

AN EVALUATION OF A HIGHSCHOOL PREPARATORY PROGRAMME

HELÉNE DU PLESSIS

DPLHEL003

A dissertation submitted in partial fulfilment of the requirements for the reward of the Degree of  
Master of Philosophy (Programme Evaluation)

Faculty of Commerce

University of Cape Town

2015

**COMPULSORY DECLARATION:**

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works of other people has been attributed, cited and referenced.

Signature:

Signed by candidate

Date: 10/12/2015

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

## **ACKNOWLEDGEMENTS**

I would like to express my sincere gratitude to the following:

My supervisor, Dr. Sarah Chapman, for her ongoing support, advice and patience.

Programme staff, for their assistance with data collection and their willingness to engage with the evaluation process.

All of the participants who took part in the focus group and interviews.

My family and loved ones, for their support and encouragement.

## EXECUTIVE SUMMARY

This dissertation reports a process and outcome evaluation of an anonymised organisation's preparatory programme. The preparatory programme is an out-of-school time (OST) programme that aims to prepare the students that are selected for the scholarship programme, for high school. The programme has been implemented since 2008 and this dissertation will focus on the cohorts which received the programme in 2014 and 2015.

The preparatory programme was investigated in terms of its plausibility and a theory of change was developed in consultation with the programme co-ordinator. Three questions related to the programme's process were posed. These questions related to whether the programme had been implemented as intended, whether the CAT instrument that the programme was using was a useful tool for informing selection, and which aspects of the programme the participants found most and least helpful. Five questions regarding outcomes were posed that enquired whether students who had received the programme improved their mathematics and English performance, their cognitive reasoning ability, and their self-efficacy; and whether students in each of the teaching streams had benefitted equally from the intervention.

The evaluation used secondary data collected throughout 2015 and included student' term 2 and term 4 report cards, their performance on internal programme assessments, CAT scores, SEQ-C results, a focus group with the teachers who taught on the 2014 iteration of the programme, and interviews with the programme co-ordinator and students that had received the programme. Data analysis methods included the use of descriptive statistics, as well as parametric and non-parametric statistical tests for quantitative data. IBM Statistics 22 was used for the analysis of quantitative data and QSR NVivo 10 was used for qualitative data.

The results revealed that several minor changes were made to the planned structure of the programme in order to maintain the quality of the intervention. Students found the mathematics and English components of the programme useful, although how well they retained and applied the content from the programme varied. Students did not find the creative writing workshops or the study skills workshop particularly useful. The CAT instrument could potentially be a useful tool for informing selection, although it is not currently being used to its full potential. Students' mathematics aggregates improved significantly, while English aggregates did not. There was no significant change in internal assessment scores for both mathematics and English. Students in the mid-stream appeared to benefit most from the preparatory programme. There was a significant

improvement in student CAT scores, as well as on each of the CAT subscales (quantitative, verbal, and non-verbal). There was a significant increase in social self-efficacy scores and a significant decrease in academic self-efficacy scores. Total self-efficacy scores and emotional self-efficacy scores did not change significantly between the pre- and post-test. However, none of these observed effects could be attributed directly to participation in the preparatory programme due to the lack of a comparison group. It is possible that changes between repeated measures on participants were due to maturation effects, regression to the mean, or another historical event which influenced the outcomes of the programme participants.

The evaluator makes the following key recommendations:

1. Use the data generated by the CAT as part of the selection process, or discontinue the use of the CAT entirely and focus on other selection criteria.
2. Review the content and method of presentation of both the creative writing workshops and the study skills workshop in order to ensure that it is evidence-based and in line with participant needs.
3. Standardise the difficulty of the internal assessments administered by programme staff in order to make the results comparable to one another. This will enable programme staff to monitor student progress.
4. The programme may be able to determine which students would benefit most from the intervention they offer by conducting further research. This would enable the programme to streamline the selection process and invest its resources in the students which would benefit most.

These recommendations for programme improvement are presented in greater detail in the discussion chapter of this report.

## TABLE OF CONTENTS

<b>LIST OF ABBREVIATIONS</b>	<b>7</b>
<b>LIST OF TABLES</b>	<b>8</b>
<b>LIST OF FIGURES</b>	<b>9</b>
<b>CHAPTER ONE: INTRODUCTION</b>	<b>10</b>
Programme Description	10
Programme Theory	16
Programme Plausibility	17
<i>Will programme activities lead to improved academic performance?</i>	18
<i>Is there evidence for a causal relationship between academic performance and level of preparedness for high school?</i>	21
<i>Is there evidence for a relationship between self-efficacy and academic performance?</i>	23
Evaluation questions	24
<i>Process Evaluation</i>	24
<i>Outcome Evaluation</i>	25
<b>CHAPTER TWO: METHOD</b>	<b>26</b>
Participants	26
Process Evaluation	26
Outcome Evaluation	30
<b>CHAPTER THREE: RESULTS</b>	<b>37</b>
Process Evaluation	37
Outcome Evaluation	47

<b>CHAPTER FOUR: DISCUSSION</b>	<b>56</b>
Process Evaluation	56
Outcome Evaluation	61
Limitations	65
Conclusion	67
<b>REFERENCES</b>	<b>69</b>
<b>APPENDIX A: SEMI-STRUCTURED FOCUS GROUP PROTOCOL</b>	<b>72</b>
<b>APPENDIX B: SEMI-STRUCTURED INTERVIEW PROTOCOL</b>	<b>74</b>
<b>APPENDIX C: CONSENT FORM FOR INTERVIEWS</b>	<b>76</b>
<b>APPENDIX D: CONSENT FORM FOR SEQ-C QUESTIONNAIRE</b>	<b>77</b>
<b>APPENDIX E: SELF-EFFICACY QUESTIONNAIRE FOR CHILDREN (SEQ-C)</b>	<b>78</b>
<b>APPENDIX F: POWER CALCULATIONS</b>	<b>80</b>

## LIST OF ABBREVIATIONS

<b><i>Abbreviation</i></b>	<b>Explanation</b>
<i>CAT</i>	Cognitive Abilities Test
<i>CAT2E</i>	Cognitive Abilities Test second Edition
<i>CAT3</i>	Cognitive Abilities Test version 3
<i>CAT4</i>	Cognitive Abilities Test version 4
<i>OST</i>	Out-of-school time
<i>SEQ-C</i>	Self-Efficacy Questionnaire for Children



## LIST OF TABLES

Table 1.	Preparatory programme logic model	16
Table 2.	Participants for the evaluation	26
Table 3.	Themes used for analysis of qualitative data from focus group with teachers and interview with programme co-ordinator	28
Table 4.	Themes used for analysis of qualitative data from interviews with students from the 2014 cohort	30
Table 5.	Summary of research design section	36
Table 6.	Descriptive statistics for participants for the CAT and subscales	40
Table 7.	General statistics for CAT quantitative subscale and final mathematics score model	42
Table 8.	Pre and post scores for mathematics and English report card results	48

## LIST OF FIGURES

Figure 1.	Structure of the preparatory programme	13
Figure 2.	Organisational structure of the programme	15
Figure 3.	Theory of change for the preparatory programme	17
Figure 4.	Design for the mathematics and English performance aspect of the outcome evaluation	31
Figure 5.	Design for the internal test performance aspect of the outcome evaluation	32
Figure 6.	Design for the cognitive reasoning aspect of the outcome evaluation	33
Figure 7.	Design for the self-efficacy component of the outcome evaluation	34
Figure 8.	Stream performance on English internal assessments	50
Figure 9.	Stream performance on mathematics internal assessments	51
Figure 10.	Stream performance for term 2 and term 4 English report card marks	52
Figure 11.	Stream performance for term 2 and term 4 mathematics report card marks	53
Figure 12.	Amended impact theory for knowledge component of the preparatory programme	61

## **Introduction**

Research has indicated that children from low-income families are at greater risk of academic underachievement and failure (Cooper, Charlton, Valentine & Muhlenbruck, 2000). Therefore these children may need to supplement what they are taught during school hours in order to attain their full academic potential. After-school programmes and vacation schools are two forms of intervention which take place outside of regular schooling hours. These types of intervention are commonly referred to as out-of-school time (OST) interventions. OST programmes have become an increasingly popular form of intervention for improving student achievement (Miller, 2007), and they have been demonstrated to be particularly beneficial to disadvantaged students (Miller, 2003). Academic performance in high school has important repercussions for admittance into tertiary education institutions, job opportunities and potential income.

This evaluation will focus on the organisation's preparatory programme, which aims to prepare students selected for the organisation's scholarship programme for high school. The organisation has been anonymised in this report at the request of programme staff and forms part of the evaluation agreement between the evaluator and the organisation.

The aim of this dissertation is to assess whether the organisation's preparatory programme improved student outcomes in terms of their mathematics and English results, their self-efficacy, and their cognitive reasoning ability. The organisation requested an evaluation of their preparatory programme in order to gain insight into the potential impact of their programme, as well as to receive recommendations for programme improvement.

This chapter provides a detailed description of the preparatory programme, the underlying assumptions within the programme theory, an analysis of the plausibility of these assumptions, as well as the evaluation questions derived from the programme theory that will be answered by this evaluation.

## **Programme Description**

The organisation was established in 2007 and is funded by a foundation. The foundation focuses on funding educational initiatives in the Western Cape.

The programme offers scholarships to academically talented students with limited resources, allowing them to attend some of the best high schools in the Cape Town area (Anonymous, n.d.). This is done in order to allow the students to attain their full potential and contribute meaningfully

to their families, communities and country. Once students are selected into the programme, the programme provides academic, financial and social support from grade 8 until they matriculate.

The high schools the students are given scholarships to attend are selected on the basis of their commitment to academic excellence, as well as their holistic approach to education. Each of the schools accepts a pre-determined number of programme students each year. The following nine schools are currently participating in the programme:

- Camps Bay High School
- Herschel Girl's High School
- Rondebosch Boy's High School
- Rustenburg Girl's High School
- South African College High School
- Springfield Convent School
- Westerford High School
- Wynberg Boy's High School
- Wynberg Girl's High School

Before the students begin the grade 8 year at their new scholarship schools they are required to participate in the preparatory programme which takes place between June and November in their grade 7 year. This evaluation will focus solely on the preparatory programme, which focuses on preparing grade 7 students for high school. The focus of this evaluation was determined in consultation with the programme co-ordinator.

### **Goals/aims of the programme**

The broad objectives of the preparatory programme are to:

- 1). Ensure that students are adequately prepared for high school
- 2). Provide academic support and skills development for students
- 3). Provide psychosocial support for students

## **Selection Process**

Students are eligible to apply to write the selection exam for the programme if they meet the following criteria:

- 1). Student is currently in grade 7.
- 2). Student is currently thirteen years old or younger.
- 3). Student achieved a 65% aggregate in grade 6.
- 4). Student must demonstrate clear financial need.

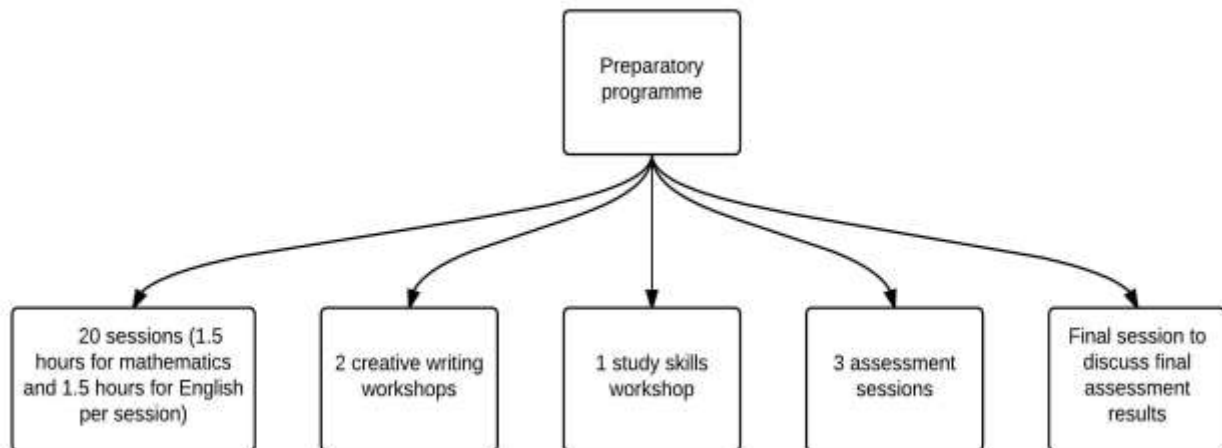
A parent or caregiver of the student applying for the programme is then required to attend a Registration Day. Once registration has been completed, learners write an examination based on the mathematics and English curriculum in grade 6. Based on the results of the examination and the demonstrated financial need of each student, approximately 120 students are shortlisted for a round table interview with programme staff. These students are also required to take the Cognitive Abilities Test (CAT). This tests the students' three main areas of reasoning: quantitative, verbal, and non-verbal reasoning. Approximately 60 students go through to the next round of interviews with the principals of participating high schools. Up to 50 students are selected for the programme by the programme staff in consultation with the headmasters of the high schools.

## **Programme Activities**

Once students are accepted into the preparatory programme, the results of the CAT are used in conjunction with their performance on the entrance examination to determine which teaching stream the students will enter. Students are divided into three streams which study the same material using slightly different teaching approaches. The first stream is for students who are particularly academically gifted and are expected to perform well. The second stream is for students who are expected to pass without difficulty, but who may not do as well as the students in the first stream. The third stream is for students who may find passing very challenging. Students are carefully monitored throughout the programme to ensure that they are in the correct stream, and they can be moved into another stream although this is rarely necessary.

The preparatory programme consists of twenty-seven, approximately 3-hour sessions which take place on Saturdays and during the school holidays at a school in the Cape Town city bowl. It typically begins in the June/July school holiday and ends in mid-November.

The structure of the preparatory programme is depicted in Figure 1 below.



*Figure 1.* Structure of the preparatory programme.

Each stream has its own teacher for mathematics as well as for English, and the classes take place concurrently. The order of the mathematics and English teaching sessions are alternated from week to week, ensuring that students get the opportunity to engage with both of the subjects while they are still fresh, over the course of the programme. The first subject takes place for approximately an hour and a half after which the students are given a short break. The second subject then commences for another hour and a half of teaching time. These mathematics and English sessions focus on the skills and content which students would be expected to master as part of the grade 7 curriculum.

The assessment sessions are spread out throughout the programme. The initial assessment takes place before students have covered any of the programme content. The mid-term assessment takes place after students have completed a portion of the programme content, in order to test their understanding and check their progress. The final assessment takes place after all of the programme content has been completed. The initial assessment focuses on content which students should have covered as part of the grade 6 curriculum at their schools. The mid-term assessment focuses on testing the students' competency in content which has been covered by the programme up until that point. The final assessment is at a higher level of difficulty than the initial and mid-term assessments, as the students are tested on all content which they are expected to have some competency as part of the grade 7 curriculum- even content which was not covered by the programme.

Once the preparatory programme has been completed, students generally progress to the next stage of the programme which is the full scholarship programme. Students are placed in one of the

participating high schools and offered academic, financial and social support for the duration of their high school career.

### **Other services**

The foundation funding the programme has several partner initiatives which work alongside the preparatory and scholarship programmes in order to ensure that the students are given every opportunity to succeed. This includes placement into a hostel if the student's home is overcrowded or if it may not be safe for the student to remain there. Counselling services are also available to students with additional emotional or adjustment needs.

The organisational structure of the programme is presented on the following page in Figure 2. This diagram was developed by the evaluator and was based on information gathered during interviews with the programme co-ordinator, as well as information gathered from the organisation's website.

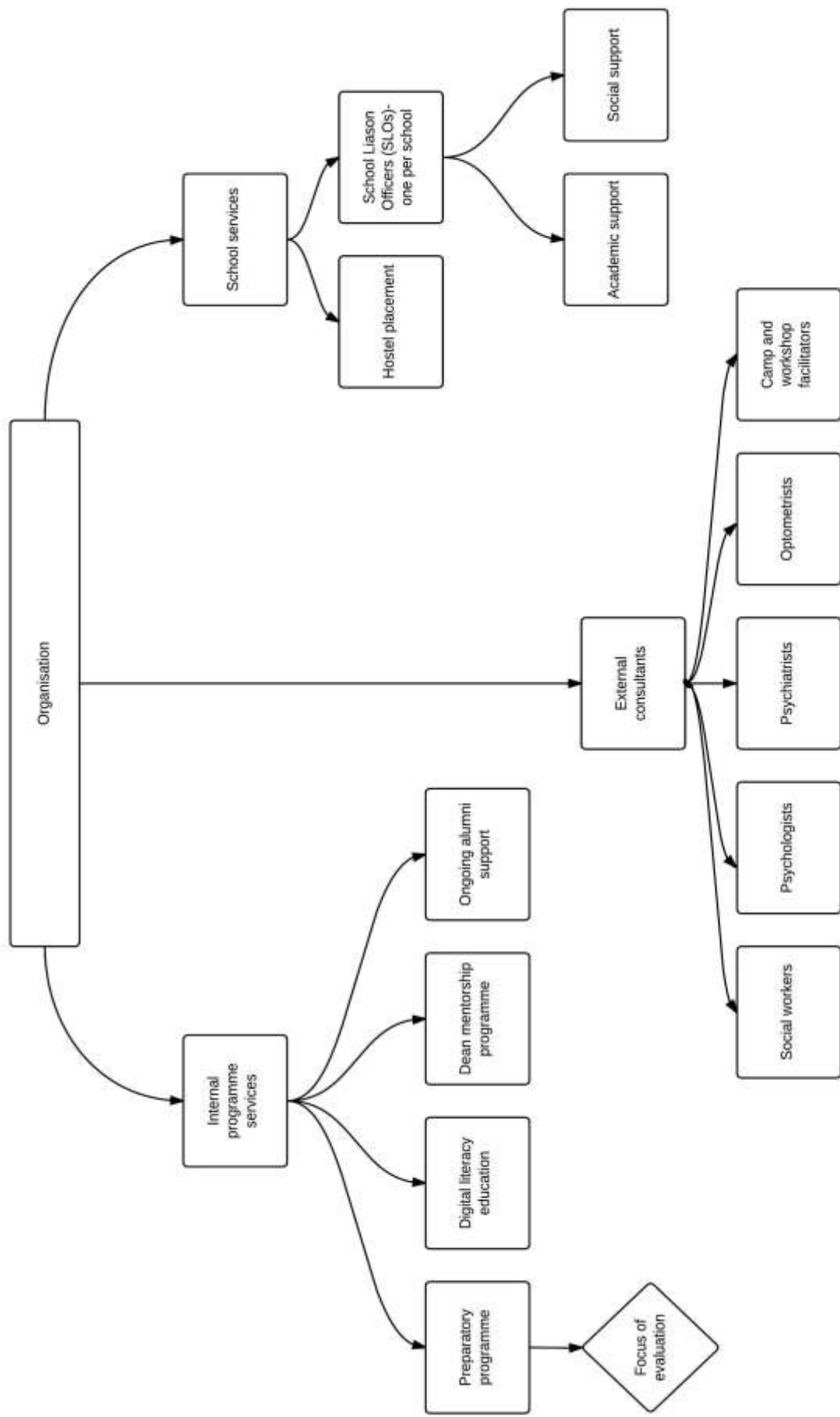


Figure 2. Organisational structure of the programme



## Programme Theory

The programme theory is the set of assumptions which underlie the design of the programme, and which serves to link the programme activities to the expected outcomes (Rossi, Lipsey & Freeman, 2004). Table 1 represents the logical framework for the preparatory programme.

Table 1

*Preparatory Programme Logic Model*

Inputs	Activities	Outputs	Initial Outcomes	Long-term Outcomes
<ul style="list-style-type: none"> <li>• Qualified middle-school teachers specialising in English and mathematics</li> <li>• Teaching material appropriate for grade 7 students based on current curriculum</li> <li>• Students apply for programme</li> </ul>	<ul style="list-style-type: none"> <li>• 2 creative writing work shops</li> <li>• 1 study skills work shop</li> <li>• 20 mathematics and English intervention sessions</li> <li>• 3 assessment sessions</li> <li>• Final feedback session</li> </ul>	<ul style="list-style-type: none"> <li>• 4 main types of intervention session completed</li> <li>• Designated number of beneficiaries reached</li> </ul>	<ul style="list-style-type: none"> <li>• Improved overall academic performance</li> <li>• Improved mathematics and English performance</li> <li>• Improved cognitive reasoning ability</li> <li>• Greater self-efficacy</li> </ul>	<ul style="list-style-type: none"> <li>• Improved level of preparedness for high school</li> <li>• Higher level of contribution to family, community and country</li> </ul>

The central assumption of the preparatory programme is that a curriculum-based mathematics and English intervention will improve academic performance for those subjects and increase self-efficacy. This then in turn leads to students being better prepared for high school, ultimately resulting in higher levels of contribution to families and communities by programme participants. This theory of change was developed by the evaluator in consultation with the preparatory programme co-ordinator. The theory of change underwent a series of revisions based on feedback received from the programme co-ordinator and other programme staff. The theory of change for the preparatory programme was used to clarify programme goals, and to develop evaluation questions which would focus on the outcomes of interest to the programme. Figure 3 on the following page represents the theory of change for the preparatory programme.

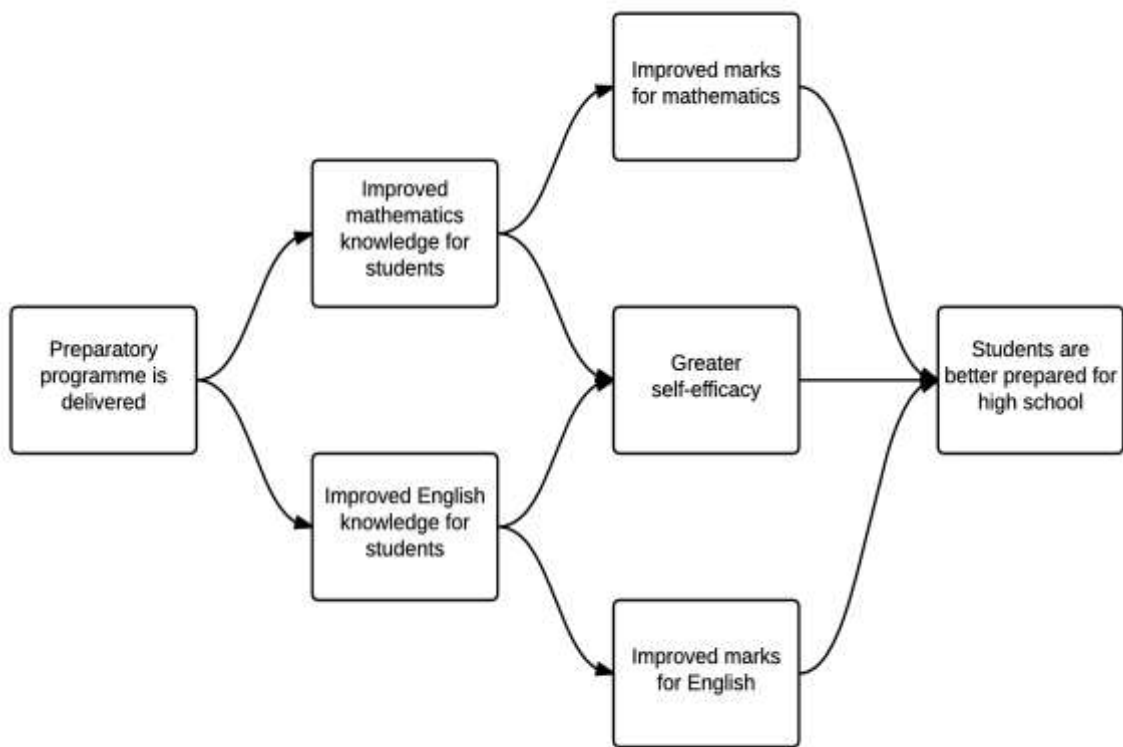


Figure 3. Theory of change for preparatory programme

### Programme Plausibility

There are three main causal assumptions underlying the preparatory programme, namely:

- 1). Programme activities will lead to improved mathematics and English performance
- 2). There is a causal relationship between academic performance and level of preparedness for high school
- 3). Programme activities will lead to greater self-efficacy

In order to assess the plausibility of these assumptions, a literature review of research on relevant topics and similar interventions was conducted. The searches were conducted using five electronic databases, namely EBSCO Host, ERIC, Google Scholar, PsychInfo, and SOCIndex. Searches were conducted between the 14 April and 5 December 2015, using the following keywords for each of the databases: “academic intervention” AND “standardised test outcomes”, “effectiveness of academic interventions”; “academic performance” AND “high school preparedness”, “self-efficacy” AND “academic performance”, “out-of-school time programmes”. This section will appraise the plausibility of the design of the programme, and not its actual performance.

## **Will programme activities lead to improved academic performance?**

Research suggests that a drop in academic performance upon entering high school is to be expected, for both the overall aggregate achieved, as well as for standardised test achievement (Cohen & Smerdon, 2009). This drop in achievement occurs for both low and high achieving students. That being said, programmes which promote academic achievement and offer support before students enter high school have been demonstrated to be successful in improving academic performance and increasing their confidence in their ability to learn and succeed at a high school level (Mizelle, 1995).

### *General programme components*

In general, the services provided by educational intervention programmes aim to counter negative school or community influences (lack of rigorous curriculum, poorly trained teachers, and a lack of role models) by providing the missing components (Gullatt & Jan, 2002).

Based on a series of meta-analyses of out-of-school time (OST) programmes, the following programme components were identified by Hammond and Reimer (2006) as contributing to a successful OST programme:

- Small classes with a low teacher to child ratio
- Educated and qualified staff members
- Continuous assessment of student progress
- Consideration of and provision for the health and mental wellbeing of the programme participants

Based on these criteria the potential impact of the preparatory programme is promising. Class sizes for the programme under evaluation are fairly small, with the largest class consisting of 18 students. All of the programme staff who teach on the preparatory programme are qualified teachers with several years of teaching experience. Student progress is monitored through the administration of the initial, mid-term and final assessments, and students have access to the services of external consultants (social workers, psychologists, psychiatrists, optometrists) once they are selected for the programme.

A common component of successful OST programmes is that their intervention content is aligned with the content covered by the students' school (Miller, 2007). This approach is utilised by the preparatory programme co-ordinator, as the material covered by the programme focuses on content and skills which students should cover in their grade 7 year.

There are several successful interventions offering similar programme activities. These are discussed in greater detail below.

'Sponsor a Scholar' is a successful programme which targets students who exhibit evidence of motivation, participation in extracurricular activities, have high levels of school attendance, and have an expressed interest in participating in the program (Gullatt & Jan, 2002). This programme offers financial aid and scholarships to participants and offers academic support services like extra tutoring; examination preparation classes; and workshops on study skills. After an outcome evaluation was conducted, researchers concluded that students who received the programme achieved better grades than those who did not. Students with lower levels of academic achievement appeared to benefit more from the programme. The evaluation used a longitudinal matched-comparison group design. Although random assignment was not used, the evaluator did attempt to control for a range of variables which were correlated with academic success (Johnson, 1999).

'Upward Bound' is a programme which targets students who have completed the 8th grade, whose families demonstrate significant financial need, and/or who are potentially the first in their families to go to college (Gullatt & Jan, 2002). Over the course of a year, students are provided with weekly academic support, particularly through additional tutoring and a mentoring relationship. An additional 6-week academic intervention is offered to participants during the school holidays. Students who received the intervention achieved a higher number of credits in non-remedial mathematics; were more likely to remain in school; and were more likely to receive financial aid in order to attend university (Gullatt & Jan, 2002). Overall, the evaluation appears to have been fairly rigorous, making use of random assignment and using propensity scores to match participants. However, there was a high rate of attrition from the intervention, with approximately 60% of the participants leaving the intervention (Myers & Schirm, 1999). The researchers used the opportunity to study the effect of dosage on the programme effect. Students who remained in the programme for longer than two years benefitted more than those who remained for less than two years (Myers & Schirm, 1999).

'Success for All' is a comprehensive academic intervention programme that has undergone regular annual evaluations since its implementation (Ross, Smith, Casey & Slavin, 1995). Although this intervention is designed for students in grades 1-3, it has several components in common with the preparatory programme, and both interventions are aimed at disadvantaged students. Similar components include grouping students by ability, small class sizes, regular assessments to monitor student progress, and social support for the students and their families. 'Success for All' offers individual tutoring in addition to the aforementioned activities. The evaluations conducted on the

programme were fairly rigorous, as students were matched with a control group on reading pre-test scores. Students in the comparison group attended a comparable school and lived in the same neighbourhood as the intervention group. Results indicated that the intervention group scored significantly higher on measures of language development than the control group and the intervention had the greatest effect on the students who scored in the lowest 25% percentile on the pre-test (Ross, *et al.*, 1995). Effect sizes for the intervention have been found to be quite consistent, averaging above .5.

### *Mathematics interventions*

Previous research indicates that middle and high school students are more likely than younger students to benefit from interventions aimed at improving their mathematics aggregate (Lauer *et al.*, 2004). The students targeted by the preparatory programme are therefore at an optimal age to benefit from an intervention focusing on mathematics.

Prior research has demonstrated that gains in mathematics were greatest when the programme combined academic content with a social aspect which allowed students to engage with the material (Lauer *et al.*, 2004). The ideal total intervention time for mathematics was a total of between 45-100 hours of teaching time (Lauer *et al.*, 2004). The preparatory programme offers participants approximately 30 hours of teaching time for mathematics. This is approximately 15 hours shorter than the minimum ideal time for the greatest programme impact.

### *English interventions*

Research and educational literature indicates that students in elementary and high school are more likely to benefit from interventions aimed at improving reading ability and English aggregates than children in middle school (Lauer *et al.*, 2004). This is in line with a body of well-established research in this area which highlights the importance of early intervention for developing reading competency and language development (Miller, 2007). The students in the preparatory programme do not fall within the ideal age range for a language intervention, as they are in middle school when they take part in the programme.

Reading and English gains were greatest for programmes which offered individual tutoring to students and when students were grouped according to ability (Lauer *et al.*, 2006). Students are divided into teaching streams according to ability for both the mathematics and English component of the preparatory programme. The preparatory programme does not currently offer individual tutoring to students.

Students benefitted most from reading/language interventions when they received between 44 and 84 hours of teaching time (Lauer *et al*, 2004). As with the mathematics component of the programme, the English component currently falls well below the minimum ideal time at approximately 30 hours of teaching time.

### *Summary*

Through looking at the meta-analyses of successful programmes and evaluations of programmes offering similar activities we are able to infer that it is feasible that the programme activities lead to improved academic performance. The following aspects of the preparatory programme are supported by the literature:

- Intervention content is aligned with content covered in regular schooling hours
- Qualified programme staff
- Small classes with a low teacher to child ratio
- Continuous assessment of student progress
- Consideration of and provision for the health and mental wellbeing of the programme participants
- Mathematics intervention is aimed at middle school students

The following aspects of the preparatory programme are not supported by the literature or could be altered to improve potential programme impact:

- Intervention does not include an individual tutoring component
- English intervention is aimed at middle school students. This form of intervention has been demonstrated to be most effective for elementary and high school students
- 30 hours of teaching time for the mathematics component of the programme. A minimum of 45 hours is recommended based on the literature
- 30 hours of teaching time for the English component of the programme. A minimum of 44 hours is recommended based on the literature

### **Is there evidence for a causal relationship between academic performance and level of preparedness for high school?**

Literature and previous research indicates that there are a range of developmental and contextual factors which contribute to adjustment to the challenges and demands of high school (Cohen & Smerdon, 2009). As children enter adolescence, they typically experience many changes in their

social contexts (Cohen & Smerdon, 2009). This may include an increased reliance on relationships with peers, and increased autonomy from their parents or guardians. Coupled with the shift in institutional context (transitioning from middle school to high school), students may experience increased levels of academic stress; difficulty with managing their time effectively; and lower self-esteem (Mizelle & Irwin, 2000).

Students with higher levels of academic performance prior to entering high school typically have a greater chance of graduating (Roderick, 2006), indicating that students who perform well academically have a greater likelihood of adjusting well to the challenges presented by high school. Furthermore, research suggests that students who are struggling academically may face additional challenges as they attempt to transition to high school, as the shift impacts further on their performance and ability to cope (Calabrese, 1987; Goodenow, 1993). As such, prior research indicates a relationship between academic performance prior to entering high school and ease of adjustment. Students who perform well academically appear to experience less difficulty in adjusting than those who receive lower grades.

### *Summary*

Prior research indicates that although the previous level of academic achievement is one of the factors which may contribute to a successful transition to high school, there are many other factors which also contribute, including emotional stability, gender, race, socioeconomic status, and characteristics of the high school attended. A drop in academic performance upon entering high school is to be expected, although students who perform well academically prior to high school seem to have less trouble adjusting than their peers.

A standardised intervention with limited flexibility and focusing only on academic performance may not be successful for all students. The research stresses that some students may require additional support in combination with an academic intervention in order to achieve the desired level of high school readiness.

Both the preparatory programme and the scholarship programme offer additional support to participants. This includes a wide range of services, including hostel placements, counselling services, or access to a social worker, optometrist or psychiatrist. These actions are supported by the literature and bode well for the potential impact of the academic portion of the programme. However, this evaluation will not focus specifically on the support aspect of the preparatory programme.

### **Is there evidence for a relationship between self-efficacy and academic performance?**

Self-efficacy is defined by Bandura (1977, 1982, 1997), as the level of confidence which individuals have in their ability to complete certain actions or attain certain levels of performance. The relationship between academic performance and general self-efficacy has been well established by research. A number of studies found evidence of a relationship between general self-efficacy and the academic performance of students (Hoge, Smit & Crist, 1995; Keltikangas-Jarvinnen, 1992; Wiltfang & Scarbecz, 1990). The strength of the relationship varies greatly across studies (J. Lane & A. Lane, 2001) and depends partly on how self-efficacy was operationalised and assessed in each study (Pajares, 1996). Most prior research found the correlation between general self-efficacy and academic performance to be fairly weak. Studies specifically investigating the relationship between academic self-efficacy and academic performance found evidence of a stronger relationship (Marsh, 1990).

The literature supports a link between academic performance and self-efficacy, in that greater belief in the ability to perform is linked to greater levels of achievement. The directionality of the relationship has been contested, with studies finding contradictory results. Muijs (1997) conducted a rigorous regression and structural equation modelling analysis, and concluded that academic performance was causally predominant over self-efficacy. This means that it may be plausible for interventions which target academic performance to improve self-efficacy, although programmes targeting self-efficacy in order to improve academic performance would be unlikely to obtain the desired outcomes.

#### *Summary*

The preparatory programme hopes to improve the self-efficacy of participants by improving their academic performance. As academic performance has been found to be causally predominant, this aspect of the programme's theory of change is supported by the literature. However, prior research is not able to provide an indication of the strength of the relationship between general self-efficacy and academic performance, as this has found to be highly variable across different studies. There is some evidence that there is a relationship between academic performance and academic self-efficacy, with greater academic performance being associated with greater academic self-efficacy.



## **Evaluation Questions**

The aim of this dissertation is to assess whether the preparatory programme adequately prepares programme recipients for high school, in terms of their overall academic performance, their mathematics and English performance, their self-efficacy, and their cognitive reasoning skills. This formative evaluation will take the form of a process and outcome evaluation.

## **Process Evaluation**

A process evaluation is used in order to determine whether a programme is being implemented as intended (Rossi *et al*, 2004). This often includes an assessment of whether the programme is reaching its intended recipients; whether the services are being delivered as intended by the design of the programme; and the fidelity of implementation of any support systems.

An assessment of service delivery may include an evaluation of whether participation in the programme is at levels which are acceptable to the programme, and whether some subgroups may be participating more or less than others (Rossi *et al*, 2004). The evaluation may also seek to determine whether the programme is being delivered in sufficient quantities, and/or whether it is delivered in a standardised fashion across different sites or to different groups.

An evaluation of the organisational functions and activities of a programme may produce information on the effectiveness of any relevant support functions (Rossi *et al*, 2004). This may include support personnel, funding, or selection procedures.

A process evaluation is often conducted in addition to an outcome evaluation, as establishing the extent of the fidelity of implementation can be helpful in explaining why a programme did or did not achieve the expected outcomes.

The process evaluation component of this evaluation will be of limited scope, as programme staff indicated that more process related questions would be of limited use to the programme in comparison to those indicated below. Programme staff indicated that an evaluation focusing on the outcomes of interest would be most useful to the programme and the information generated would be used in order to improve the programme.

The following questions were produced in consultation with the programme manager.

The specific questions related to process and fidelity of implementation which will be answered in this evaluation will be:

Service delivery:

1. Was the programme delivered as intended for the intended duration?

Organisational support:

2. Is the CAT instrument that the programme is using a useful tool for informing selection?

Service utilisation:

3. What aspects of the preparatory programme did the 2014 cohort find more/less effective?

### **Outcome Evaluation**

An outcome evaluation is conducted in order to determine the effects of the programme on the target population and seeks to determine to what extent the changes can be attributed only to the programme (Rossi *et al*, 2004).

The proximal outcomes for this intervention include mathematics and English academic performance, self-efficacy, and cognitive reasoning scores.

The specific questions related to the outcomes of the programme include:

1. Can observed improvements in learners' mathematics and English performance be attributed to the intervention?
2. Did students from the 2014 cohort in each of the streams benefit equally from the intervention? Did some streams benefit more (or less) than others?
3. Can observed improvements in learners' cognitive reasoning ability be attributed to the intervention?
4. Did the students in the 2015 cohort show improvements in self-efficacy scores?

## METHOD

### Participants

There were 181 participants in the evaluation. Participants were from the 2014 and 2015 cohorts, as per Table 2 below.

Table 2

#### *Participants for the Evaluation*

Cohort	Group	Participants
2014	Intervention group	47 students that received the programme
	Comparison group	78 students who were short-listed, but were not finally selected for the programme
	Teachers	6 teachers which taught on the 2014 programme
2015	Intervention group	50 students that received the programme

### Process evaluation.

This component of the evaluation was undertaken within either an exploratory or descriptive framework, or a combination of both. The specific approach used depended on the specific evaluation question being answered. Further detail on the approach used for each evaluation question related to process is provided below.

#### **1. Was the programme delivered as intended for the intended duration?**

##### ***Research design.***

In order to answer this question a combination of the descriptive and exploratory frameworks was utilised. These approaches were deemed appropriate, as the researcher intended gain an understanding of how the preparatory programme was run, as well as describing how the programme was delivered. A qualitative interview/focus group based approach was favoured instead of a survey or questionnaire based approach, as focus groups and interviews allow the researcher to clarify certain points with the participants, and to investigate additional information

that may arise from the discussion. Open-ended questions also allow participants to raise points which the researcher may not have initially considered.

### ***Materials.***

This component of the evaluation made use of a focus group. The focus group was comprised of the six teachers which taught the preparatory programme to the 2014 cohort, specifically three English teachers and three mathematics teachers. The focus group used a semi-structured format. Each of the questions were discussed by the group and additional points which were raised by the group were pursued further if they were judged to be pertinent to the evaluation. The focus group protocol is presented in Appendix A.

### ***Procedure.***

The focus group took place on 12 September 2015, directly after one of the intervention sessions. The focus group was conducted in a classroom at the school where the intervention takes place. The evaluator arranged for tea and refreshments. The evaluator led an informal discussion around the questions which are outlined in Appendix A. The focus group session took approximately 1 hour. The programme co-ordinator who teaches on the programme was interviewed separately. This interview took place on the 9 September 2015.

Consent for participation was obtained orally from the teachers prior to the group interview. They were assured that their participation is entirely voluntary and that their responses will be kept confidential. The focus group and interview were not recorded in order to make the participants more comfortable with the process. The evaluator took written notes during the session.

### ***Data analysis.***

The qualitative data from the focus group and interview were analysed using QSR NVivo 10 software in order to extract meaningful themes. The themes were organised around the sessions outlined in the focus group protocol. The themes are outlined in more detail in Table 3.

Table 3

*Themes used for analysis of qualitative data from focus group with teachers and interview with programme co-ordinator*

<b>Overall impressions</b>	<b>Structure</b>	<b>Duration</b>	<b>Resources and management</b>	<b>Challenges</b>	<b>Recommendations</b>
Mathematics	Order of sessions	Individual sessions	Resources	Mathematics	Mathematics
English	Streams	Programme length	Management	English	English
Other				Other	Other

**2. Is the CAT instrument that the programme is using a useful tool for informing selection?**

***Research design.***

This question was approached using the descriptive framework. This approach was used as the researcher aimed to describe how the CAT instrument was being used by the programme. A quantitative inferential statistics based approach was favoured as this would give the best indication of how the instrument was actually being used, and would provide more useful information than if only descriptive statistics were used.

***Materials.***

This component of the evaluation made use of secondary data which had already been collected by programme staff. Specifically, this evaluation used the pre-intervention results from the Cognitive Abilities Test 3 (CAT3) which had been collected as part of the selection process for the programme. The CAT3 measures students' underlying academic ability through three subscales- quantitative, verbal, and non-verbal reasoning ability (GL Education, 2015). More detail on the reliability and structure of the CAT3 is provided in the outcome evaluation portion of this section.

This evaluation also used the December 2014 mathematics and English results which were obtained from school reports. These reports were generated by the schools which the students attended.

### ***Procedure.***

The necessary data were obtained directly from the programme staff as it had already been collected and it formed part of the programme's data archive. It was not deemed necessary to obtain additional consent as this data formed part of the programme's routine data collection procedure.

### ***Data analysis.***

All quantitative data for this report was analysed using IBM Statistics SPSS version 22 software. A regression analysis between mathematics/English report card results and CAT standardised age scores and each of the CAT subscales (verbal, non-verbal, and quantitative reasoning) was run in order to determine whether the CAT is useful as a diagnostic tool for the programme. Mathematics and English results were analysed separately.

## **3. What aspects of the preparatory programme did the 2014 cohort find more/less effective?**

### ***Research design.***

This component of the evaluation utilised an exploratory framework. This approach was selected as the researcher aimed to gain a better understanding of which components students found most and least useful, as well as exploring how the preparatory programme could potentially be improved upon. Interviews were selected as the appropriate avenue for data collection, rather than surveys or questionnaires, as interviews allow the researcher to clarify certain points with the participants, and to investigate additional information that may arise from the discussion. Open-ended questions also allow participants to raise points which the researcher may not have initially considered. Interviews were also deemed more appropriate due to the potentially sensitive nature of the topic discussed.

### ***Materials.***

A series of semi-structured interviews with students from the 2014 cohort were conducted. The interview schedule is presented in Appendix B.

### ***Procedure.***

Four interviews were conducted in total. Two interviews were conducted with students who improved their mathematics and English marks markedly after the intervention, and two with students whose marks dropped after the intervention. A list of candidates was compiled for each,

and then the required number was randomly selected using a random number generator. The interviews themselves were conducted on the University of Cape Town campus. The dates of the interviews were determined in consultation with the interviewee. The interviews each took approximately 45 minutes. The evaluator took written notes during each of the interviews.

Consent for participation was obtained from the parents/guardians of the child prior to the interview (see Appendix C). The identity of the students who participated in the interviews will be kept confidential and no identifying information is provided in the final report. The interviews were not recorded in order to make the participants more comfortable with the process.

**Data Analysis.**

The qualitative data from the interviews was analysed using QSR NVivo 10 software in order to extract meaningful themes. The themes were organised around the sessions outlined in the interview protocol. The themes are outlined in more detail in Table 4.

Table 4

*Themes used for analysis of qualitative data from interviews with students from the 2014 cohort*

<b>Adapting to high school</b>	<b>Mathematics</b>	<b>English</b>	<b>Study skills workshop</b>	<b>Creative writing workshops</b>	<b>Recommendations</b>
Fitting in	Level of ability	Level of ability	Usefulness	Usefulness	Mathematics
Mathematics classes	Most helpful	Most helpful			English
English classes	Least helpful	Least helpful			Other
What has helped	Usefulness	Usefulness			

**Outcome evaluation.**

The quasi-experimental design selected for the outcome evaluation component of this evaluation varied and depended on the specific evaluation question being answered. Three outcomes are explored; (1) mathematics and English performance, (2) Cognitive reasoning ability and, (3) Self-efficacy. Further detail on the approach used for each evaluation question related to outcome is

provided below. A comparison group was not available for the majority of the outcomes, due to the complexity of matching and time constraints.

**4. Can observed improvements in learners' mathematics and English performance be attributed to the intervention?**

***Research design.***

A quasi-experimental retrospective, single group repeated measures study was used in order to assess whether there had been an improvement in the mathematics and English performance of the 2014 cohort. The T in the figure below indicates that there was only a single treatment group available, with no comparison group. The N indicates that selection into the treatment (the preparatory programme) was non-random. The X indicates the preparatory programme intervention which took place between June 2014 (after school examinations had been completed) and December 2014. O<sub>1</sub> and O<sub>2</sub> indicate the June and December report card results, respectively.

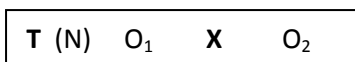


Figure 4. Design for the mathematics and English performance aspect of the outcome evaluation

***Materials.***

Mathematics and English performance between June 2014 and December 2014 were compared based on reports generated by the students' schools.

***Procedure.***

The parents of the 2014 and 2015 cohorts had given their consent at the interview stage of the selection process, for their children's reports and results to be used for research purposes. This permission was deemed necessary as the data formed part of routine monitoring and will be used for programme improvement.

The school report cards were collected from the students via email.

***Data analysis.***

The change in mathematics and English scores from the report cards were assessed using paired *t*-tests.



**5. Did students from the 2014 cohort in each of the streams benefit equally from the intervention? Did some streams benefit more (or less) than others?**

***Research design.***

A quasi-experimental retrospective, repeat measures study was used in order to assess whether the 2014 cohort showed an improvement in the scores achieved in the internal tests administered during the course of the programme. The internal test performance of the three streams which the students were divided into were compared with one another to determine if students at different levels of ability benefitted equally from the programme. The X indicates the preparatory programme intervention which took place between June 2014 (after school examinations had been completed) and December 2014. O<sub>1</sub> indicates the initial assessment, O<sub>2</sub> the mid-term assessment, and O<sub>3</sub> the final assessment.

T	(N)	O <sub>1</sub>	X	O <sub>2</sub>	O <sub>3</sub>
T	(N)	O <sub>1</sub>	X	O <sub>2</sub>	O <sub>3</sub>
T	(N)	O <sub>1</sub>	X	O <sub>2</sub>	O <sub>3</sub>

Figure 5. Design for the internal test performance aspect of the outcome evaluation

***Materials.***

The results of the three assessments conducted by the programme were used in order to assess the improvement of students in mathematics and English as they progressed through the programme. These assessments are compiled and administered by programme staff before the programme begins, mid-programme, and after the programme is completed. All three streams write the same tests.

***Procedure.***

The results of the three internal mathematics and English assessments form part of the data archive of the programme, and these data were made available to the evaluator.

***Data analysis.***

The results of the internal assessments were assessed using a mixed-design ANOVA. A disaggregated analysis which contrasts progression for the three different teaching streams was also conducted to determine whether one group benefitted more from the programme than the other.

**6. Can the observed changes in cognitive reasoning ability be attributed to the intervention?**

***Research design.***

A retrospective single-group repeat measures study was used in order to track the development of cognitive reasoning of students in the 2014 cohort between baseline and programme completion. It would have been ideal if a post-intervention measure could also have been obtained for the comparison group. However, due to financial constraints for the programme, this was not feasible. The X indicates the preparatory programme intervention which took place between June 2014 (after school examinations had been completed) and December 2014. O<sub>1</sub> indicates the CAT assessment administered during the selection process and O<sub>2</sub> indicates the post-test after programme completion.

<b>T</b>	(N)	O <sub>1</sub>	<b>X</b>	O <sub>2</sub>
<b>C</b>	(N)	O <sub>1</sub>		

Figure 6. Design for the cognitive reasoning aspect of the outcome evaluation

***Materials.***

The Cognitive Abilities Test, version 3 (CAT3) was used in order to assess the cognitive reasoning skills of the students. This test could not be included as an appendix as the programme does not have permission from the publisher to reprint the test.

The CAT3 has three separate batteries, assessing verbal reasoning, quantitative reasoning and non-verbal reasoning (Strand, 2004). A score for each of the batteries is calculated, as well as the mean score across all three batteries. The reliability of the CAT is reportedly high, with internal consistency estimates averaging .94 for verbal reasoning, .90 for quantitative reasoning, and .92 for non-verbal reasoning (Strand, 2004).

***Procedure.***

The cognitive assessments had already been collected by programme staff for all 125 students which were shortlisted for the programme. These results were used in order to establish if selected and non-selected students were equivalent at baseline. The intervention group was asked to repeat the CAT after the preparatory programme intervention had been completed. This second CAT assessment was also administered by programme staff. This assessment is routinely conducted by programme staff, so it was not deemed necessary to obtain additional consent for this component of the evaluation.

### ***Data analysis.***

An independent samples *t*-test was used in order to assess whether the group selected for the programme differed significantly on baseline CAT scores to those who were interviewed but not selected for the intervention. A paired sample *t*-test was used at pre-and post-programme to assess whether there had been an improvement in cognitive reasoning ability for the intervention group. Paired sample *t*-tests were also conducted on each of the subscales of the CAT.

## **7. Did the students in the 2015 cohort show improvements in self-efficacy scores?**

### ***Research design.***

A single group repeated measures study was used in order to assess whether there was an increase in the self-efficacy scores of students who received the programme. The X indicates the preparatory programme intervention which took place between June 2015 (after school examinations had been completed) and December 2015.

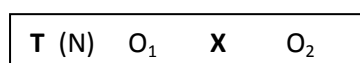


Figure 7. Design for the self-efficacy component of the outcome evaluation

### ***Materials.***

The Self-efficacy Questionnaire for Children (SEQ-C) was used in order to compare the self-efficacy scores of the 2015 cohort before and after the intervention had taken place (see Appendix E). The SEQ-C consists of twenty-four items which represent three domains of self-efficacy: social self-efficacy, academic self-efficacy, and emotional self-efficacy (Muris, 2001). The scale has been adjusted from the original in order to better suit the current research. Several of the items have been reworded, although the content has not been significantly altered. Each item is scored on a five-point scale ranging from 1 = never , to 5 =always. The internal consistency of the scale was within the acceptable range, with .88 for the total self-efficacy score and between .85 and .88 for subscale scores (Muris, 2001).

***Procedure.***

The SEQ-C was administered to the 2015 cohort by programme staff prior to the commencement of the programme in June 2015 and then again after the programme had been completed at the end of October 2015.

Parental consent for participation was obtained prior to the completion of the questionnaire (see Appendix D).

***Data analysis.***

Paired *t*-tests were used to compare general self-efficacy scores for the intervention group pre- and post. Paired sample *t*-tests were also conducted on each domain of self-efficacy (academic, social, and emotional self-efficacy).

Table 5

*Summary of research design section*

<b>Evaluation Question</b>	<b>Procedure</b>	<b>Number of participants</b>	<b>Type of participants</b>	<b>Materials</b>
<b>1</b>	Focus group conducted by evaluator	6	3 Mathematics teachers 3 English teachers	Semi-structured focus group protocol Figure 4
<b>2</b>	Pre-intervention cognitive abilities test and school report cards	47	2014 cohort	CAT3 and December 2014 school report card
<b>3</b>	Interviews conducted by evaluator	4	2014 cohort	Semi-structured interview protocol Figure 5
<b>4</b>	Pre- and post intervention school report cards	47	2014 cohort	June 2014 and December 2014 school report cards
<b>5</b>	Pre- mid- and post intervention mathematics and English assessments administered by the programme staff	47	2014 cohort	Baseline, mid-intervention and post intervention assessments for mathematics and English
<b>6</b>	Pre- and post intervention cognitive abilities tests for intervention group. Pre-test only for comparison group Administered by programme staff.	125	2014 cohort  Comparison group (reached interview stage but weren't selected)	CAT3
<b>7</b>	Pre- and post intervention questionnaires administered by programme staff.	50	2015 cohort	Self-Efficacy Questionnaire for Children (SEQ-C)  Appendix C

## RESULTS

The results in this section are presented according to the evaluation questions posed in the method section.

### Process Evaluation

#### 1. Service Delivery:

##### **Was the programme delivered as intended for the intended duration?**

A focus group was conducted with four out of the six teachers who taught on the preparatory programme in 2014. The focus group took place approximately 11 months after the final session of the 2014 iteration of the programme, and took approximately one hour to complete. The focus group schedule is presented as Appendix B.

The programme co-ordinator, who also teaches on the programme, was interviewed individually. The final teacher no longer teaches on the programme, could not be reached and was excluded from the analyses. The final sample consisted of three English teachers and two mathematics teachers.

### *Structure*

All of the teachers who participated agreed that overall the programme ran as intended for the 2014 iteration. The structure of the programme did not deviate greatly from what was intended, although there were a few alterations.

One creative writing workshop was planned, which took place at the beginning of the programme. After it had taken place, the programme co-ordinator was not satisfied with the quality of the workshop and another creative writing workshop was arranged with a different facilitator. This second workshop was an addition to the original planned programme activities.

Another change to the planned structure of the programme was the introduction of several 'block days', where students were taught only mathematics, or only English, for that session of the programme. This was done in order to accommodate the schedules of some of the teachers who had to go away on school camps. Although the structure of the sessions was altered, the overall amount of teaching time for each subject was not altered and students still received the intended amount of hours focusing on each subject.

### *Streams*

Each of the students was assigned to one of three streams for mathematics and English as planned. This was based on their initial assessment scores. The students remained in these streams for the duration of the preparatory programme. There is a consensus among the teachers that the streaming system works well and allows them to tailor their teaching approach to the level of ability in the group they are teaching.

### *Content*

Each teacher in the programme received all of the content that needed to be covered right at the beginning of the programme. Generally speaking, each of the streams moves through this content at their own pace, and additional content is added at the teacher's discretion. The content was not divided into sessions and it was up to the teacher to determine how much content would be covered by their class within a particular session. This matches what was done in previous iterations of the programme. English teachers expressed confidence that they had covered the necessary content by the end of the programme. The mathematics teachers indicated that they would have liked to cover the content in greater detail and had difficulty covering all of the necessary content with their students.

### *Duration*

All of the teachers indicated that they used the full amount of time allocated to them for teaching as there was a lot of content to cover.

Although the teachers indicated that they would like more time with the students, they agreed that it would not be practical to extend the programme duration by beginning earlier in the year, or by having more sessions within the current programme period. The current programme duration is already considered to be demanding of the students, and teachers indicated that any benefits obtained from additional teaching time would likely be countermanded by the additional demands placed on students in terms of time and homework. It was also considered to be impractical to extend the duration of the current sessions. Teachers indicated that they thought the students would not be able to concentrate if the duration were extended, and they would therefore not benefit from the additional time.

### *Facilities and resources*

The teachers all agree that they had all of the necessary resources in order to teach. The classrooms were always accessible and there were enough tables and chairs for all of the students. The programme provided the students and teachers with any necessary stationery and workbooks, and these were all distributed as planned. The programme also provided the students with breakfast upon arrival at the session, a mid-morning snack, and a packed lunch to take home with them. Overall the food provided was well received by the students, but one teacher indicated that the breakfast protein bars went uneaten by the majority of the students.

Several teachers identified the smart boards in some of the classrooms as problematic. On several occasions the smart boards were not functioning correctly, or they had been written on with a white board marker (which damages the smart board). The mathematics teachers in particular rely on the smart boards in order to cover the content, and on these occasions they had to make use of alternative methods in order to proceed with the session.

### *Summary*

The structure of the programme did not deviate greatly from what was planned, and changes were made in order to ensure that the quality of certain programme components was sufficient. Although the planned structure of the programme was not strictly adhered to, participants still received the intended amount of teaching time for each subject. The streaming system was implemented as planned and students remained in their streams for the duration of the programme. Teachers and students were supplied timeously with adequate resources and overall the teaching facilities were deemed to be satisfactory. The smart board systems in several of the classrooms were identified as problematic.

## **2. Organisational support:**

### **Is the CAT instrument that the programme is using a useful tool for informing selection?**

Two sets of CAT3 data were analysed in order to determine whether students selected for the programme ( $n = 47$ ) scored significantly more on the CAT3 and its subscales than students who were not selected for the programme ( $n = 78$ ). Participants ( $n = 125$ ) completed the CAT3 testing in a single day of testing. Participants received a set amount of time to complete each of the three subscales of the CAT. Each participant received an overall mean standardised age score, as well as a



standardised age score for the quantitative, non-verbal and verbal subscales. Standardised age scores describe how a student performed compared to a representative sample of students from the general population. It is important to note that the CAT3 was standardised on populations outside of South Africa, and that the mean standardised age score of 100 for the CAT3 indicates the average in those populations, and that it may not necessarily be representative of the average score in a South African sample.

IBM SPSS Statistics 22 was used to generate descriptive statistics and perform both parametric and non-parametric tests. Table 6 displays the descriptive statistics for each group.

Table 6

*Descriptive statistics for participants for the CAT and subscales*

	Selected for the programme ( <i>n</i> = 47)			Not selected for the programme ( <i>n</i> = 78)		
	<i>M</i>	<i>Mdn</i>	<i>SD</i>	<i>M</i>	<i>Mdn</i>	<i>SD</i>
Mean standardised age score	105.3	106	5.89	103.5	103	9.63
Quantitative standardised age score	108.9	110	8.32	107.7	107	10.2
Verbal Standardised age score	102.9	102	9.17	100.8	101.5	12.2
Non-verbal standardised age score	104.2	105	7.97	101.8	102	10.2

An independent samples *t*-test was run in order to determine whether the mean standardised age score of the students selected for the programme was significantly different from those not selected for the programme. Although the distribution of each of the two groups was approximately normal, Levene’s test for homogeneity of variance was significant,  $F = 9.43$ ,  $p = .003$ . Therefore, the results reported are where equal variances were not assumed.

The mean standardised age score for students who were selected for the programme ( $M = 105.3$ ,  $SE = 0.82$ ) did not differ significantly from students not selected for the programme ( $M = 103.5$ ,  $SE = 1.12$ ),  $t(123) = 1.35$ , 95% BCa CI [-0.866, 4.63],  $p = .178$ . There was a small effect,  $d = 0.12$ .

A series of Mann-Whitney non-parametric tests were run on the quantitative, verbal and non-verbal standardised age subscales of the CAT3 in order to determine whether students selected for the programme differed significantly from those not selected for the programme. The distribution for each of the subscales for students not selected for the programme was approximately normal. However, for each of the three subscales the distribution for students selected for the programme

deviated markedly from normality. The non-verbal and the quantitative subscales for this group were leptokurtic, and the verbal subscale was positively skewed. In addition, Levene's test for homogeneity of variance was significant for the verbal standardised age score subscale,  $F = 5.03$ ,  $p = .02$ .

The quantitative standardised age score for students who were selected for the programme ( $Mdn = 110$ ) did not differ significantly from students not selected for the programme ( $Mdn = 107$ ),  $U = 1662$ ,  $z = -.837$ ,  $p = .383$ ,  $r = -.07$ .

The verbal standardised age score for students who were selected for the programme ( $Mdn = 102$ ) did not differ significantly from students not selected for the programme ( $Mdn = 101.5$ ),  $U = 1663$ ,  $z = -.867$ ,  $p = .386$ ,  $r = -.07$ .

The non-verbal standardised age score for students who were selected for the programme ( $Mdn = 105$ ) did not differ significantly from students not selected for the programme ( $Mdn = 102$ ),  $U = 1553$ ,  $z = -1.428$ ,  $p = .153$ ,  $r = -.12$ .

Linear regression analyses were conducted using the mean standardised age score generated by the CAT instrument as the predictor variable, and the final report marks for mathematics and English as the outcome variables. The mean standardised age score is the average of the verbal, quantitative, and non-verbal subscales for each student. The final mathematics/English score for each student was obtained from their final report card generated by their school. Only students who were selected for the programme were included in this analysis ( $n = 47$ ). Analyses were conducted using IBM SPSS Statistics 22.

The mean standardised age score generated by the CAT did not contribute significantly to the prediction of either the final mathematics report result,  $r = .196$ , 95% BCa CI  $[-.074, .422]$ ,  $p = .188$ , or the final English report result,  $r = .069$ , 95% BCa CI  $[-.216, .363]$ ,  $p = .644$ .

Another linear regression was conducted to determine whether the verbal subscale on the CAT could predict the final English report card score. The verbal standardised age score subscale on the CAT instrument did not contribute significantly to the prediction of the final English result,  $r = .15$ , 95% BCa CI  $[-.125, .381]$ ,  $p = .315$ .

A final linear regression was conducted to determine whether the quantitative subscale on the CAT could predict the final mathematics report card score. The CAT quantitative subscale scores were mean centred in order to make the interpretation more meaningful. The quantitative reasoning standardised age score subscale on the CAT instrument contributed significantly to the prediction of the final mathematics report result,  $r = .362$ , 95% BCa CI  $[.05, .578]$ ,  $p = .012$ , and accounted for

13.1% of the variation in final mathematics results. Scores received ranged between 87 and 128 in the sample. For each additional point on the CAT quantitative score above the sample mean, the final mathematics score increases by 0.46%. The following regression equation was derived:

$$\text{Final mathematics score} = 82.17 + (0.467 \times \text{CAT quantitative score})$$

Tables 7 displays the coefficients for the CAT quantitative subscale.

Table 7

*General statistics for CAT quantitative subscale and final mathematics score model*

	B	SE B	$\beta$
Constant	82.17	1.41	
Quantitative subscale score (mean centred)	.467	0.179	.362

*Note:  $R^2 = .131, p = .012$ .*

Because the data was mean centred, the constant in Table 7 represents the average mathematics grade when the CAT score is equal to the sample mean. The score itself (82.17%) is also noteworthy, as it indicates that the students selected for the programme achieved high marks for mathematics as a cohort.

Student gender and the primary school attended in grade 7 were added to the model, but they were not associated with either the quantitative CAT score or the final mathematics mark. The model was also not significantly improved by the inclusion of these variables. They were therefore excluded from the analysis. Additional data on student race and family income were unavailable.

### **3. Service utilisation:**

#### **What aspects of the preparatory programme did the 2014 cohort find more/less effective?**

Individual interviews were conducted with three of the students who took part in the programme in 2014. The interviews took place approximately 11 months after the final session of the 2014 iteration of the programme, and each took approximately forty minutes to complete. The interview schedule is in Appendix C.

Students were identified as candidates for the interview based on the increase or decrease in their mathematics and English marks between June 2014 and December 2014. All students which showed

a marked (5% or greater) increase/decrease were identified and ten candidates were randomly selected from this list. All students were contacted via email and asked if they were willing to be interviewed. Eight students responded to the initial email. The final sample consisted of one student which increased their mathematics and English marks between June and December 2014, and two students whose marks decreased.

It is important to note that all five of the students whose marks were flagged for decreasing substantially had the same English teacher in the preparatory programme.

### *Social support*

Social support from peers which participated in the preparatory programme and had been placed in the same high school were identified by all of the students as a key contributor to adapting successfully to high school. This was especially important at the beginning of high school, as students felt as though they did fit in with their new peers and relied on their preparatory programme peers for support and friendship. As the year progressed, students made new friends and relied less on their programme peers.

Older children who had participated in the programme were also identified by one student as a valuable source of support. The diversity of the students in the preparatory programme was identified as a major strength of the programme overall. Students indicated that the exposure to other students from different areas and cultures had a positive impact on them.

### *Creative writing and study skills workshops*

One student indicated that they found the study skills work shop interesting at the time, but that they haven't used any of the skills that they were taught. Another student indicated that the study skills work shop was not helpful to them. Students indicated that they preferred to keep using the study methods that they used prior to attending the workshop to adopting new methods.

Two of the students indicated that they have difficulty with essays at their new school, and that the creative writing workshops they completed while in the preparatory programme did not furnish them with the skills they need to overcome this difficulty. Another student indicated that they did not find the creative writing workshop to be helpful as they have not used any of the techniques which were covered in either of the workshops. Students did not indicate a preference for one of the creative writing workshops over the other.

One student indicated that the additional content in the programme (creative writing and study skills workshops) focused mainly on English. The student would have preferred additional content which focused on mathematics, as this was the aspect which they thought they needed the most help with.

### *Breaks*

One student identified the break times between the first and second session as being unnecessarily long, sometimes up to one hour. The student suggested that a shorter break would still be sufficient time to rest, and it would allow them to finish earlier.

### *Mathematics*

One student indicated that they found the mathematics portion of the programme very helpful. They found the content to be highly relevant, and found their teacher to be highly competent at explaining the concepts in a way that they understood. They found the use of examples by the teacher to be particularly helpful in aiding understanding. The same student found the content covered in the programme useful, and some of the content covered in the programme was used in the first six months of high school. This student improved both their mathematics and English aggregates between June and December 2014.

Another student found the mathematics content useful, and found their notes from the preparatory programme particularly useful for revision when certain content was covered in high school. This student enjoyed the mathematics classes and appreciated the teaching style employed by the teacher.

One student found the mathematics portion to be unstimulating and indicated that the knowledge gained in the mathematics portion of the preparatory programme was very useful to them while they were in grade 7. However, they have not used this knowledge since they entered high school. The student did not identify the knowledge or skills taught in the preparatory programme as applicable to the new content covered in grade 8. This students' mathematics and English aggregates decreased between June and December 2014.

### *English*

Two of the students found the English portion of the preparatory programme to be helpful. They found it useful for strengthening their understanding of grammar and language. One of these students identified that some of the content covered in the programme was used in the first six months of high school and was able to provide the evaluator with specific examples. This student improved both their mathematics and English aggregates between June and December 2014.

One of the students indicated that the knowledge gained in the English portion of the preparatory programme was very useful to them while they were in grade 7. However, they have not used this knowledge since they entered high school. This students' mathematics and English aggregates decreased between June and December 2014. The student described the English portion of the programme as 'complicated', although they do indicate that it helped with their understanding overall.

Two of the students identified the 'breaking down' portion of the preparatory programme as useful. Students were able to apply this skill to their new academic work and found it useful in understanding new content.

One student indicated that their English teacher at their new school taught them grammar which directly contradicts what they were taught while in the preparatory programme.

One student thought that the English component of the programme could be improved upon by using more examples in order to demonstrate concepts, and also by explaining concepts in greater detail.

### *Natural science*

Two of the students identified the addition of natural science as a subject as a possible improvement to the preparatory programme. This subject was identified as being particularly challenging, and students described a substantial increase in difficulty in this subject once they had entered high school. Students indicated that they thought they would benefit substantially from this subject being included in the preparatory programme.

### *Internal assessments*

All of the students indicated that the internal tests were difficult, particularly the last test. One student indicated that some of the content had not been covered in sufficient detail for the internal assessments.

### *Homework*

One student identified the homework as a portion of the preparatory programme that could be improved upon. The student indicated that additional examples given along with the homework would aid their understanding and make the homework easier to complete on their own. They also thought that it would be beneficial if they were allowed to ask questions about the next week's homework while in the preparatory class.

Another student identified the amount of homework given as problematic. They found the homework given by the mathematics and English teachers to be challenging to complete in addition to the homework that they are given by the school that they attend.

### *Structure*

One student indicated that they would prefer it if there were a set structure to the preparatory programme sessions- mathematics in the morning session and English in the afternoon session. The student indicated that they found it easier to concentrate in the mornings, and preferred to do mathematics first as they found it to be the more challenging subject.

### *Summary*

Overall, students were very positive about the preparatory programme and all of the students indicated that they thought they would not have coped as well with the demands of high school without it. Students were able to retain and apply their knowledge from the programme to differing degrees. The social support provided by other programme recipients was identified as a key contributor to their ability to cope with high school.

Students identified the break times, homework and the structure of the programme as aspects which could be improved upon. Students also identified the natural science subject as particularly challenging and indicated that having it as a subject on the programme would be beneficial. The

internal assessments were identified as very challenging and there was concern about whether the content had been covered sufficiently before the testing sessions.

Students were ambivalent about the study skills workshop and were not able to provide any examples of how they had applied the knowledge gained in the workshop. Students enjoyed the creative writing workshops, but did not find that the workshops enabled them to address specific weaknesses in their own writing.

## **Outcome Evaluation**

### **4. Can the observed changes in learners' mathematics and English performance be attributed to the intervention?**

Two sets of pre and post-test report card scores were analysed in order to determine whether students had improved their academic performance. The pre-intervention score was obtained from students' term 2 report card generated by the school after their May/June examinations. The post-test score was obtained from their final report card after their November/December examinations. Participants ( $n = 47$ ) received a percentage score out of 100 for their mathematics and English aggregates.

The mathematics and English results were analysed separately. IBM SPSS Statistics 22 was used to generate descriptive statistics and perform parametric tests. Histograms, Q-Q plots, and the Kolmogorov-Smirnov test indicated that the distribution of the differences between pre- and post-tests for English,  $D(47) = 0.107$ ,  $p = .2$ , and mathematics,  $D(47) = 0.103$ ,  $p = .2$ , did not deviate significantly from normality.

Table 8 displays these results.



Table 8

*Pre and post scores for mathematics and English report card results*

English score ( <i>n</i> = 47)				Mathematics score ( <i>n</i> = 47)			
Participant	Pre-test score	Post-test score	Difference	Participant	Pre-test score	Post-test score	Difference
1	87	91	4	1	97	97	0
2	89	88	-1	2	91	93	2
3	77	77	0	3	83	82	-1
4	75	81	6	4	91	86	-5
5	75	80	5	5	91	90	-1
6	86	86	0	6	87	92	5
7	92	88	-4	7	94	79	-15
8	72	84	12	8	61	78	17
9	90	89	-1	9	91	94	3
10	58	77	19	10	63	65	2
11	94	93	-1	11	86	94	8
12	83	84	1	12	65	84	19
13	75	66	-9	13	55	69	14
14	89	88	-1	14	93	92	-1
15	89	92	3	15	86	99	13
16	55	63	8	16	49	66	17
17	84	87	3	17	76	76	0
18	80	82	2	18	96	96	0
19	74	83	9	19	64	80	16
20	88	86	-2	20	76	81	5
21	82	74	-8	21	83	87	4
22	81	81	0	22	92	87	-5
23	84	89	5	23	85	74	-11
24	79	88	9	24	62	65	3
25	82	80	-2	25	89	86	-3
26	75	75	0	26	69	74	5
27	64	64	0	27	84	79	-5
28	84	88	4	28	98	96	-2
29	83	82	-1	29	76	69	-10
30	75	63	-12	30	61	77	16
31	82	88	6	31	86	90	4
32	97	90	-7	32	82	93	11
33	64	63	-1	33	83	84	1
34	67	71	4	34	84	79	-5
35	92	89	-3	35	85	72	-13
36	89	85	-4	36	91	91	0
37	86	92	6	37	83	98	15
38	94	87	-7	38	69	76	7
39	87	93	6	39	94	91	-3
40	74	82	8	40	91	89	-2
41	63	63	0	41	68	74	6
42	74	65	-9	42	73	76	3
43	75	69	-6	43	49	64	15
44	78	73	-5	44	63	59	-4
45	71	79	8	45	59	62	3
46	92	85	-7	46	80	91	11
47	88	90	2	47	70	86	16
Total	3774	3813	39	Total	3704	3862	158
<i>M</i>	80.29	81.12	0.82	<i>M</i>	78.8	82.17	3.36
<i>SD</i>	9.78	9.27	6.13	<i>SD</i>	13.4	10.74	8.39

A paired-samples *t*- test was conducted to investigate the difference between pre-test English school examination performance ( $M = 80.29$ ,  $SE = 1.44$ ) and post-test performance ( $M = 81.12$ ,  $SE = 1.32$ ). This difference, 0.82, 95% BCa CI [-0.97, 2.62] was not significant,  $t(46) = 0.92$ ,  $p = .358$ . There was a small effect,  $d = 0.134$ .

Another second paired –samples *t*- test was conducted to investigate the difference between pre-test mathematics academic performance ( $M = 78.8$ ,  $SE = 1.95$ ) and post-test performance ( $M = 82.17$ ,  $SE = 1.54$ ). This difference, 3.36, 95% BCa CI [0.89, 5.82] was significant,  $t(46) = 2.74$ ,  $p = .009$ . There was a medium effect,  $d = 0.37$ .

##### **5. Did the students in the 2014 cohort in each of the streams benefit equally from the intervention? Did some streams benefit more (or less) than others?**

Performance scores on the three assessments administered by the programme were analysed in order to determine whether the three streams scored significantly differently from one another at various points in the programme. These assessments were administered before the programme began, approximately mid-way through the programme, and right at the end of the programme. All students, irrespective of their stream placement, wrote the same tests. Mathematics and English scores were analysed separately. IBM SPSS Statistics 22 was used in order to generate descriptive statistics and perform parametric and non-parametric tests.

Data for English were analysed using a mixed-design ANOVA with a within-subjects factor of assessment score (initial assessment, mid-term assessment and final assessment scores) and a between-subjects factor of stream (bottom stream, mid-stream and top stream).

Mauchly's test indicated that the assumption of sphericity had been upheld for English scores,  $\chi^2(2) = 0.97$ ,  $p = .639$ . There was a significant effect of stream placement on performance for the assessments,  $F(2,44) = 29.81$ ,  $p < .001$ ,  $\omega = .758$ . Planned contrasts indicate that the top stream scored significantly more than the mid-stream,  $F(2,44) = 11.41$ ,  $p < .001$ ,  $r = .453$ , and the mid-stream scored significantly more than the bottom stream,  $F(2,44) = 6.4$ ,  $p = .01$ ,  $r = .356$ .

There was a significant effect of the assessment type on performance,  $F(2, 88) = 69.1$ ,  $p < .001$ ,  $\omega = .781$ . Planned contrasts indicate that students scored significantly higher on the mid-term test than they did in the final assessment,  $F(1,44) = 99.28$ ,  $p < .001$ ,  $r = .831$ , and that initial assessment scores did not differ significantly from final assessment scores,  $F(1,44) = 0.013$ ,  $p = .909$ ,  $r = .017$ . Figure 8 displays the mean scores for each stream on the three assessments for English.

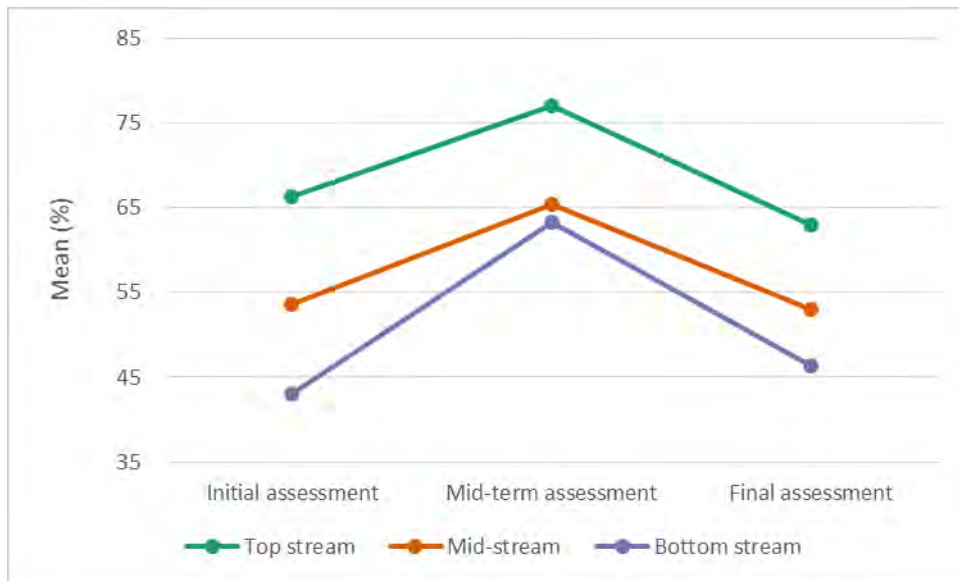


Figure 8. Stream performance on English internal assessments

Data for mathematics were analysed using a mixed-design ANOVA with a within-subjects factor of assessment score (initial assessment, mid-term assessment and final assessment scores) and a between-subjects factor of stream (bottom stream, mid-stream and top stream).

Mauchly's test indicated that the assumption of sphericity had been upheld for mathematics scores,  $\chi^2(2) = 0.935, p = .238$ . There was a significant effect of stream placement on performance for the assessments,  $F(2,44) = 25.71, p < .001, \omega = .734$ . Planned contrasts indicate that the top stream scored significantly more than the mid-stream,  $F(2,44) = 9.88, p < .001, r = .428$ , and the mid-stream scored significantly more than the bottom stream,  $F(2,44) = 7.15, p = .006, r = .373$ .

There was a significant effect of the assessment type on performance,  $F(2, 88) = 41.81, p < .001, \omega = .697$ . Planned contrasts indicate that students scored significantly higher on the mid-term test than they did in the initial assessment,  $F(1,44) = 18.76, p < .001, r = .546$ , and significantly lower on the final assessment than the mid-term assessment,  $F(1,44) = 18.94, p < .001, r = .548$ . Figure 9 displays the mean scores for each stream on the three assessments for mathematics.

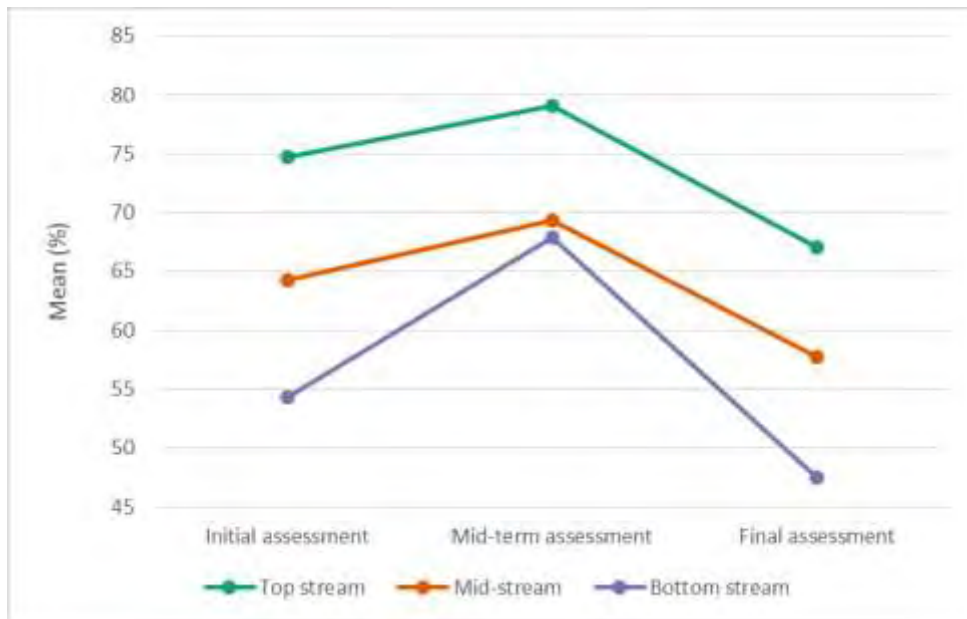


Figure 9. Stream performance on mathematics internal assessments

Two sets of pre and post-test report card scores were aggregated in order to determine whether students within each of the streams had improved their academic performance. The pre-intervention score was obtained from students' term 2 report card generated by the school after their May/June examinations. The post-test score was obtained from their final report card after their November/December examinations. The mathematics and English results were analysed separately. IBM SPSS Statistics 22 was used to generate descriptive statistics and perform parametric tests.

Kolmogorov-Smirnov tests indicated that the distribution of the differences between English pre- and post-tests for the top stream,  $D(17) = 0.168$ ,  $p = .2$ , mid-stream,  $D(16) = 0.115$ ,  $p = .2$ , and bottom stream,  $D(14) = 0.222$ ,  $p = .06$ , did not deviate significantly from normality.

A paired-samples  $t$ -test was conducted to investigate the difference between pre-test English school report card scores ( $M = 83$ ,  $SE = 1.8$ ) and post-test report card scores ( $M = 82.7$ ,  $SE = 1.81$ ) for the top stream. This difference,  $-0.3$ , 95% BCa CI  $[-1.27, 1.86]$  was not significant,  $t(16) = 0.39$ ,  $p = .697$ . There was a small effect,  $d = 0.151$ .

A paired-samples  $t$ -test was conducted to investigate the difference between pre-test English school report card scores ( $M = 82.18$ ,  $SE = 2.82$ ) and post-test report card scores ( $M = 84$ ,  $SE = 2.08$ ) for the mid-stream. This difference,  $1.82$ , 95% BCa CI  $[-5.69, 2.06]$  was not significant,  $t(15) = 0.99$ ,  $p = .335$ . There was a small effect,  $d = 0.249$ .

A paired-samples  $t$ - test was conducted to investigate the difference between pre-test English school report card scores ( $M= 74.85, SE = 2.34$ ) and post-test report card scores ( $M = 75.92, SE = 2.83$ ) for the bottom stream. This difference, 1.07, 95% BCa CI [-5.44, 3.3] was not significant,  $t(13) = 0.92, p = .606$ . There was a small effect,  $d = 0.256$ .

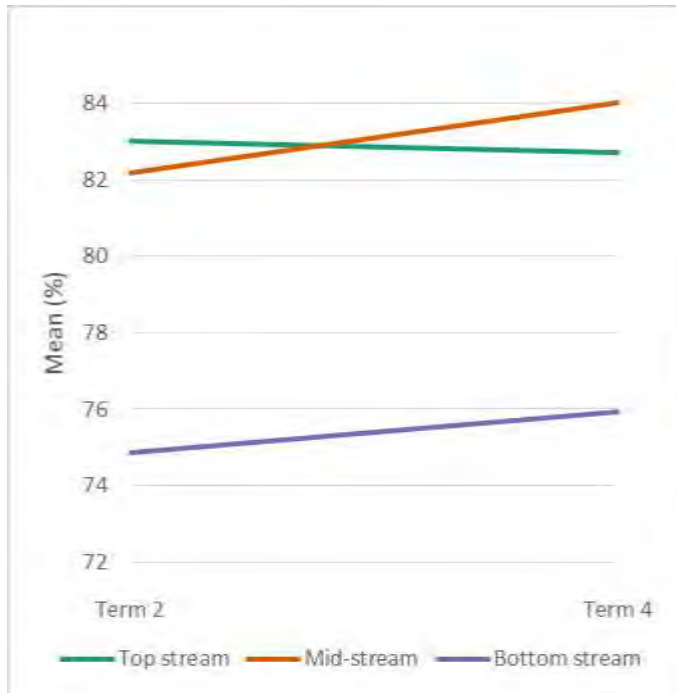


Figure 10. Stream performance for term 2 and term 4 English report card marks

Kolmogorov-Smirnov tests indicated that the distribution of the differences between mathematics pre- and post-tests for the top stream,  $D(18) = 0.165, p = .2$ , mid-stream,  $D(17) = 0.136, p = .2$ , and bottom stream,  $D(12) = 0.153, p = .2$ , did not deviate significantly from normality.

A paired-samples  $t$ - test was conducted to investigate the difference between pre-test mathematics report card scores ( $M= 83.33, SE = 2.38$ ) and post-test report card scores ( $M = 85, SE = 2.14$ ) for the top stream. This difference, 1.67, 95% BCa CI [-6.3, 2.96] was not significant,  $t(17) = 0.75, p = .458$ . There was a small effect,  $d = 0.205$ .

A paired-samples  $t$ - test was conducted to investigate the difference between pre-test mathematics report card scores ( $M= 78.64, SE = 3.5$ ) and post-test report card scores ( $M = 82.47, SE = 2.78$ ) for the mid-stream. This difference, 3.83, 95% BCa CI [-7.12, -0.52] was significant,  $t(16) = 2.45, p = .026$ . There was a small effect,  $d = 0.364$ .

A paired-samples *t*-test was conducted to investigate the difference between pre-test mathematics report card scores ( $M = 72.25$ ,  $SE = 4.17$ ) and post-test report card scores ( $M = 77.5$ ,  $SE = 3.26$ ) for the bottom stream. This difference, 5.25, 95% BCa CI [-11.33, 0.83] was not significant,  $t(11) = 1.9$ ,  $p = .08$ . There was a small effect,  $d = 0.393$ .

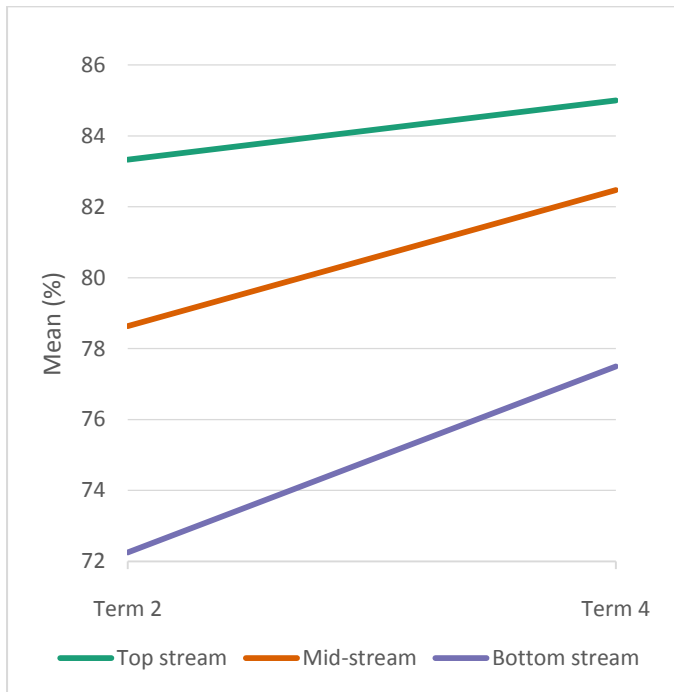


Figure 11. Stream performance for term 2 and term 4 mathematics report card marks

## 6. Can the observed improvements in cognitive reasoning ability be attributed to the intervention?

CAT pre- and post-test scores were obtained for the 2014 cohort which completed the programme last year ( $n = 47$ ). The pre-test was completed in April 2014 as a part of the selection process for the programme. The post-test was completed approximately 1 year and 6 months later. IBM Statistics SPSS 22 was used in order to generate descriptive statistics and perform both parametric and non-parametric tests.

Kolmogorov-Smirnov test indicated that the distribution of the differences between pre- and post-tests for the mean standardised age scores, and the verbal and non-verbal subscales did not deviate significantly from normality. Paired-samples *t*-tests were conducted in order to determine whether the pre- and post-test scores differed significantly for the mean standardised score and the subscales.

A paired-samples *t*- test was conducted to investigate the difference between pre-test mean CAT standardised age scores ( $M = 105.38$ ,  $SE = 0.85$ ) and post-test mean scores ( $M = 111.27$ ,  $SE = 1.12$ ). This difference, 5.89, 95% BCa CI [4.23, 7.54] was significant,  $t(46) = 7.16$ ,  $p < .001$ . There was a large effect,  $d = 0.725$ .

A paired-samples *t*- test was conducted to investigate the difference between pre-test verbal standardised age scores ( $M = 102.89$ ,  $SE = 0.08$ ) and post-test scores ( $M = 109.25$ ,  $SE = 0.1$ ). This difference, 6.36, 95% BCa CI [4.44, 8.61] was significant,  $t(46) = 6.07$ ,  $p < .001$ . There was a medium effect,  $d = 0.666$ .

A paired-samples *t*- test was conducted to investigate the difference between pre-test non-verbal standardised age scores ( $M = 104.21$ ,  $SE = 0.004$ ) and post-test scores ( $M = 111.19$ ,  $SE = 0.003$ ). This difference, 6.97, 95% BCa CI [4.4, 9.55] was significant,  $t(46) = 5.45$ ,  $p < .001$ . There was a medium effect,  $d = 0.626$ .

According to the Kolmogorov-Smirnov test, the distribution of the quantitative subscale did deviate significantly from normality,  $D(47) = 0.243$ ,  $p < .001$ . A non-parametric test was used in order to investigate this variable.

A Wilcoxon-signed rank test was conducted to investigate the difference between pre-test quantitative standardised age scores and post-test scores. Students scored significantly higher on the quantitative subscale at post-test ( $Mdn = 113$ ) than they did at pre-test ( $Mdn = 110$ ),  $T = 12$ ,  $p = .003$ ,  $r = -.307$ .

## **7. Did the students in the 2015 cohort show improvements in their self-efficacy scores?**

SEQ-C pre- and post-test scores were obtained for the 2015 cohort which completed the programme this year. One student was excluded from the analysis as only one page of the two page questionnaire was completed. The final sample size was 49 students. The pre-test was completed in June 2015, and the post-test was completed approximately 6 months later at the end of October.

If a question was left blank in the pre-test, a score of 'sometimes' was assumed. The same procedure was followed for the post-test, except if the question had been previously answered in the pre-test. In this case, the answer was assumed to be unchanged from what had been previously indicated in the pre-test.

The total SEQ-C scores were analysed, as well as each of the three subscales for the SEQ-C, specifically the academic, social and emotional self-efficacy subscales. IBM Statistics SPSS 22 was used in order to generate descriptive statistics and perform both parametric and non-parametric tests.

Kolmogorov-Smirnov tests indicated that the distribution of the differences between pre- and post-tests for the total SEQ-C scores and the emotional self-efficacy subscale did not deviate significantly from normality. Paired-samples *t*-tests were conducted in order to determine whether the pre- and post-test scores differed significantly for the emotional self-efficacy subscale and the total SEQ-C scores.

A paired-samples *t*-test was conducted to investigate the difference between pre-test total SEQ-C scores ( $M = 95.97$ ,  $SE = 1.42$ ) and post-test scores ( $M = 97.1$ ,  $SE = 1.34$ ). This difference, 1.13, 95% BCa CI [-1.03, 3.27] was not significant,  $t(48) = 1.046$ ,  $p = .301$ . There was a small effect,  $d = 0.148$ .

A paired-samples *t*-test was conducted to investigate the difference between pre-test emotional self-efficacy subscale scores ( $M = 29.2$ ,  $SE = 0.74$ ) and post-test scores ( $M = 30.24$ ,  $SE = 0.61$ ). This difference, 1.04, 95% BCa CI [-0.27, 2.35] was not significant,  $t(48) = 1.587$ ,  $p = .119$ . There was a small effect,  $d = 0.223$ .

According to the Kolmogorov-Smirnov test, the distribution of the academic self-efficacy subscale,  $D(49) = 0.136$ ,  $p = .023$ , and the social self-efficacy subscale  $D(49) = 0.15$ ,  $p = .007$ , did deviate significantly from normality. Non-parametric tests were used in order to investigate these subscales.

A Wilcoxon-signed rank test was conducted to investigate the difference between pre-test academic self-efficacy subscale scores and post-test scores. Students scored significantly lower on the academic self-efficacy subscale at post-test ( $Mdn = 34$ ) than they did at pre-test ( $Mdn = 35$ ),  $T = 330$ ,  $p = .033$ ,  $r = -.215$ .

A Wilcoxon-signed rank test was conducted to investigate the difference between pre-test social self-efficacy subscale scores and post-test scores. Students scored significantly higher on the social self-efficacy subscale at post-test ( $Mdn = 33$ ) than they did at pre-test ( $Mdn = 33$ ),  $T = 516$ ,  $p = .034$ ,  $r = .214$ .



## DISCUSSION

This discussion chapter follows the same order as the evaluation questions presented in the results chapter.

### Process Evaluation

#### 1. Service Delivery:

##### **Was the programme delivered as intended for the intended duration?**

Based on the discussion from the interview conducted with the programme co-ordinator and the focus group conducted with the other teachers, the majority of the 2014 iteration of the programme was implemented as planned. However, there were changes made to the structure of the sessions which were delivered, and an extra creative writing workshop was delivered in addition to what was planned. Teachers also encountered problems with the smart boards which were in the classrooms, as these were not in working order or had been damaged.

Structural changes include the introduction of “block days” (days focusing entirely on mathematics or English), rather than the planned structure of a 50/50 split between mathematics and English on a given day. This was done in order to accommodate some of the teachers who would be away for a portion of the preparatory programme. Although the structure of the sessions changed, mathematics and English still received the planned amount of teaching time by the end of the programme.

An extra creative writing workshop was delivered in addition to the one workshop which was planned. The programme co-ordinator was not satisfied with the quality of the planned workshop after it had been completed, and a second workshop was run by a different facilitator in order to maintain the quality of the intervention.

Although the smart boards had been damaged or were not set up correctly, the teachers identified this as more of a minor annoyance than a serious impediment to teaching. All of the teachers were able to find another way to present the relevant material, and lessons went ahead regardless of whether the smart board was working or not. The smart boards have been repaired and were set up correctly for the 2015 iteration of the programme, so this problem has been rectified.

A process evaluation has significant diagnostic value, as it can help to ascertain why a programme did not achieve the desired outcomes. If the expected programme activities were not delivered, then the benefits of taking part in those activities could not occur (Rossi *et al*, 2004). This is known as implementation failure. In the case of the preparatory programme, there were changes to the

programme activities which were delivered, as well as to their structure. However, the amount of teaching time allocated to mathematics or English remained the same, and additional workshops were added in order to maintain the quality of the workshop. It can be concluded that the desired activities were delivered to the programme participants and that these were of sufficient quality to satisfy the programme co-ordinator. If the preparatory programme fails to achieve the desired outcomes it is unlikely to be due to implementation failure.

## **2. Organisational support:**

### **Is the CAT instrument that the programme is using a useful tool for informing selection?**

Programme staff indicated that the CAT3 was identified as a useful measure for informing the selection process as it was not heavily dependent on language, did not rely on past academic learning, and would give an indication of the academic potential of individual students.

Students selected for the programme did not differ significantly from students not selected for the programme in terms of their performance on the CAT instrument. For the 2014 cohort, the CAT instrument was unable to differentiate the students that were selected from those who were not selected. This was true of the mean standardised age score, and each of the three subscales, verbal, nonverbal and quantitative. These results suggest that the data generated by the CAT3 instrument is not currently being used to meaningfully inform the selection process.

The ability of the CAT instrument to predict academic performance is well-established for developed countries with well-resourced schools (Deary, Strand, Smith & Fernandes, 2007). In a study which included over 70 000 British school children, there was a strong correlation between CAT scores and academic performance, and CAT scores accounted for 58.6% of the variation in scores for mathematics and 48% of the variation in scores for English (Deary *et al*, 2007). This study was based on the CAT2E, which is an earlier edition of the CAT than was used in this evaluation. However, the CAT3 is structurally similar, and CAT2E scores can be converted into CAT 3 scores. As with the CAT3, the CAT2E consists of three subscales (verbal, nonverbal, and quantitative reasoning), and a mean standardised age score is calculated.

The CAT3 was able to predict the final mathematics aggregates of students who were selected for the programme. The regression equation outlined in the results chapter could be used to predict applicant's final mathematics aggregate for grade 7 and this information could potentially be used to inform selection. However, there are several caveats which are important to note.

The regression equation makes use of the quantitative subscale score of the CAT. The company which distributes the CAT does not currently allow for programmes to administer only a single subscale- the entire CAT package has to be purchased. This would represent a significant expense for the programme, and only a third of the data is used in the regression equation.

It is also important to note that this analysis was based on the CAT3, which is currently being phased out by the company which distributes it. From the next testing cycle the programme will need to use the CAT4, which consists of four subscales instead of three. The CAT4 includes a visuospatial reasoning subscale in addition to the verbal, nonverbal and quantitative subscales used by the CAT3 (GL Education, 2015). Although it has been indicated by the distribution company that the CAT4 and CAT3 mean standardised age scores are comparable to one another, it is unclear to what degree the three subscales of the CAT3 have been altered in the CAT4 and whether the CAT4 would still be able to predict the mathematics aggregate based on the quantitative subscale score.

It is likely that factors other than performance on this measure are having a large impact on who gets selected for the programme. The selection criteria outlined by the programme state that students need to obtain a minimum aggregate of 65% for their final grade 6 results, and they also need to demonstrate clear financial need. With the current data which are available, it is not possible to determine to what extent these other factors influenced programme selection, and whether they would successfully differentiate between those selected and those not selected.

In lieu of these results, the programme has two viable options for streamlining the selection process.

1. Use the CAT scores to inform the selection process.

The CAT scores are not currently being used in a manner which is useful to the selection process, as the students who are selected and who are not selected do not differ significantly from one another in terms of their scores on the CAT instrument. The CAT score could either be assigned a greater weighting, and CAT scores given greater consideration when considering which students to select for the programme, or the CAT instrument could be applied only to students which have met the criteria for clear financial need and the grade 6 65% aggregate. The CAT score could then be used to make the final selection for the programme, with students scoring higher on the CAT gaining preference over students which score lower.

2. Eliminate the CAT from the selection process.

Another option is to focus more on other selection criteria, and eliminate the CAT entirely from the selection procedure. Programme staff could rely on other selection criteria, such as

financial need and previous academic performance in order to determine which students receive the programme. This option would represent a significant reduction in costs for the programme.

Finally, the regression equation which was developed can only account for a relatively small amount of the variation in mathematics aggregates (13%), leaving 87% of the variation unexplained. This is worth noting, as previous research indicates that CAT scores can predict up to approximately 60% of the variation in mathematics or English scores (Deary *et al.*, 2007) in developed countries. In this sample, the CAT was unable to predict the English aggregates at all, and only a relatively small portion of the variance in mathematics scores was explained. This difference may be due to the nature of the population that the CAT has been standardised on, which includes European and North American students. The CAT4 will include additional testing options which were not available for the CAT3. There will be an option for programme staff to select “if-challenged” grades, rather than standard grades. The “if-challenged” grades have been standardised for use in schools which are under-resourced (J. Neill, personal communication, May 11, 2015). If the programme decides to continue using the CAT, the new grades may be able to account for more for the variation in scores as they were standardised on a population more similar to the sample of interest.

### **3. Service utilisation:**

#### **What aspects of the preparatory programme did the 2014 cohort find more/less effective?**

All of the students indicated that they enjoyed aspects of the creative writing workshops. However, the degree to which they found the content useful varied. Several creative writing instruction practices have been demonstrated to produce a significant improvement in how well students write. These include explicit teaching of writing processes (*e.g.* strategy instruction, text structure instruction, creative/imagery instruction) and interventions focused on supporting students’ writing (*e.g.* pre-writing activities, peer assistance, writing assessment)(Graham, McKeown, Kiuahara & Harris, 2012). Only interventions found to focus on grammar instruction did not produce significant improvements. Based on this research, a creative writing workshop is a logical component to include in an academic intervention. However, two of the students indicated that they still had trouble with writing essays after completing the workshops, and that the workshop had not addressed the particular challenges that they faced in their own writing. Without further knowledge of the content of the workshops and the teaching approaches used, the evaluator is unable to comment on why the students still had difficulty with essays. The evaluator recommends a review of the approaches used

in the creative writing workshop in order to refine the presentation of this component of the programme and ensure that an evidence-based approach is used.

None of the students interviewed indicated that they found the study skills workshop useful or had applied any of the knowledge gained from the workshop. As the link between effective study skills and academic performance has been well established by research (Deshler & Schumaker, 1993), a study skills workshop is a logical component to include in an academic intervention. However, without more knowledge of the exact content or approaches used in the workshop it is difficult to ascertain why students did not find it useful. There are several varieties of strategies (*e.g.* repetition/rehearsal-based, procedural/organisation-based, cognitive-based, or metacognitive-based) and approaches (*e.g.* Strategies Intervention Model, or the Good Information Processing Approach) which could be used to impart study skills to students (Gettinger & Seibert, 2002). Each of these strategies has its own strengths and weaknesses, and the manner in which the content is taught is as important as the type of strategy used. A review of the content and approaches used in the study skills workshop would be useful in aligning the workshop with student needs.

All of the students which were interviewed indicated that they found the preparatory programme useful. However, there were clear differences between the student whose mathematics and English performance improved, and the students whose performance decreased. The student which improved their performance was able to provide clear and specific examples of content or skills covered during the preparatory programme which they were able to apply in their first year of high school. In contrast, the other students could not recall which content had been covered, or indicated that this content was only useful to them while they were in grade 7. This is particularly salient, as the student whose marks improved and the student who indicated that the preparatory programme content was only useful for primary school now attend the same high school, and have been placed in the same class for mathematics. Therefore, both students received the same content in high school, but only one of the students was able to recall and apply the knowledge from the preparatory programme to the new situation. In light of this new information, an amendment to the impact theory depicted in chapter 1 of this report is proposed.

An impact theory describes a cause and effect sequence in which certain programme activities are linked with the expected outcomes (Rossi *et al*, 2004). Figure 12 depicts an amended impact theory for the knowledge portion of the preparatory programme based on the results of the interviews.

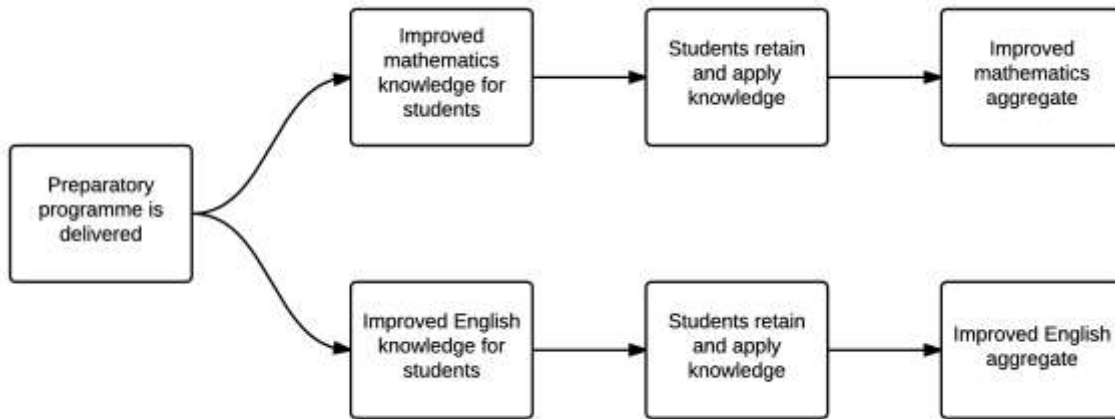


Figure 12. Amended impact theory for knowledge component of the preparatory programme

Differences in performance may be due to any number of factors, including differences in levels of motivation, locus of control, self-efficacy or metacognitive approaches to learning. A more systematic analysis of the differences between students who improve their academic performance and those whose academic performance decreases is recommended. Gaining a better understanding of these differences would enable programme staff to formulate targeted interventions for students who are not obtaining the expected benefits from the preparatory programme, or to specifically target students during the selection process which would benefit most from the current formulation of the intervention.

### Outcome Evaluation

#### 4. Can the observed changes in learners' mathematics and English performance be attributed to the intervention?

Examination of the term 2 and term 4 report cards indicated that students did not significantly improve their English aggregates, but did improve their mathematics aggregates. Mathematics aggregates improved by an average of 3.3%.

As the programme was mostly implemented as planned (see section pertaining to evaluation question 1), this makes the failure to achieve the desired outcomes in English unlikely to be due to implementation failure.

It is possible that the English aggregates did not improve due to difficulties with language- although data regarding the home and additional languages spoken by the students were unavailable. It is also possible that students did not receive a sufficient dose of this component of the intervention.

Previous research indicates that the largest effects for OST programmes targeting English academic achievement were observed when students had received between 44 and 84 hours of teaching time (Lauer *et al*, 2004). The programme participants only received approximately 30 hours of teaching time for English, and this may have been insufficient to bring about the desired changes.

The nature of the sample adds to the complexity of interpretation. The students attend a wide variety of primary schools, and there are no data available about what forms of assessment those schools used and what level of difficulty the tasks and examinations were set at. As the level of difficulty for the assessment tasks and examinations were not standardised, it makes it difficult to compare the scores of students across different primary schools.

Without a control group, it is difficult to interpret these results meaningfully. It may be possible that if the intervention had not occurred, the aggregates for mathematics and English could have decreased, remained the same, or increased. The marks are most likely to have decreased due to regression to the mean, as the group performed very well on the pre-test, with an average score of 80.29% for English and 78.8% for mathematics. Regression to the mean is the tendency of extreme scores to regress to the mean on subsequent measurements. Due to the cohort's high performance on the pre-test, it is most plausible that without an intervention the marks on the post-test would be lower and closer to the population average. The slight increase in the cohort's English average (0.83%) and the significant increase in the mathematics average (3.3%) are both promising in terms of the programme effect. A ceiling effect could also account for the modest increases in the English and mathematics aggregates. It is plausible to suggest that as the students are already performing particularly well academically, significant improvement is improbable. Although unlikely, it is possible that the intervention had a negative effect on students' academic outcomes, but their innate abilities were sufficient to counteract this and bring about a net positive effect.

Despite these observations, the current evaluation design does not allow for the increase in results to be attributed to participation in the programme. As the students who are selected for the programme receive a scholarship and are aware that their academic performance is monitored by programme staff, bias may be introduced into the results due to the Hawthorne effect. This occurs when the awareness of being observed brings about a change or improvement in behaviour. The observed positive effect may be due to the effects of being monitored, rather than as a result of the content of the preparatory programme itself.

**5. Did the students in the 2014 cohort in each of the streams benefit equally from the intervention? Did some streams benefit more (or less) than others?**

Results indicate that the streaming system is working as planned, as each of the streams scored significantly differently from one another on each of the internal assessments for mathematics and English, except for on the second internal assessment. This result was to be expected, as the programme co-ordinator indicated that the second assessment was easier than the initial and final assessment. This accounts for the initial rise in performance across all three streams for the mid-term assessment, and the drop in performance for the final assessment. These results indicate that students have been placed into the correct streams and provides evidence for the continued use of the streaming system.

It is recommended for future iterations of the programme that the three assessments be of a standardised difficulty across all three internal assessments in order to allow for greater insight into student progress. If both the second assessment and the final assessment consisted only of content covered by the programme it would make student results more comparable than if content not covered by the programme were included. Direct comparison of student scores on the final assessment is not currently very meaningful as students may have covered material to varying degrees in their respective primary schools.

Analysis of the English report card results disaggregated by stream indicate that none of the streams significantly improved their English aggregates. Based on the gradient of the slopes it appears as though the middle stream benefitted the most from the intervention. It is worth noting that the top English stream was the only one of the six streams (three mathematics and three English) which were analysed which decreased their overall aggregate as a group. It is recommended that data from previous iterations of the programme be analysed to determine whether this is a trend in the data.

Analysis of the mathematics report card result disaggregated by stream indicate that only the mid-stream group significantly improved their mathematics aggregate. The bottom stream improvement was borderline significant, and the top stream also showed an improvement. This finding is contrary to previous research, which indicates that students in the top stream (J. Kulik & C. Kulik, 1992) or bottom stream (Ross *et al*, 1995, Gullatt & Jan, 2002) typically benefit more than students in the mid-stream.



## **6. Can the observed changes in the learners' cognitive reasoning ability be attributed to the intervention?**

Students from the 2014 cohort showed significant improvements in their cognitive reasoning ability. There was a significant increase in the scores on all three of the subscales and the standardised age score between April 2014 and October 2015. In practical terms, the CAT scores increased by between 4 to 6 points on each of the subscales on average.

Previous research on the CAT has demonstrated that CAT scores should remain stable for up to three years after testing (Strand, 2004). CAT scores have been demonstrated to be highly reliable, with internal consistency estimates averaging .94 for verbal reasoning, .90 for quantitative reasoning, and .92 for non-verbal reasoning (Thorndike, Hagen & France, 1986). Test-retest correlations are also high, ranging from between .76 for the non-verbal subscale and .89 for the overall CAT score (Strand, 2004). As students were re-tested within 1 year and 6 months of their initial CAT testing, and taking the stability and reliability of the CAT into account, it makes a maturation effect unlikely. This falls within the three year period during which test scores should have remained stable. The large increases in cognitive ability is not likely to be due to a testing effect, as the CAT has been demonstrated to be a highly reliable measure.

This is reasonable evidence that the programme has had an effect on the cognitive reasoning ability of the students. However, it is not possible to separate the effects of the preparatory programme from the historical effects of the scholarship programme. By the time the post-test took place, students had been attending their new schools for approximately nine months, and it is not possible to determine whether the improvement in cognitive reasoning ability is due to participation in the preparatory programme, or attendance at their scholarship high school.

## **7. Did the students in the 2015 cohort show improvements in their self-efficacy scores?**

Total self-efficacy scores and scores on the emotional self-efficacy subscale did not change significantly between the pre- and post-test. Scores on the social self-efficacy subscale were significantly higher at post-test, whereas scores on the academic self-efficacy subscale were significantly lower at post-test. As efficacy beliefs have been found to be an important moderator in developmental outcomes for adolescents (Allen, Leadbeater & Aber, 1994), the observed decreases in academic self-efficacy may seem to be cause for concern.

As students enter a new academic environment they re-evaluate their own perceived competency in comparison to their new group of peers (Eccles & Midgley, 1989). Research suggests that their

ratings of competence in mathematics and English tend to decrease as children age, as more emphasis is placed on relative ability in higher grade levels and in competitive environments (Anderman & Midgley, 1997). The preparatory programme is a new academic environment for the students, and student evaluations of their academic competence may have decreased as a result of exposure to their gifted peers. Students with high academic ability, such as the students which are selected for the preparatory programme, are particularly vulnerable to decreases in academic self-efficacy (Wigfield, Eccles, Mac Iver, Reuman & Midgley, 1991). As a decrease in academic self-efficacy appears to occur naturally as adolescents age, the average decrease of 0.79 does not warrant concern or further intervention. The average academic self-efficacy for the students was still within the range which could be considered to be high, with an average score of approximately 33 out of a possible 40.

Social self-efficacy is related to personal beliefs about the ability to perform the specific behaviours which underlie personal relationships (Connelly, 1989). The significant increase in the social self-efficacy scores of the students is challenging to explain as there were no programme activities specifically focused on improving this aspect of development. However, social self-efficacy has been found to be positively correlated to peer and family support (McFarlane, Bellissimo & Norman, 1995), and it is possible that the additional support provided by programme staff and peers in the preparatory programme influenced the positive development of this aspect of self-efficacy. Students which were interviewed (see section pertaining to evaluation question 3) all emphasized that the support provided by their programme peers had been a key factor in their adjustment to high school and that their peers formed a vital component of their social network.

It is difficult to speculate about what would have occurred in the absence of the programme, as there was no control group for comparison and there is a dearth of longitudinal research relating to the development of social self-efficacy in adolescents. It is also possible that this was due to a maturation effect, and that this change would have occurred naturally over time regardless of programme participation.

### **Limitations**

The central limitation of this evaluation is the lack of a comparison group which did not receive the programme. The evaluator is therefore unable to attribute any of the effects discussed in the results section of this report to participation in the programme. Thus, it is possible that changes between repeated measures on participants were due to maturation effects, regression to the mean, the

Hawthorne effect, or another historical event which influenced the outcomes of the programme participants. Maturation effects are naturally occurring changes to the participants as they age. Regression to the mean is the tendency of extreme scores to be closer to average on subsequent measures. The Hawthorne effect occurs when participants' behaviour is modified or improved due to the fact that they are being observed or monitored. Historical events, like other community interventions or increased funding for schools, could also produce changes in the participants. Statistical tests, like the *t*-tests used in many of the analyses in this evaluation, can only indicate that there is a significant difference between pre- and post-assessment scores. However, *t*-tests cannot by themselves demonstrate causality. Theoretically, maturation effects, regression to the mean, the Hawthorne effect, and historical effects would have the same impact on the intervention and the comparison group if the groups were comparable at baseline. This would allow the evaluation to make determinations about the impact of the programme on the participants. The lack of a control group limited the strength of the evaluation design and the resultant strength of any statements made about the programme.

Reliance on student report cards for evaluation question four limited the strength of any resultant statements. Although all of the students were assessed on the same subjects (specifically mathematics and English), the specific assessment methods would differ widely from school to school. Different content would be covered for the pre- and post-assessments and are not necessarily of a standardised difficulty. This limits the ability to compare students from different schools to one another, as well as limiting the value of comparing a single students' pre- and post-assessment scores.

The generalisability of the results of the evaluation are limited due to the nature of the sample. The sample consisted of students who already perform very well academically, and these kinds of students may be more motivated than their average peers to improve their academic performance over time. Therefore, the results of the evaluation cannot be reliably generalised beyond the population the sample was taken from.

Post-hoc power analyses revealed that several aspects of the outcome evaluation were under-powered due to small effect sizes in combination with small samples sizes. As a result, the evaluation may have failed to detect a significant effect that was present. All of the power calculations were conducted using the Power and Sample Size calculator. Details of the post-hoc power analyses are presented in Appendix D. This limitation could be potentially be overcome in future evaluations by combining data from several iterations of the programme.

Despite these limitations, this evaluation should be able to offer meaningful recommendations which would enable a more rigorous evaluation in the future if acted upon.

## **Conclusion**

Notwithstanding the lack of a comparison group, this evaluation was still able to offer compelling evidence that the preparatory programme shows promise in terms of its expected outcomes. Students that attended the programme showed a significant improvement in cognitive reasoning ability, social self-efficacy, and their mathematics aggregates. Although the design of this evaluation does not allow for these changes to be attributed directly to participation in the programme, these results are still encouraging in terms of the potential programme impact and at the very least they are an indication that the programme is not harming its participants. This evaluation also demonstrated that the programme is well implemented, with a focus on maintaining the quality of the intervention. A more robust evaluation, with a comparison group and a larger sample size may be able to demonstrate that the observed positive effects are due only to participation in the preparatory programme.

The evaluator makes the following key recommendations:

1. Use the data generated by the CAT as part of the selection process. The programme could switch to the CAT4 in the next testing cycle and then use the results to determine which students get the programme once all of the other criteria have been considered. If this approach is used it is recommended that regression analyses are conducted on the CAT4 scores and its subscales with students' mathematics and English aggregates for term 4 in grade 7. This could help to determine whether the CAT4 offers further utility for future selection processes by enabling programme staff to predict the mathematics and English aggregates of their participants based on their CAT scores. OR
2. Discontinue the use of the CAT entirely and focus on other selection criteria. The programme could focus on their other current selection criteria. This currently includes previous academic performance and financial need.
3. Review the content and method of presentation of both the creative writing workshops and the study skills workshop in order to ensure that it is evidence-based and in line with participant needs.
4. Standardise the difficulty of the internal assessments administered by programme staff in order to make the results comparable to one another. This will enable programme staff to

monitor student progress. Ensuring that the content of the second and final assessments consist only of the material covered by the programme will give programme staff a better indication of student retention of the material and allow direct comparison of student scores in a meaningful way.

5. The programme may be able to determine which students would benefit most from the intervention they offer by conducting further research. A more systematic analysis of the differences between students who improve their academic performance and those whose academic performance decreases is recommended. This could include research on the potential differences in levels of motivation, locus of control, self-efficacy or metacognitive approaches to learning. This would enable the programme to streamline the selection process and invest its resources in the students which would benefit most.

Overall, the preparatory programme is associated with positive student outcomes and it may have a positive impact on how prepared they are for the challenges of high school.

## REFERENCES

- Allen, J.P., Leadbeater, B.J., & Aber, J.L. (1994). The development of problem behaviour syndromes in at-risk adolescents. *Development and Psychopathology, 6*, 323-342.
- Anderman, E.M., & Midgley, C. (1997). Changes in achievement goal orientations, perceived academic competence and grades across the transition to middle-level schools. *Contemporary Educational Psychology, 22*, 269-298.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioural change. *Psychological Review, 84*, 191-215.
- Bandura, A. (1982). Self-efficacy mechanisms in human agency. *American Psychologist, 37*, 122-147.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Calabrese, R.L. (1987). Adolescence: A growth period conducive to alienation. *Adolescence, 22*, 929-938.
- Cohen, J.S., & Smerdon, B.A. (2009). Tightening the dropout tourniquet: Easing the transition from middle to high school. *Preventing School Failure: Alternative Education for Children and Youth, 53*(3), 177-184.
- Connolly, J. (1989). Social self-efficacy in adolescence: Relations with self-concept, social adjustment and mental health. *Canadian Journal of Behavioural Science, 21*(3), 258-269.
- Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development, 65*(1), 1-118.
- Deary, I.J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*, 13-21.
- Deshler, D. D., & Schumaker, J. B. (1993). Strategy mastery by at-risk students: Not a simple matter. *The Elementary School Journal, 94*, 153-167.
- Eccles, J.S., & Midgley, C. (1989). Stage/environment fit: Developmentally appropriate classrooms for early adolescents. In R. E. Ames & C. Ames (Eds.), *Research on motivation in education (Vol. 3)*. New York: Academic Press.
- Gettinger, M., & Seibert, J.K. (2002). Contributions of study skills to academic competence. *School Psychology Review, 31*(3), 350-365.
- GL Education (2015). *Cognitive Abilities Test: Fourth edition (CAT4)*. Retrieved from <http://www.gl-education.com/international-products/cognitive-abilities-test-fourth-edition-cat4>

- Goodenow, C. (1993). Classroom belonging among early adolescent students: Relationships to motivation and achievement. *Journal of Early Adolescence, 13*, 21-43.
- Graham, S., McKeown, D., Kiuahara, S., & Harris, K.R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*(4), 879-896.
- Gullatt, Y., & Jan, W. (2002). *How do pre-collegiate academic outreach programmes impact college-going among underserved students?* Boston: Pathways to College Network and Clearinghouse.
- Hammond, C., & Reimer, M. (2006). *Essential elements of quality after-school programs*. Clemson, SC: Communities in Schools and National Dropout Prevention Center/Network.
- Hoge, D.R., Smit, E.K., & Crist, J.T. (1995). Reciprocal effects of self-concept and academic achievement in sixth and seventh grade. *Journal of Youth and Adolescence, 24*, 295-314.
- Johnson, A.W. (1999). *Sponsor-a-Scholar: Long-term impacts of a youth mentoring program on academic performance*. Plainsboro, NJ: Mathematica Policy Research, Inc.
- Keltikangas-Jarvinen, L. (1992). Self-esteem as a predictor of future school achievement. *European Journal of Psychology of Education, 7*, 123-130.
- Kulik, J.A., & Kulik, C.C. (1992). Meta-analytic findings on grouping programs. *Gifted Child Quarterly, 36*(2), 73-77.
- Lane, J., & Lane, A. (2001). Self-efficacy and academic performance. *Social behaviour and Personality, 29*(7), 687-694.
- Lauer, P.A., Akiba, M., Wilkerson, S.B., Apthorp, H.S., Snow, D., & Martin-Glenn, M. (2004). *The effectiveness of out- of-school-time strategies in assisting low-achieving students in reading and mathematics: A research synthesis* (Updated ed.). Aurora, CO: Mid-continent Research for Education and Learning.
- Lauer, P.A., Akiba, M., Wilkerson, S.B., Apthorp, H.S., Snow, D., & Martin-Glen, M.L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research, 76*(2), 275-313.
- Marsh, H.W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*, 646-656.
- McFarlane, A.H., Bellissimo, A., & Norman, G.R. (1995). The role of family and peers in social self-efficacy: Links to depression in adolescence. *American Journal of Orthopsychiatry, 63*(3), 402-410.
- Miller, B. M. (2003). *Critical hours: Afterschool programs and educational success*. Brookline, MA: Miller Midzik Research Associates for the Nellie Mae Education Foundation. Retrieved November 27, 2015, from <http://www.nmefdn.org/CriticalHours.htm>.

- Miller, K. (2007). The benefits of out-of-school time programs. *Principal's Research Review*, 2(2), 1-6.
- Mizelle, N.B. (1995, April). *Transition from middle school into high school: The student perspective*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Mizelle, N.B., & Irwin, J.L. (2000). Transition from middle school to high school. *Middle School Journal*, 31(5), 57-61.
- Muijs, R.D. (1997). Symposium: Self-perception and performance predictors of academic achievement and academic self-concept: A longitudinal perspective. *British Journal of Educational Psychology*, 67, 263-277.
- Muris, P. (2001). A brief questionnaire for measuring self-efficacy in youths. *Journal of Psychopathology and Behavioural Assessment*, 23(3), 145-149.
- Myers, D., & Schirm, A. (1999). *The impacts of Upward Bound: Final report for phase I of the national evaluation*. Plainsboro, NJ: Mathematica Policy Research, Inc.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4), 543-578.
- Roderick, M. (2006, April). *Closing the aspirations-attainment gap: Implications for high school reform*. Paper presented at the MDRC: High School Reform Committee conference, San Diego, CA.
- Ross, S.M., Smith, L.T., Casey, J., & Slavin, R.E. (1995). Increasing the academic success of disadvantaged children: An examination of alternative early intervention programs. *American Educational Research Journal*, 32(4), 773-800.
- Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A systematic approach (7<sup>th</sup> ed.)*. Sage.
- Strand, S. (2004). Consistency in reasoning test scores over time. *British Journal of Educational Psychology*, 74, 617-631.
- Thorndike, R. L., Hagen, E., & France, N. (1986). *Cognitive Abilities Test Second Edition: Administration Manual*. Windsor: nferNelson.
- Wigfield, A., Eccles, J.S., Mac Iver, D., Reuman, D.A., & Midgley, C. (1991). Transitions during early adolescence: Changes in domain specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology*, 27, 552-565.
- Wiltfang, G. L., & Scarbecz, M. (1990). Social class and adolescents' self-esteem: Another look. *Social Psychology Quarterly*, 53, 174-183.



## APPENDIX A

### Semi-structured focus group protocol

For Focus Groups held with preparatory programme teachers

#### SESSION A: INTRODUCTIONS and ORAL CONSENT (10 MINUTES)

- Purpose of the focus group

#### **ORAL CONSENT PREAMBLE: Confidentiality and how the information will be used –**

As you may be aware, the organisation is trying to improve its preparatory programme.

I am a student currently doing my Master's in programme evaluation at UCT, and I am working on the programme as a component of my degree.

I would like to ask you a few questions about your experience of teaching students in the preparatory programme.

Please note that what you say will be used to inform the evaluation report which will be circulated amongst programme staff, and which will ultimately form the basis of a 50% dissertation which will be publically accessible via UCT's dissertation database. Non-sensitive quotes and excerpts which recall what you say in this Focus Group may be paraphrased in the report. I will, however, always keep confidential the source from which data is collected, and the evaluation report will not name individual informants. No-one at the programme will know from whom a specific quote or paraphrased comment came from in this report.

Please note that your participation is voluntary and you can withdraw at any time. The focus group will probably take up to 55 minutes.

If you have any further questions or concerns about what the evaluation is about, and how data will be used, please feel free to email me about it at [heleneamelia.duplessis@gmail.com](mailto:heleneamelia.duplessis@gmail.com)

Are you happy to continue with the focus group on these terms?

- Get teachers to give their names (do not record names)
- Questions from teachers
- Opportunity to get a beverage / food

SESSION B: OVERALL IMPRESSIONS (15 minutes)

- Overall, did the 2014 preparatory programme run smoothly?
- Is there anything about how the preparatory programme ran last year that you think is worth mentioning?

SESSION C: STRUCTURE OF THE PROGRAMME (10 minutes)

- Did the order of the planned sessions change at all?
- How well do you think the structure worked?
- Do you think anything could be done in order to improve it?
- Do you think there were enough sessions?
- Did you find the division of students into three streams useful?

SESSION C: DURATION OF THE SESSIONS (5 minutes)

- How long did it normally take to cover the content for a session?
- Was there enough/too much time allocated to each session?

SESSION D: RESOURCES AND MANAGEMENT (10 minutes)

- Were you provided with the necessary resources for each session?
- Are there any additional resources you would have found helpful?
- Do you think the preparatory programme was well organised?

SESSION E: CHALLENGES (10 minutes)

- Did you experience any challenges in presenting the maths/English content as planned?
- Were there any sessions which were particularly challenging for you as teachers? Why?

SESSION F: RECOMMENDATIONS (5 minutes)

- Do you have any suggestions for how the preparatory programme could be improved?

## APPENDIX B

### Semi-structured interview protocol

For interviews held with preparatory programme recipients from 2014 cohort

#### SESSION A: INTRODUCTIONS and ORAL CONSENT (5 MINUTES)

- Purpose of the interview

#### **ORAL CONSENT PREAMBLE: Confidentiality and how the information will be used –**

As you may be aware, the organisation is trying to improve its preparatory programme. I am a student currently doing my Master's in programme evaluation at UCT, and I am working on the programme as a component of my degree.

I would like to ask you a few questions about your experiences of the preparatory programme last year.

Please note that what you say will be used to inform the evaluation report which will be circulated amongst programme staff, and which will ultimately form the basis of a 50% dissertation which will be publically accessible via UCT's dissertation database. Non-sensitive quotes and excerpts which recall what you say in this interview may be paraphrased in the report. I will, however, always keep confidential the source from which data is collected, and the evaluation report will not name individual informants.

Please note that your participation is voluntary and you can withdraw at any time. Parental consent for your participation in this interview has already been granted.

The interview will probably take up to 40 minutes.

If you have any further questions or concerns about what the evaluation is about, and how data will be used, please feel free to contact me about it on 083 453 1975.

Are you happy to continue with the interview on these terms?

- Get student to give their name (do not record name)
- Questions from the student

SESSION B: ADAPTING TO HIGH SCHOOL (15 minutes)

- When you started high school, what was your first impression?
- How did you feel about making friends?
- Did you feel like you fit in? Why / why not?
- Has this gotten better or worse as the year has progressed?
- What has helped you to adjust to high school?
- How have you found the mathematics classes at your new school?
- How have you found the English classes at your new school?
- Which are your favourite / least favourite subjects at school? Why?

SESSION C: ROLE OF THE PREPARATORY PROGRAMME (15 minutes)

- Is being able to cope with schoolwork important to adapting to high school?
- How did it help you to adapt?
- Did a lot of what you learned in the maths component of the programme get used in the first six months at your new school?
- Do you think that you were at a similar level of ability to your classmates in maths?
- Did a lot of what you learned in the English component of the programme get used in the first six months?
- Do you think that you were at a similar level of ability to your classmates in English?
- How well do you think you would have coped without it?
- What aspects did you find the most helpful?
- What aspects did you find least helpful?

SESSION D: RECOMMENDATIONS (5 minutes)

- How do you think the preparatory programme could be improved?

**APPENDIX C**



**Consent form for interviews**

**University of Cape Town**

**Faculty of Commerce**

Dear parent/guardian

Congratulations on the selection of your child into the scholarship programme. The organisation is currently aiming to improve the preparatory programme, and you can help to accomplish this by giving permission for your child to attend an interview. The interview will take approximately 40 minutes. It will be conducted at a time which is convenient for you and your child.

The aim of this interview is to determine which aspects of the preparatory programme your child found beneficial and which aspects could be improved upon.

This research has been approved by the UCT Commerce Faculty Ethics in Research Committee.

Your participation in this research is voluntary. You can choose to withdraw from the research at any time.

Due to the nature of the study your child will need to provide the researchers with their name- but all responses will be confidential and used for the purposes of this research only. Programme staff will not know which children took part in the interviews and no identifying information will be provided in the report.

Should you have any questions regarding the research, please feel free to contact the researcher.

Heléne du Plessis: 083 453 1975

I have read and understood the information above, and give permission for my child to complete the interview.

---

Name and surname

Date

Parent/ guardian signature

---

Child's name and surname

**APPENDIX D**



**Consent form for SEQ-C Questionnaire**

**University of Cape Town**

**Faculty of Commerce**

Dear parent/guardian

Congratulations on the selection of your child into the preparatory programme. The organisation is currently aiming to improve the preparatory programme, and you can help to accomplish this by giving permission for your child to complete a questionnaire. The questionnaire will take approximately 15 minutes to complete. It will be administered once before the programme begins and again after the programme has finished in November.

The aim of this research is to investigate the effect of the preparatory programme on self-efficacy.

This research has been approved by the UCT Commerce Faculty Ethics in Research Committee.

Your participation in this research is voluntary. You can choose to withdraw from the research at any time.

Due to the nature of the study your child will need to provide the researchers with their name- but all responses will be confidential and used for the purposes of this research only.

Should you have any questions regarding the research, please feel free to contact the researcher.

Heléne du Plessis: [heleneamelia.duplessis@gmail.com](mailto:heleneamelia.duplessis@gmail.com)

I have read and understood the information above, and give permission for my child to complete the questionnaires.

---

Name and surname

Date

Parent/ guardian signature

---

Child's name and surname

Name of primary school

## APPENDIX E

### Self-Efficacy Questionnaire for Children

Adapted from Muris (2001)

Name: \_\_\_\_\_

Please read each statement and tick the appropriate block for each question.

	Never	Rarely	Sometimes	Often	Always
1. I ask a teacher to help me when I get stuck on schoolwork.					
2. I can argue back well when classmates disagree with me.					
3. I can cheer myself up when something bad has happened.					
4. I can study even when there are other interesting things to do.					
5. I succeed in becoming calm again when I am very scared.					
6. I can become friends with other people my age.					
7. I can study a chapter for a test.					
8. I can have a chat with someone I don't know.					
9. I can stop myself from feeling nervous.					
10. I succeed in finishing all my homework every day.					
11. I can work well with my classmates.					

	<b>Never</b>	<b>Rarely</b>	<b>Sometimes</b>	<b>Often</b>	<b>Always</b>
12. I can control my feelings.					
13. I can pay attention during all my subjects.					
14. I can tell other children that they are doing something that I don't like.					
15. I can give myself a pep talk when I feel low.					
16. I succeed in passing all my subjects.					
17. I can tell a funny story to a group of children.					
18. I can tell a friend that I don't feel well.					
19. My schoolwork makes my parents happy.					
20. I succeed in staying friends with people.					
21. I succeed in keeping unpleasant thoughts away.					
22. I succeed in passing school tests.					
23. I am able to stop arguments between people.					
24. I succeed at not worrying about things that might happen.					



**APPENDIX F**  
**POWER CALCULATIONS**

All power calculations were conducted using the online Power and Sample Size calculator.

**Evaluation Question 2: Organisational support: Is the CAT instrument that the programme is using a useful tool for informing selection?**

**CAT standardised age score means**

$\alpha = 0.05$

Group A mean: 105.3

Group B mean: 103.5

Standard deviation: 8.44

Sampling ratio (A/B): 0.6

Sample size: 125

Power  $(1-\beta) = 0.31$

**CAT quantitative standardised age scores**

$\alpha = 0.05$

Group A mean: 108.9

Group B mean: 107.7

Standard deviation: 9.5

Sampling ratio (A/B): 0.6

Sample size: 125

Power  $(1-\beta) = 0.13$

**CAT verbal standardised age scores**

$\alpha = 0.05$

Group A mean: 102.8

Group B mean: 100.8

Standard deviation: 11.19

Sampling ratio (A/B): 0.6

Sample size: 125

Power  $(1-\beta) = 0.23$

### **CAT non-verbal standardised age scores**

$\alpha = 0.05$

Group A mean: 104.2

Group B mean: 101.9

Standard deviation: 9.4

Sampling ratio (A/B): 0.6

Sample size: 125

Power  $(1-\beta) = 0.38$

### **Evaluation Question 4: Can the observed changes in learners' mathematics and English performance be attributed to the intervention?**

#### **Mathematics**

$\alpha = 0.05$

Group mean 1: 78.8

Group mean 2: 82.1

Standard deviation: 8.3

Sampling ratio (A/B): 1

Sample size: 47

Power  $(1-\beta) = 0.48$

#### **English**

$\alpha = 0.05$

Group mean 1: 80.2

Group mean 2: 81.1

Standard deviation: 6.1

Sampling ratio (A/B): 1

Sample size: 47

Power  $(1-\beta) = 0.11$

**Evaluation Question 5: Did the students in the 2014 cohort in each of the streams benefit equally from the intervention? Did some streams benefit more/less than others?**

**Mathematics:**

**Top stream**

$\alpha = 0.05$

Group mean 1: 83.3

Group mean 2: 85

Standard deviation: 9.3

Sampling ratio (A/B): 1

Sample size: 18

Power  $(1-\beta) = 0.08$

**Mid-stream**

$\alpha = 0.05$

Group mean 1: 78.6

Group mean 2: 82.4

Standard deviation: 6.4

Sampling ratio (A/B): 1

Sample size: 17

Power  $(1-\beta) = 0.41$

**Bottom stream**

$\alpha = 0.05$

Group mean 1: 72.2

Group mean 2: 77.5

Standard deviation: 9.5

Sampling ratio (A/B): 1

Sample size: 12

Power (1- $\beta$ ) = 0.27

**English:**

**Top stream**

$\alpha$  = 0.05

Group mean 1: 83

Group mean 2: 82.7

Standard deviation: 3

Sampling ratio (A/B): 1

Sample size: 17

Power (1- $\beta$ ) = 0.06

**Mid-stream**

$\alpha$  = 0.05

Group mean 1: 82.1

Group mean 2: 84

Standard deviation: 7.2

Sampling ratio (A/B): 1

Sample size: 16

Power (1- $\beta$ ) = 0.11

**Bottom stream**

$\alpha$  = 0.05

Group mean 1: 74.8

Group mean 2: 75.9

Standard deviation: 7.5

Sampling ratio (A/B): 1

Sample size: 14

Power (1- $\beta$ ) = 0.06

**Evaluation Question 6: Can the observed changes in the learners' cognitive reasoning ability be attributed to the intervention?**

**Standardised age score means**

$\alpha = 0.05$

Group A mean: 105.3

Group B mean: 111.2

Standard deviation: 5.6

Sampling ratio (A/B): 1

Sample size: 47

Power  $(1-\beta) = 0.99$

**Quantitative standardised age scores**

$\alpha = 0.05$

Group A mean: 108.9

Group B mean: 111.3

Standard deviation: 16.4

Sampling ratio (A/B): 1

Sample size: 47

Power  $(1-\beta) = 0.1$

**Verbal standardised age scores**

$\alpha = 0.05$

Group A mean: 102.8

Group B mean: 109.2

Standard deviation: 7.1

Sampling ratio (A/B): 1

Sample size: 47

Power  $(1-\beta) = 0.99$

### **Non-verbal standardised age scores**

$\alpha = 0.05$

Group A mean: 104.2

Group B mean: 111.1

Standard deviation: 8.7

Sampling ratio (A/B): 1

Sample size: 47

Power  $(1-\beta) = 0.97$

### **Evaluation Question 7: Did the students in the 2015 cohort show improvements in their self-efficacy scores?**

#### **Total self-efficacy**

$\alpha = 0.05$

Group mean 1: 95.9

Group mean 2: 97.1

Standard deviation: 7.5

Sampling ratio (A/B): 1

Sample size: 49

Power  $(1-\beta) = 0.12$

#### **Academic self-efficacy**

$\alpha = 0.05$

Group mean 1: 34.5

Group mean 2: 33.7

Standard deviation: 3.3

Sampling ratio (A/B): 1

Sample size: 49

Power  $(1-\beta) = 0.22$

**Social self-efficacy**

$\alpha = 0.05$

Group mean 1: 32.1

Group mean 2: 33

Standard deviation: 2.8

Sampling ratio (A/B): 1

Sample size: 49

Power  $(1-\beta) = 0.35$

**Emotional self-efficacy**

$\alpha = 0.05$

Group mean 1: 29.2

Group mean 2: 30.2

Standard deviation: 4.5

Sampling ratio (A/B): 1

Sample size: 49

Power  $(1-\beta) = 0.19$