuality and abortion. There is a tension between a progressive and tolerant legislative framework and often conservative and intolerant religious views.

We believe that **arepp:Theatre for Life** staff could usefully explore the output of the data analysis in more detail than has been possible here. There is a commitment from **arepp:Theatre for Life** to incorporating some form of statistical data collection and analysis into its work in a more routine way in future. This should include:

- ◆ Revision of the indicators and the survey questionnaire;
- ◆ Clarity on the variables about which information is needed and improved methods of collecting this information;
- ◆ Review of the administrative systems that support the process;
- ◆ Ongoing entering of the data on a programme like SPSS;
- ◆ Regular analysis.

The full report deals also with some of the issues around embedding the work of **arepp:Theatre for Life** in the schools and local context. Despite the positive view of all the stakeholders with whom we spoke about **arepp:Theatre for Life** and its work, the organisation has not developed strong relationships within the schools or with local support structures. The lack of impact on the schools as a whole is reflected in the findings of the quantitative study and is something which **arepp:Theatre for Life** is aware that it needs to address. This may well require that it focus more on the aspect, at the expense of quantity in terms of numbers reached, in the future. Some of the legwork has been done in departments, schools and with other non-governmental organisations, but the programme design needs to incorporate this aspect directly, with the necessary investment in time and money to make it work.

The building up of a comparative statistical data set over a three year period is a unique attempt for an organisation such as **arepp:Theatre for Life** to move “beyond gut feeling.” We believe it has been a very worthwhile exercise and one that has increased the organisation’s understanding of what it is trying to achieve. We believe that the emphasis of the **arepp:Theatre for Life** model on the importance of a sense of self-efficacy in enabling young people to make informed choices is a valid one.

Our findings support this opinion although neither we nor **arepp:Theatre for Life** can take this further to measure the implications in terms of behaviour change. In any event, attributing behaviour change to any one intervention is not feasible in societies like ours in which there is a multiplicity of interventions and influences. Moreover, behaviour change itself needs to be explored as a concept. It is not as simple as “condom use” or even “abstinence”, although such choices will clearly have an impact on rates of HIV infection and teenage pregnancies. In

a society where child and woman abuse, violence in general, HIV/AIDS, unemployment, crime, racism, homophobia, corruption and religious intolerance (to mention but a few of our problems) are rampant and often define the parameters of our behaviour, a more nuanced view of behaviour change is needed.

Finally, we believe that the following questions posed for us as external evaluators of the evaluation project have been answered:

- ◆ Is it possible for an organisation such as arepp:Theatre for Life to go beyond gut level in assessing its impact?
- ◆ Did arepp:Theatre for Life succeed in doing this?
- ◆ Did the results show an impact?

The answer to all three questions is “yes”.

### **Summary of recommendations**

That arepp:Theatre for Life:

- ◆ Revisit how it deals with issues such as abstinence, condom use, homophobia, gender, reinforcement of accurate information.
- ◆ Use the findings of the evaluation project to facilitate a more differentiated approach to schools.
- ◆ Review how best to embed its work in a school and local context.
- ◆ Explore the output of the data analysis in more detail than has been possible in the external evaluation.
- ◆ Revise its indicators of a sense of self-efficacy and the related statements through which impact in this regard is measured.
- ◆ Clarify the variables about which information is needed in profiling schools and improve its methods of collecting information in this regard.
- ◆ Improve its administrative systems for supporting data collection. This should include the establishment of a database for entering monitoring information and the institution and implementation of better systems for recording school data.
- ◆ Agree on a limited number of external indicators that it believes will support a finding that its work has an impact and then invest the necessary time and money on keeping data on these indicators accurate and up to date. To some extent, such indicators are also variables in profiling the schools and can be used as such (e.g. pregnancy rates, violent incident rates).
- ◆ Continue to monitor its impact rigorously as a way of ensuring ongoing improvement of its work. Monitoring of impact should, wherever possible, be done against a comparable control group. The gathering of quantitative data about impact should always be supplemented by qualitative data gathered through observation, interviews and focus groups.
- ◆ Address the issue of how better to brand itself so that the name and the intervention are inextricably linked.