BRITISH COUNCIL **A N** ? **E V A L U A T I O N** 2 **O F** Z W B :: 1 7 B = 6 P W 8 F W 2 4 > ΙΜΡΑΟΤ ENGLISHK ENGLISH8LANGUAGE CAPABI FEBRUARY AND INSWOR



FOREWORD

CONTRIBUTORS

- EXECUTIVE SUMMARY 1.
- 2. INTRODUCTION
- RESEARCH METHODOLOGY 3. The sample design: a stratified two-stage cluster sa Method of delivery The English language assessment The student context questionnaire Motivational scale analysis
- 4. IMPLEMENTATION IN BOGOTÁ, COLOMBIA Educational context in Colombia Educational context in Bogotá The comparison unit
- 5. LANGUAGE LEARNING ENVIRONMENT
- SAMPLING RESULTS 6.
- 7. ASSESSMENT OUTCOMES

Receptive skills Productive skills Comparing achievement by gender Comparing achievement by urban and rural schoo Comparing achievement by socio-economic status

ENGLISH LANGUAGE LEARNING MOTIVATION 8.

Construct validation findings Multi-group analysis findings Links between motivation and proficiency Summary of English language learning motivation

9. IN CONCLUSION

REFERENCES

APPENDIX A - OVERVIEW OF THE APTIS FOR TEENS T APPENDIX B -FINAL CFA MEASUREMENT MODEL EST

H Y 4 H 0 1

	04
	07
	08
	10
	12
sample design	12
	16
	16
	19
	24
	28
	28
	29
	30
	32
	38
	42
	43
	44
	45
ol location	46
S	47
	48
	49
	55
	63
findings	71
	72
	74
TEST STRUCTURE	78
IMATES	84

LIST OF TABLES

Table 2: Reference table of motivational scales and associated descriptions2Table 3: Stratification variables chosen3Table 4: Demographic variables of participating students from Bogotá3Table 5: Target population definition in Bogotá3Table 6: Coverage and exclusions3Table 7: School sample size3Table 8: Student sample size4Table 9: Population and sample size4Table 10: Weighted school and student participation rates4Table 11: Distribution of student population for overall test performance by CEFR level4Table 12: Mean population score by skills and overall achievement4Table 14: Estimated distribution of student population for Istening and reading skills4by CEFR level4Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by gender, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups5Table 24: T-tests to check for differences between gender groups5Table 25: Corelations between dender groups5 <th>Table 1: Level of localisation for Aptis for Teens (from O'Sullivan and Dunlea, 2015)</th> <th>18</th>	Table 1: Level of localisation for Aptis for Teens (from O'Sullivan and Dunlea, 2015)	18
Table 3: Stratification variables chosen3Table 4: Demographic variables of participating students from Bogotá3Table 5: Target population definition in Bogotá3Table 5: Target population definition in Bogotá3Table 6: Coverage and exclusions3Table 7: School sample size3Table 8: Student sample size4Table 9: Population and sample size4Table 10: Weighted school and student participation rates4Table 11: Distribution of student population for overall test performance by CEFR level4Table 12: Mean population score by skills and overall achievement4Table 13: Estimated distribution of student population for listening and reading skills4by CEFR level7Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances5between error terms7Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups6(standardised estimates)6Table 25: Correlations between groups according to their pro	Table 2: Reference table of motivational scales and associated descriptions	21
Table 4: Demographic variables of participating students from Bogotá3Table 5: Target population definition in Bogotá3Table 6: Coverage and exclusions3Table 7: School sample size3Table 8: Student sample size4Table 9: Population and sample size4Table 10: Weighted school and student participation rates4Table 11: Distribution of student population for overall test performance by CEFR level4Table 12: Mean population score by skills and overall achievement4Table 13: Estimated distribution of student population for listening and reading skills by CEFR level4Table 14: Estimated distribution of student population for speaking and writing skills dy CEFR level4Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status5Table 18: Improvements in adjusted chi-square values following addition of covariances5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 22: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between	Table 3: Stratification variables chosen	30
Table 5: Target population definition in Bogotá3Table 6: Coverage and exclusions3Table 7: School sample size3Table 8: Student sample size4Table 9: Population and sample size4Table 10: Weighted school and student participation rates4Table 11: Distribution of student population for overall test performance by CEFR level4Table 12: Mean population score by skills and overall achievement4Table 13: Estimated distribution of student population for listening and reading skills4by CEFR level7Table 14: Estimated distribution of student population for speaking and writing skills4by CEFR level7Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status5between error terms5Table 19: Error covariances included in the measurement model5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 25: Latent covariances allowed to vary between gender groups6(standardised estimates)6Table 25: Latent covariances between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender g	Table 4: Demographic variables of participating students from Bogotá	32
Table 6: Coverage and exclusions3Table 7: School sample size3Table 8: Student sample size4Table 8: Student sample size4Table 10: Weighted school and student participation rates4Table 11: Distribution of student population for overall test performance by CEFR level4Table 12: Mean population score by skills and overall achievement4Table 13: Estimated distribution of student population for listening and reading skills4by CEFR level7Table 14: Estimated distribution of student population for speaking and writing skills4by CEFR level7Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status5Table 18: Improvements in adjusted chi-square values following addition of covariances5between error terms7Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between groups according to their proficiency Table 26: Significant differenc	Table 5: Target population definition in Bogotá	38
Table 7: School sample size3Table 8: Student sample size4Table 9: Population and sample size4Table 10: Weighted school and student participation rates4Table 11: Distribution of student population for overall test performance by CEFR level4Table 12: Mean population score by skills and overall achievement4Table 13: Estimated distribution of student population for listening and reading skills4by CEFR level7Table 14: Estimated distribution of student population for speaking and writing skills4by CEFR level7Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances5between error terms7Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 24: T-tests to check for dif	Table 6: Coverage and exclusions	38
Table 8: Student sample size4Table 9: Population and sample size4Table 10: Weighted school and student participation rates4Table 11: Distribution of student population for overall test performance by CEFR level4Table 12: Mean population score by skills and overall achievement4Table 13: Estimated distribution of student population for listening and reading skills4by CEFR level7Table 14: Estimated distribution of student population for speaking and writing skills4by CEFR level7Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances5between error terms7Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 23: Questions used in the construction of SES variable5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between groups according to their proficiency6Table 26: Significant differences between groups according to their proficiency6Table 26: Significant differences between groups accord	Table 7: School sample size	39
Table 9: Population and sample size4Table 10: Weighted school and student participation rates4Table 11: Distribution of student population for overall test performance by CEFR level4Table 12: Mean population score by skills and overall achievement4Table 13: Estimated distribution of student population for listening and reading skills4by CEFR level4Table 14: Estimated distribution of student population for speaking and writing skills4by CEFR level5Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 18: Improvements in adjusted chi-square values following addition of covariances5between error terms5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 26: Significant differences between groups according to their proficiency6Table 26: Significant differences between groups according to their proficiency6Table 26: Significant differences between groups according to their proficiency6Table 26: Significant differences between groups according to their prof	Table 8: Student sample size	40
Table 10: Weighted school and student participation rates4Table 11: Distribution of student population for overall test performance by CEFR level4Table 12: Mean population score by skills and overall achievement4Table 13: Estimated distribution of student population for listening and reading skills4by CEFR level5Table 14: Estimated distribution of student population for speaking and writing skills4by CEFR level6Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances5between error terms5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 24: T-tests to check for differences between gender groups5Table 25: Latent covariances allowed to vary between gender groups6(standardised estimates)6Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 9: Population and sample size	40
Table 11: Distribution of student population for overall test performance by CEFR level4Table 12: Mean population score by skills and overall achievement4Table 13: Estimated distribution of student population for listening and reading skills4by CEFR level7Table 14: Estimated distribution of student population for speaking and writing skills4by CEFR level7Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances5between error terms5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 22: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between gender groups (standardised estimates)6Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between groups according to their proficiency Table 26: Significant differences between groups according to their proficiency Table 26: Significant differences between groups according to	Table 10: Weighted school and student participation rates	41
Table 12: Mean population score by skills and overall achievement4Table 13: Estimated distribution of student population for listening and reading skills by CEFR level4Table 14: Estimated distribution of student population for speaking and writing skills by CEFR level4Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances between error terms5Table 19: Error covariances included in the measurement model5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between groups according to their proficiency Table 26: Significant differences between groups according to their proficiency6Table 26: Significant between motivational variables and overall total score	Table 11: Distribution of student population for overall test performance by CEFR level	42
Table 13: Estimated distribution of student population for listening and reading skills by CEFR level4Table 14: Estimated distribution of student population for speaking and writing skills by CEFR level4Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances between error terms5Table 19: Error covariances included in the measurement model5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 26: Significant differences between groups according to their proficiency (standardised estimates)6Table 26: Significant differences between groups according to their proficiency (standardised estimates)6Table 26: Significant differences between groups according to their proficiency (standardised estimates)6Table 26: Significant differences between groups according to their proficiency 	Table 12: Mean population score by skills and overall achievement	42
Table 14: Estimated distribution of student population for speaking and writing skills by CEFR level4Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances between error terms5Table 19: Error covariances included in the measurement model5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between groups according to their proficiency Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 13: Estimated distribution of student population for listening and reading skills by CEFR level	43
Table 15: Mean performance score by gender, skills and overall achievement4Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances5between error terms5Table 19: Error covariances included in the measurement model5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 22: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 23: Questions used in the construction of SES variable5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between gender groups (standardised estimates)6Table 27: Correlations between motivational variables and overall total score6	Table 14: Estimated distribution of student population for speaking and writing skills by CEFR level	44
Table 16: Mean performance score by school location, skills and overall achievement4Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances5between error terms5Table 19: Error covariances included in the measurement model5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 22: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between groups according to their proficiency6Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 15: Mean performance score by gender, skills and overall achievement	45
Table 17: Comparing overall mean score performance by socio-economic status4Table 18: Improvements in adjusted chi-square values following addition of covariances5between error terms5Table 19: Error covariances included in the measurement model5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 22: Latent covariances allowed to vary between gender groups5(standardised estimates)5Table 23: Questions used in the construction of SES variable5Table 25: Latent covariances allowed to vary between gender groups5(standardised estimates)5Table 25: Latent covariances allowed to vary between gender groups6(standardised estimates)6Table 26: Significant differences between groups according to their proficiency6Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 16: Mean performance score by school location, skills and overall achievement	46
Table 18: Improvements in adjusted chi-square values following addition of covariances between error terms5Table 19: Error covariances included in the measurement model5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 22: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between gender groups (standardised estimates)5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between groups according to their proficiency Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 17: Comparing overall mean score performance by socio-economic status	47
Table 19: Error covariances included in the measurement model5Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 22: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between groups according to their proficiency6Table 26: Significant differences between motivational variables and overall total score6	Table 18: Improvements in adjusted chi-square values following addition of covariances between error terms	50
Table 20: Standardised (STDYX standardisation) correlations between variables in the final model5Table 21: T-tests to check for differences between mean factor scores calculated from the gender model5Table 22: Latent covariances allowed to vary between gender groups (standardised estimates)5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 24: T-tests to check for differences between mean factor scores calculated from the gender model5Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)6Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 19: Error covariances included in the measurement model	53
Table 21: T-tests to check for differences between mean factor scores calculated5from the gender modelTable 22: Latent covariances allowed to vary between gender groups5(standardised estimates)Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated5Table 25: Latent covariances allowed to vary between gender groups6(standardised estimates)6Table 25: Latent covariances allowed to vary between gender groups6Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 20: Standardised (STDYX standardisation) correlations between variables in the final model	54
Table 22: Latent covariances allowed to vary between gender groups5(standardised estimates)5Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated5from the gender model5Table 25: Latent covariances allowed to vary between gender groups6(standardised estimates)6Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 21: T-tests to check for differences between mean factor scores calculated from the gender model	56
Table 23: Questions used in the construction of SES variable5Table 24: T-tests to check for differences between mean factor scores calculated5from the gender model7Table 25: Latent covariances allowed to vary between gender groups6(standardised estimates)7Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 22: Latent covariances allowed to vary between gender groups (standardised estimates)	57
Table 24: T-tests to check for differences between mean factor scores calculated5from the gender modelTable 25: Latent covariances allowed to vary between gender groups6(standardised estimates)Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 23: Questions used in the construction of SES variable	58
Table 25: Latent covariances allowed to vary between gender groups6(standardised estimates)7Table 26: Significant differences between groups according to their proficiency6Table 27: Correlations between motivational variables and overall total score6	Table 24: T-tests to check for differences between mean factor scores calculated from the gender model	59
Table 26: Significant differences between groups according to their proficiency 6 Table 27: Correlations between motivational variables and overall total score 6	Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)	62
Table 27: Correlations between motivational variables and overall total score	Table 26: Significant differences between groups according to their proficiency	65
	Table 27: Correlations between motivational variables and overall total score	66

LIST OF FIGURES

// G W O 2 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á , P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H O 3 3 Q J H

Table 28: Correlations between motivational variables and scale scores for individual components of the proficiency test	68
Table 29: Correlations between overall proficiency and motivational variables, by gender	69
Table 30: Correlations between overall proficiency and motivational variables, by SES	70
OF FIGURES	
Figure 1: The socio-cognitive model for test design and validation	17
Figure 2: The grade participating students from Bogotá started learning English	33
Figure 3: Time spent learning English per week at school	34
Figure 4: Activity undertaken when studying English outside of school	34
Figure 5: Time spent learning English outside of school	35
Figure 6: Future pathway when finished compulsory schooling	35
Figure 7: Language most regularly used to watch, read or listen to different types of media	36
Figure 8: Activity most often done on computer, tablet or smart phone	37
Figure 9: Time spent using digital devices on average per week	37
Figure 10: Responses to a sample of questionnaire items	48
Figure 11: Hypothesised measurement model	51
Figure 12: Measurement model following modifications (final model)	52
Figure 13: Differences in mean factor scores on motivational variables between 'higher' and 'lower' SES groups	61
Figure 14: Mean factor scores on motivational variables for SES percentile groupings	62
Figure 15: Motivational profiles for learners at different levels of proficiency	64
Figure 16: Correlation between overall proficiency and English self-concept	67
Figure 17: Correlation between overall proficiency and the ought-to L2 self	67

FOREWORD

Tom Miscioscia

Country Director, British Council, Colombia

There has been an increasing interest around English language learning and teaching in Colombia during the last three decades. Two major factors have played a key role within this context: the Colombian constitution of 1991 which explicitly describes the nation as multilingual and multicultural and the country's increasing global visibility and need to be more competitive.

A number of national and local initiatives have demonstrated the importance the English language represents for the government, education institutions, researchers and parents – and in enhancing the next generation of Colombians' employment opportunities. The current government has stated a clear goal in its Development Plan (2014-2018) regarding English learning: at least 8% of the students in the public sector are expected to reach a B1 (CEFR) level at the end of their high school studies.

The British Council Colombia has supported the Ministry of Education, Local Education Authorities and the National Bilingual programme, "Colombia Bilingüe", since 2004 through the implementation a series of strategic actions: assisting the adoption of the CEFR as a scale for measuring language competency and progress, developing the foreign languages standards, conducting regular assessment projects with samples of high school students and English teachers and leading continuous professional development locally and abroad. It is important to recognise that there have been challenges in Colombia's drive to improve English standards and that there is a need to continuously measure Colombia's performance in this area through robust evaluation systems. International studies and reports have recently highlighted ongoing areas for development in Colombian students' English language proficiency in terms of global comparatives, although there is tangible evidence that the country continues to advance.

The English Impact project allows us to understand better the achievements of the Bogotá bilingual programme in the context of the "Colombia Bilingüe" initiative, both in terms of the progress made by students and the impact of the diverse teacher-training strategies developed in the city. This project is particularly important for us given our commitment to enable more widespread and better quality training, teaching, learning and assessment of English worldwide. At the British Council, we hope that the results of the English Impact report will provide the national and local government with essential data and evidence to inform its future development.

Barry O'Sullivan

Head of Assessment Research and Development, British Council

English Impact, perhaps the British Council's most ambitious language-related research undertaking to date, has the capacity to offer ministries of education an accurate and objective diagnostic of the language capability of their country, region or city. This methodology is particularly powerful as it is based on expert statistical design and analysis combined with contextually appropriate interpretation of data that is a precise representation of the general population for the age group examined.

We have learned so much in the process of designing and delivering this project that it is clearly impossible to highlight all. The sampling work undertaken with Martin Murphy and his team from the Australian Council for Educational Research (ACER) is the single exception. Their professionalism helped to take our vision to a new and elevated level. Their clear and thoughtful consideration of how comparison units should be defined and how to ensure that the final test population was truly representative brought significant challenges, but also significant improvement to our understanding of the processes involved in complex sampling. The results described in this report are testament to the impact that Martin's thinking has had on our work.

The reality of delivering English Impact was a challenge that would stretch the project team to its limits. In fact, without the key players who undertook the research design and delivery, it is highly unlikely that I would be writing this foreword.

A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 P 9 B E B 1 7 P 9 B E A 1 7 P 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 P 9 B E B 1 7 P 9 B E B 1 7 P 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 P 9 B E B W 8 F

We were incredibly lucky to work with exceptional local British Council and Secretary of Education teams in Bogotá who did a great deal to ensure the success of the data collection. The determination of these two teams to deliver the project to the highest possible level of quality was critical to its eventual success.

I must confess to feeling great pride in the completion of the English Impact project in Bogotá. This report demonstrates the highest level of professionalism and will come to be recognised as a major achievement both within the British Council and in the world of English language education and policy. I expect that it will help the Secretary to continue to conceive and pursue successful policies for many years to come.

CONTRIBUTORS

Martin Murphy

Senior Research Fellow, Australian Council for Educational Research

At the heart of all good educational policy and practice are teachers, schools, and school systems working to improve the learning outcomes of students. Every day educators and policymakers globally are striving to achieve this goal in very different circumstances.

Learning from experience is an established method of improving performance. I believe all good teachers learn from their colleagues as all good schools share their experiences with other schools within their system. The same can be said for developing policies and practice at the system level. This is where English Impact aims to contribute high quality international comparative outcomes data on English language learning for this purpose.

Education systems are complex. They are shaped by many factors such as geographic location and social and economic background. By mapping the British Council's Aptis assessment outcomes onto a common population framework and by quantifying national and regional variations against that international framework, English Impact aims to identify educational policies and practices associated with the successful teaching and learning of English.

Australian Council for Educational Research (ACER) has thoroughly enjoyed its collaboration with British Council in this endeavour, helping to bring to English Impact the same methodologies underlying major international surveys, such as the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA), in the development of this population framework, sampling, weighting and variance estimation.

ACER congratulates the Secretary of Education of Bogotá and the British Council team for their very successful implementation of English Impact. The quality of the survey implementation – evidenced by the very high rates of participation and coverage, and levels of precision that meet or exceed the standards of TIMSS or PISA – should give every confidence to readers of this report, and those keen to learn from Bogotá's experiences in the increasingly important field of English language teaching and learning. We would like to acknowledge the invaluable contributions made by British Council staff in Colombia and consultation provided by the Australian Council for Educational Research and the University of Bath. Without the hard work and expert involvement of the following people, English Impact Bogotá would not have been possible.

British Council, Colombia for its implantation of English Impact:

- Ricardo Romero, Director of English
- Carolina Cruz Corzo, Academic Consultant
- Viviana Caicedo Triana, Projects Coordinator
- Takeshi Shibasaki, Aptis, Professional & Overseas Exams Administrator.

The Australian Council for Educational Research for its specialist contribution to sampling methodology and analysis:

- Martin Murphy, Senior Research Fellow
- Clare Ozolins, Research Fellow
- Kate O'Malley, Research Fellow
- Jorge Fallas, Research Fellow
- Tim Friedman, Research Fellow
- Bethany Davies, Research Officer.

The University of Bath for its contribution to motivational scale questionnaire design and analysis:

• Dr Janina Iwaniec, Lecturer in TESOL.

:: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E // 1 7 B = 2 V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E // 1 7 B = 2 0 V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 W B 8 F W ? 9 B E B 1 7 ? B = 0 P W 8 F W ? 9 B E B 1 7 ? B = 0 P W 8 F W ? 9 B E B 1 7 ? B = 0 P W 8 F W ? 9 B E B 1 7 ? B = 0 P W 8 F W ? 9 B E B 1 7 ? B = 0 P W 8 F W ? 9 B E B 1 7 ? B = 0 P W 8 F W ? 9 B E B 1 7 ? B = 0 P W 8 F W ? 9 B E B 1 7 O A G W B :: 1 7 B = 0 P W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B W 8 F W ? 9 B

Many thanks also to the following people for their valued support and input.

Dr Karen Dunn, Senior Researcher, British Council for her contribution to statistical analysis.

Viviana Caicedo Triana, Research Assistant, British Council for managing the collation and publication of the reports.

Members of the British Council English Impact Board

Cherry Gough	Barry O'Sullivan
Rebecca Hughes	Anna Searle
John Knagg	Greg Selby
Maureen McDonagh	Duncan Wilson

Members of the British Council Assessment Research Group

Vivien Berry	Judith Fairbairn
Victoria Clarke	Barry O'Sullivan
Jamie Dunlea	Mina Patel
Karen Dunn	Richard Spiby

Key members of the British Council Tests Team

Dan Elsworth	Jermaine Prince
Catherine Hughes	

1. EXECUTIVE SUMMARY

English Impact aims to provide robust policy-relevant data evidencing English language capability in Bogotá

English Impact aims to assess capability by measuring the current ability of a targeted sample of the school population from the Bogotá government sector using an English language assessment; and evaluating future potential through an in-depth analysis of students' language learning opportunities in and outside the classroom, language learning motivations and socio-economic background.

To ensure the English Impact Bogotá data collected can be used to inform and support education system and policy development, the British Council worked in collaboration with the Secretary of Education tailoring the research to meet local needs.

The rationale for undertaking this research is supported by the British Council's Royal Charter and charitable objective to develop a wider knowledge of the English language, and looks to build upon the organisation's rich heritage of global English language research. A pioneer of the study of English language, the British Council has significant experience contributing analysis and insight, while advancing knowledge across the field. While previous research has explored and expanded existing understanding of how growth in the use of the English language could shape the world economy, English Impact creates new baseline data to measure levels of English language capability.

The research was carried out by the British Council, with contributions from the Australian Council for Educational Research and the University of Bath. In 2016–17, the research was piloted within four regional and national level education systems: Sri Lanka; Bangladesh; the Community of Madrid in Spain; and the Metropolitan District of Bogotá in Colombia.

English Impact in Bogotá

- English Impact employs a two-stage cluster sample design used by other recognised largescale international surveys, sampling schools at the first stage and students at the second stage.
- 150 government-funded schools and 1,800 students were sampled to participate in English Impact 2016–17, with 149 schools and 1,479 students in Bogotá ultimately taking part. (Two of these schools did not meet the required student participation rate of 50 per cent, so they are not included in the school-level sampling estimates.)
- Students were sampled from Grade 10 in compulsory secondary education. This grade represents 10 years of schooling, counting from the first year of International Standard Classification of Education (ISCED) Level 1, and a mean age at the time of testing was at least 15.9 years.
- Students sampled were studying English as part of their studies at this grade level. A minimum of 90 minutes of formal English study per week as part of the school program was required for eligibility to the target population.
- Students completed the British Council's Aptis for Teens English Language assessment, testing reading, writing, speaking, listening, and grammar and vocabulary.

 A questionnaire comprising 53 items, delivered in Spanish, captured opinions and data from students on their school and language learning backgrounds, their language learning motivations and socio-economic status.

Key findings

- Overall school and student participation in Bogotá met the English Impact international participation standard of at least 84 per cent of sampled students in 98 per cent of sampled schools.
- Of the participating students, 30 per cent started learning English in pre-school or first grade.
- In their current grade, 18 per cent of participating students chose to study English at school, while it was a compulsory school subject for 76 per cent. Almost two thirds of participating students (63 per cent) spent between three and four hours a week studying English.
- Over half (55.4 per cent) of the participating students said they used English regularly to play computer games. A quarter (25 per cent) watched films in English, while 15.8 per cent and 16.7 per cent respectively interacted with social networks and the Internet in English.
- Over two-thirds (73 per cent) of the Bogotá students achieved at A1 CEFR level in their overall English language test performance, while 21 per cent achieved at A2 level, 4 per cent at B1 level, and 1 per cent at B2 level.
- The skill of listening achieved the highest mean scale score of 21.63 out of 50. Over two-thirds (72 per cent) of students achieved at A2 level, while 23 per cent achieved at B1 level.
- Reading achieved the second highest mean scale score of 14.35. More than half (57 per cent) of students achieved at A2 level on the CEFR, while 30 per cent were at A1 level.
- Participating students achieved their lowest performance scores for the productive skills.
 Speaking skills had the lowest mean scale score with 3.39 out of 50, and 79 per cent of students achieved at A0 level. The mean scale

B E B 1 7 B = Z V 2 8 T Z Q U <</td> 2 5 3 5 % A 8 F W ? 9 B E B 1 7 B ? 9 B E B 1 7 ? 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 0 9 3 Q J H 2 V 2 8 T I Q U<</td> 2 5 3 5 % A X G % L 1 U M 8 ? ? 9 B E B W 8 F W ? ? 8 B W ? ? 8 B W ? ? 8 S W ? ? 8<

score for writing skills was 7.38 out of 50, and 50 per cent achieved at A0 level.

- Male and female students were almost equivalent on the overall scores, with males marginally higher.
- Female students reported a higher level of motivation to learn English, than male students, on four of the seven motivational scales. The findings imply that having confidence in one's ability to study English is more closely related to the positive language learning experience for female learners than males.
- The confidence of boys in language learning was found to be more closely related to their proficiency than that of girls. This finding is in contrast to those from previous studies; further investigation is necessary to understand the processes explaining this finding.
- Schools in Bogotá's urban areas outperformed those in the rural areas (on the outskirts of the city) in all domains. Urban schools had an overall mean score of 53.90 out of 200, while the overall score for rural schools was 40.22 out of 200.
- Schools with medium socio-economic status had a higher performance in all domains, with an overall mean score of 58.91 out of 200, compared to 50.44 for schools with low socio-economic status.
- There are clear effects of socio-economic status on language learning motivation, with students from more advantaged backgrounds reporting higher motivation than their peers from less advantaged families. While significant differences were found on four of the seven aspects of motivation, there were no significant differences between the two SES groups in terms of their language learning attitudes, their perceived societal pressure to study English, and their reported effort to study English.
- The findings on motivation suggest that students from the higher SES group are more likely to translate their motivation into proficiency gains than students from the lower SES group.

2. INTRODUCTION

English language learning now plays a significant role in many national and regional education systems, increased proficiency having been identified by policymakers as contributing to economic prosperity. Examples of government policies that prioritise the improvement of English proficiency can be seen across the world. Considerably harder to find are good data that provide a comparable baseline of evidence showing levels of English language capability at the heart of where government policy makes an impact – in publicly-funded school classrooms.

Highly influential sources of data assessing academic achievement across public education systems do exist in the shape of PISA (the Programme for International Student Assessment), TIMSS (Trends in International Mathematics and Science Study), and PIRLS (Progress in International Reading Literacy Study). Collectively known as international large-scale surveys and administered by the Organisation for Economic Co-operation and Development (OECD) and the International Association for the Evaluation of Educational Achievement (IEA) respectively, their results are at the same time eagerly awaited and severely criticised for their deeply influential impact on educational practices in many countries. To date, none have included the assessment of language, but much can be learnt from the processes designed to sample and implement large-scale research of this kind from decades of experience.

This global best practice in research, and experience of data collection, is emulated within the design of the English Impact methodology that will be detailed in the following chapters. When designing this research, we have also tried to learn from the potentially damaging effect that international large-scale surveys can have. By identifying the best and, by default, the worst performing education systems, international large-scale surveys can, at times, have a negative impact. In anticipation of this perhaps inevitable 'horse race', an adaptation of the concept of capability underpins our research design.

The theoretical basis used to define English language capability is derived from an adaptation of Amartya Sen's capabilities approach. Eminent economist, philosopher and driver of social change, Sen's revolutionary contribution to development economics involved defining the concept of capability. First conceived in the 1980s as an approach to welfare economics, the theory become predominant as a paradigm for human development and inspired the creation of the UN's Human Development Index. Sen describes the capabilities approach to human development as 'a concentration on freedom to achieve in general and the capabilities to function in particular'. The core concepts within his theory surround functionings that are explained in relation to achievements, and capabilities, as people within societies possessing the opportunity to achieve (Saito, 2003).

Central to Sen's concept of capability was combining functionings and capabilities, achievement and opportunity. This adaptation of English language capability can, therefore, be described in terms of the level of achievement, or proficiency, reached by a defined population; and the opportunities provided to them to achieve greater proficiency via teaching and learning practice derived from a policy or national guideline.

Achievement, proficiency, progress or aptitude of individual English language learners are most commonly measured by a language test. Bachman (1990) suggests that as research instruments, language tests can support investigations into the nature of language proficiency and language teaching practice and perform a role in programme evaluation, only when combined with other forms of data. Critical language testing theorists also believe the knowledge created via a test is 'narrow and simplistic [...] it is mono-logic based on one instrument which is used on one occasion, detached from a meaningful context'. They suggest that using a test can provide 'a quick fix' (Shohamy, 1998), and an instant solution. However, analysis of data captured via this method alone overlooks the complexities of broader subject matter and is meaningless for the reform of education policy.

This evaluation of English language capability, presented here as an adaptation of Sen's capabilities approach, does therefore not only involve the measurement of English language proficiency captured by a test. Other data were captured and combined to provide full context to our analysis: language policy, language learning environment, language proficiency, and language learning motivations. The presentation of these supporting data is intended to provide depth and insight into students' assessment outcomes and go some way to showing the impact of the English language policy in Bogotá.

B E B 1 7 B = Z V 2 8 T Z Q U <</td> 2 5 3 5 % A 3 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 1 1 3 Q J H 2 V 2 8 T I Q U 2 5 3 5 % A X G % L 1 U M 8 ? ? 9 B E B W 8 F W ? ? 8 S % A X G % L 1 U M</

The research aims outlined and investigated were to:

- evaluate the English language capability of students studying at public schools in Bogotá
- compare the capability of students attending urban and rural schools on the outskirts of Bogotá, and the socio-economic status
- understand the relationship between English language learning motivation and increased proficiency.

To achieve these research aims, the British Council brought together world-leading research specialists, in collaboration with our own expertise in English language assessment, to create the English Impact research methodology.

3. RESEARCH METHODOLOGY

The English Impact research methodology was designed in direct response to the research aims outlined in Chapter 2, to provide the most credible evaluation of the English language capability in Bogotá. The concept of capability is characterised by the unique combination of understanding both current achievement and future opportunity, by its nature involving analysis of multiple data to capture students' current ability and future potential to succeed. Also fundamental to the evaluation of English language capability within a national or regional education system is an appropriate sampling methodology employed to accurately reflect the population of interest and supply sufficiently precise estimates from the survey. Reflecting the theoretical framework of capability, the research methodology involves three central components:

- the sample design
- the English language assessment
- the student context questionnaire.

THE SAMPLE DESIGN: A STRATIFIED TWO-STAGE CLUSTER SAMPLE DESIGN

The sampling methodology was designed by the Australian Council for Educational Research (ACER) based on its extensive experience in large-scale international educational surveys. The procedures used were drawn extensively from the practices and experiences of major comparative educational surveys that have been operating internationally for well over a decade, in particular, surveys of the International Association for the Evaluation of Educational Achievement (IEA), principally the Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS), as well as surveys of the OECD, specifically the Programme for International Student Assessment (PISA). These surveys are highly regarded internationally for their quality, and have become major contributors to educational research and policy development around the world.

The British Council team in Bogotá participated in a detailed sampling process designed by ACER and modified locally to ensure all procedures were feasible. An overview of the two-stage cluster sampling activities is given below:

- Preparation
- define the comparison unit
- identify exclusions
- determine stratification variables
- obtain database of schools and agree access
- agree the sample design
- School sampling
- select the school sample
- School liaison and student sampling
- obtain student data from schools
- select student sample
- inform schools of selected students
- arrange dates for English Impact test participation
- Data tracking
- track school participation
- track student participation.

The British Council team in Bogotá worked directly with the Secretary of Education and the Melbourne-based ACER research team to gather the school and student level data needed to complete the sampling process described above.

Participation standards

The English Impact research was guided by an established set of participation standards drawn from those used within established international surveys described above. Use of these standards enabled precision around the major estimates of the research, namely English Language capability, and to maximise the comparability of outcomes across participating countries.

The following participation standards were applied throughout the sampling implementation and analysis procedures.

Standard 1.1 Students in all schools within the comparison unit – including all educational sub-systems – who meet the criteria documented below, are part of the international target population. Students who meet the international target population are referred to as the 'eligible students'.

Standard 1.2 The target population should provide the most exhaustive coverage of students. Any deviation from full coverage of the comparison unit needs to be described and quantified in advance.

Standard 1.3 The total of combined schoollevel exclusions and within-school exclusions within the comparison unit will be no greater than five per cent of the comparison unit target population.

Standard 1.4 Only students within the comparison unit target population participate in the test.

Standard 1.5 The school sample for English Impact Bogotá will be drawn using established and professionally recognised principles of scientific sampling.

Standard 1.6 A minimum of 150 schools will be drawn for English Impact Bogotá from the comparison unit.

B E B 1 7 B = Z V 2 8 T Z Q U <</td> 2 5 3 5 % A 8 F W ? 9 B E B 1 7 B ? 9 B E B 1 7 ? 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 1 3 3 Q J H 2 V 2 8 T I Q U 2 5 3 5 % A X G % L 1 U M 8 # W 2 S 3 5 % A X G % L 1 U M 8 F W ? 9 B E B W 8 F W ?</

- **Standard 1.7** The English Impact Bogotá school response rate is at least 85 per cent of sampled schools. If a response rate is below 85 per cent, then a pre-determined, systematic use of replacement schools will be implemented.
- **Standard 1.8** The English Impact Bogotá student response rate is at least 85 per cent of all sampled students across responding schools. This response rate includes students from replacement schools.
- **Standard 1.9** Absent sampled students cannot be replaced by non-sampled students.
- The international target population is as follows:
- Students within the comparison unit enrolled in the grade that represents 10 years of schooling counting from the first year of International Standard Classification of Education (ISCED) Level 1, providing the mean age at the time of testing is at least 15.5 years, and who are currently studying English as part of their studies at this grade level, for a minimum of 90 minutes of formal study per week as part of the school programme.

The target grade and mean age

The international target population is defined to ensure comparability across education systems. It is important that students participating in the survey are at equivalent stages of schooling, as well as of comparable age. The naming of grades and the age of entry into formal schooling varies between countries. Therefore, the target grade was aligned across countries to allow for accurate reporting.

The International Standard Classification of Education (ISCED)

UNESCO's International Standard Classification of Education (ISCED) is an internationally recognised classification of the levels of schooling across countries, ranging from pre-primary education (ISECD 0) all the way through to tertiary education (ISCED 6). As with IEA studies such as TIMSS, use of this classification will align the levels of education within individual countries to a common international framework. ISCED 1 is commonly referred to as 'primary schooling'.

Comparison unit

The term 'comparison unit', used throughout the description of the English Impact research sample design, is an integral part of the research concept and measurement of English language capability to inform more effective policy development, as described in the report's Introduction.

Many aspects of educational policy development, such as English language learning, often occur at sub-national levels, e.g. provinces and states. Within provinces or states, there may be further divisions – for example, between public and private sectors. There is increasing recognition that, at the national level, the focus of international largescale surveys can be limited with respect to exploring aspects of educational provision that can vary within participating countries.

Ten years from the start of ISCED Level 1

Drawn directly from TIMSS, this part of the population definition is in recognition that the starting age of students into ISCED 1 varies, with students in some countries beginning primary school at a younger age than in other countries.

Ninety minutes of formal English per week

This definition means that the survey provides an estimate of English language capability for all Bogotá's students meeting this definition and studying at least 90 minutes of formal English learning per week, rather than for the entire student population of Bogotá.

Where educational provision is primarily the responsibility of provinces or states, using a province or state as the comparison unit can allow for local policies and practices to be clearly related to results, rather than diluted by a national result where variation in conditions between states or provinces can mask these local effects.

Precedents established in PISA's inclusion of 'adjudicated regions' and TIMSS' use of 'benchmarking entities' alongside national level units of comparison have informed the comparison unit policy implemented throughout English Impact. Close adherence to the participation standards and population definition described above, in combination with concisely described and internationally recognisable units of comparison, informed the decision to allow both national and subnational comparison units to participate within the research project.

Precision of estimates

The primary basis for the determination of sample sizes is the desired precision of major outcomes from the survey. Common practice is the presentation of this measurement in the form of standard errors and/or confidence intervals around survey estimates, as will be followed in presentation of English Impact research outcomes. The following minimum sample size for each comparison unit was recommended for every participating comparison unit:

- a minimum of 150 schools
- a target of 12 students from each sampled school
- a target of 1,800 students overall.

Drawing further on established standards used in large-scale international surveys such as TIMMS and PIRLS, thresholds for desired standard errors measurements were established. TIMSS and PIRLS report scores on a scale with a mean of 500 and a standard deviation of 100. To achieve this level of precision, these surveys aim to achieve a sample size such that the standard error is no larger than .035 standard deviation units. This equates to a standard error no larger than 3.5 score points. This standard error means a 95 per cent confidence interval of ±7.0 score points around the estimated mean.

For percentage estimates, like the percentage of students in each CEFR level for English Impact, the maximum standard error desired was set at 1.75 per cent of the percentage estimate. This means that the confidence interval around population percentage estimates should be less than ±3.5 per cent.

Coverage and exclusions

All students enrolled in the target grade, studying at least 90 minutes of English per week and within the comparison unit belong to the target population. The target population is intended to provide full coverage of all eligible students

A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 P 9 B E B 1 7 P 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 P 9 B E B 1 7 P 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 P 9 B E B 1 7 P 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 P 8 9 8 1 % > ? 9 B E B 1 7 P 9 B E B 1 7 P 8 9 8 1 % > ? 9 B E B 1 7 P 9 B E B 1 7 P 8 9 8 1 % P ? 9 B E B 1 7 P 8 9 8 1 % > ? 9 B E B 1 7 P 8 9 8 1 7 P 8 1 P 8 1 P 8 1 P 8 1 P 8 1 P 8 1 P 8 1 P 8 1 P 8 1 P 8

within the comparison unit. Any deviation from full coverage of the comparison unit was described and quantified in advance of the data collection phase. Every effort was made to ensure complete coverage of the whole population, however, in all established sampling exercises of this kind there are often practical reasons invoked for excluding schools and students:

- school-level exclusions may include schools that are very remote or very small
- student-level exclusions include students with either functional or intellectual disabilities that prevent them from taking part in the assessment, fitting the predefined criteria.

To ensure comparability and maximum coverage of the eligible population, the standards for English Impact require that school and withinschool exclusions should not exceed 5 per cent.

Stratification

A process of implicit stratification was implemented throughout the English Impact sampling methodology. Implicit stratification has the effect of sorting the school sampling frame by a set of implicit stratification variables. It is an effective way of ensuring a proportional allocation of schools across all implicit strata in the sample. Common stratification variables include urban or rural school status, geographic region or school funding type. Stratification can lead to improved reliability of survey estimates, provided the stratification variables are related to those survey outcomes.

METHOD OF DELIVERY

To carry out the assessments in every sampled school in Bogotá, a pioneering digital method of delivery was developed. Every English language assessment and student questionnaire were completed by students via a completely offlineenabled tablet. Other large-scale assessments, such as PISA and TIMSS, have made initial steps towards computer-based assessment. English Impact has pioneered completion of a largescale assessment using 100 per cent computerbased delivery.

Data were collected via two applications (apps) on each tablet in fully invigilated conditions. A keyboard was used for the writing component to make this process as easy as possible. Individual headphones with a microphone were used for the speaking and listening components. This delivery method aimed to ensure all students were tested as consistently as possible despite location, Internet access or available in-school facilities. Fully computer-based delivery allows like-for-like comparison of results and research outcomes that are robust, reliable and consistent.

The two research tools used to capture data via the tablet apps, the English language assessment and the student context questionnaire, are described below.

THE ENGLISH LANGUAGE ASSESSMENT

The English language assessment instrument used to measure the English proficiency of participants in the English Impact research was British Council's Aptis for Teens test assessing four skills: reading, writing, speaking and listening, as well as grammar and vocabulary.

The Aptis test system

Aptis is a computer-based test of general English proficiency and currently has four main variants:

- Aptis General
- Aptis Advanced
- Aptis for Teachers
- Aptis for Teens.

No specific cultural or first language background is required, and test content is developed to be appropriate for English language learners in a variety of contexts. Aptis General, Aptis for Teachers, and Aptis Advanced are designed for adults and young adults aged 16 years or over. Aptis for Teens is for 13 to 17-year olds. An important feature of the tests developed within the Aptis test system is their alignment with the Common European Framework of Reference for Languages (CEFR), a widely used international framework of language proficiency providing detailed descriptions of what language learners are able to do with a language at six different levels of proficiency (Council of Europe, 2001). Incorporating the CEFR in the development of the Aptis test system helps to interpret results by linking the test to an internationally recognised set of proficiency benchmarks.

All Aptis test variants are designed to provide information on the ability of test takers to participate in a wide range of general language use situations. The Aptis test system is an approach to test design, development and delivery devised by the British Council to provide flexible English language assessment options to test users. There are five components: core (knowledge of grammar and vocabulary), reading, listening, writing and speaking. Although the core component is always administered, organisations are able to select any combination of the other components according to their needs. For English Impact in Bogotá, all five components were taken.

Theoretical model underpinning the test system

The theoretical model of test development and validation that underpins the Aptis test system is based on the socio-cognitive model proposed by O'Sullivan (2011, 2015), O'Sullivan and Weir (2011), and Weir (2005). As O'Sullivan (2015) notes: 'the real strength of this model of validation is that it comprehensively defines

Figure 1: The socio-cognitive model for test design and validation

THE TEST SYSTEM

Test task

Performance parameters Linguistic demands

Test Administration

THE SCORING SYSTEM

Theoretical fit Accuracy of decisions Value of decisions

A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E B 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E // 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 W B 8 F W ? 9 B E B 1 7 O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 O A G W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B W 8 F W

each of its elements with sufficient detail as to make the model operational'. The socio-cognitive model is based around three elements:

- the test taker
- the test system
- the scoring system.

The model drives design decisions by specifying how these three elements combine to result in a measure of candidate performance that is meaningful in terms of the English language ability being assessed. This in turn allows the test developers to collect evidence in a systematic way in the creation of a validation argument to support claims about the test. Figure 1, taken from O'Sullivan (2015), demonstrates how the three elements feed into the test takers' performance.



Research evidence supporting the validity of the test system

An important part of the Aptis test system has been the commitment of the British Council to support an active and robust validation research and dissemination agenda. A dedicated team carries out research and statistical analyses at the design and development stage. Operational test delivery data are regularly analysed to ensure the tests perform to demanding technica performance criteria. The Assessment Research Awards and Grants scheme actively funds research into the tests from leading international researchers. An impressive body of published documentation, covering an extensive and diverse range of validation projects, contribute important evidence to the validity argument supporting the uses of the Aptis test system.

Localisation: Adapting tests for particular uses

The term 'localisation' is used within the Aptis test system to refer to the ways in which the Aptis test is adapted for use in particular contexts with particular populations to allow for particular decisions to be made. The model identifies different levels of localisation depending on the degree of change from the original underlying framework used in the development of Aptis, and the amount of resources required to realise that change. Aptis for Teens is considered to be a Level 4 localisation based on the five-level model described in O'Sullivan and Dunlea (2015). The description for Level 4 is reproduced below in Table 1. Table 1: Level of localisation for Aptis for Teens(from O'Sullivan and Dunlea, 2015)

h	Level 4	Partial re- definition of target construct	Developing new task types that are more
		from existing	relevant for a
		involve	population of
al		developing different task	test takers, while remaining
٦		types to elicit	within the
al		different aspects of	overall framework of
		performance.	the Aptis test
			system (e.g. Aptis for Teens).

The Aptis for Teens test

The Aptis for Teens test variant used in the English Impact project has been designed specifically to meet the needs of younger language learners by testing their English language skills through familiar scenarios. Task parameters such as topic, genre and the intended audience are relevant to the target use domain of a teenager. Questions reflect activities that occur in everyday life, such as social media, homework, school events and sport. For example, instead of writing a complaint letter to a company – a task used in Aptis General for adults but something a teenager may not yet have yet experienced they might be asked to write about the benefits and drawbacks of a social issue relevant to teenagers and likely to be discussed in classrooms. The cognitive competencies of the age group are also taken into consideration. Tasks are tailored to provide support needed for this age group to give them the chance to perform to the best of their ability. For an overview of the structure of each component of the Aptis for Teens test, please see Appendix A.

Aptis for Teens was designed for a specific age group, young learners aged from 13 to 17 years. As with the other main variants in the Aptis test system, the test is designed to be used with test takers irrespective of culture, country of origin or residence, gender or first language. This means that background knowledge is not tested, bias is reduced and language skills are isolated for testing.

An important part of the features impacting on the test system also relates to the test delivery environment. The English Impact project tests were invigilated by a British Council employee, who visited each school to conduct the testing. This additional level of quality assurance ensured the security and uniformity of the test delivery.

Scoring and reporting

The scoring system is the final area of validation. The core, reading and listening components are scored automatically within the computer delivery system. Trained raters mark the speaking and writing components, using an online rating system.

Aptis for Teens test results are reported on a numerical scale (0–50) and as a CEFR level for each component. An overall CEFR level is also given if all components are completed by the test taker. The CEFR describes English language proficiency across six levels (A1–C2). In Aptis for Teens, results are reported for levels A1 to B2, and if a test taker demonstrates an ability beyond B2, this is reported as C (C1 and C2 are not differentiated in Aptis for Teens).

The core, reading and listening components use selected-response formats such as multiple choice, gap fill and matching tasks. Speaking and writing components require test takers to provide samples of spoken and written performance. The speaking test is a semi-direct test in which test takers record responses to pre-recorded prompts. The writing test approximates online written communication. The focus of the

speaking and writing marking scales is on test taker communicative competence; these are marked by trained raters. See Appendix A for a detailed overview of the task types contained in each component.

THE STUDENT CONTEXT QUESTIONNAIRE

Context questionnaires form an integral part of most large-scale surveys. While accurate information on student performance is central to such assessments, the factors that are linked to performance are of particular interest, not only for researchers, but also for practitioners and policymakers in education.

Alongside the four-skill English language assessment, students complete a background questionnaire to gather contextual information to support the English language capability data. The questionnaire was translated into Spanish.

The student questionnaire comprises 53 items in three sections.

Demographic background, including socioeconomic status variables

Items within this section of the questionnaire include: grade, gender, age; prior schooling, language spoken within the home, and country of birth. A number of questionnaire items were used to measure the latent variable of socioeconomic status (SES). This will be explained and explored fully in Chapter 8.

English language learning inside and outside school

Items within this section of the questionnaire include: the grade at which students began learning English; time spent in subject area lessons learning English; and whether students study English outside of school.

English language learning motivation

This section of the student context questionnaire was designed by Dr Janina Iwaniec from the University of Bath. Dr Iwaniec is a secondlanguage learning motivational specialist who conducted a review of the most relevant and influential theories and constructs used to measure language learning motivation.

Motivation is one of the most influential of all individual differences, trumping factors such as language learning aptitude (Gardner and Lambert, 1972) in explaining gains in proficiency in certain contexts. Recently, it has been shown that motivation is more important than the age of onset, with students who start later developing higher levels of motivation and guickly catching up with the proficiency of learners who start English instruction early (Pfenninger and Singleton, 2016). Motivation is also considered to be responsive to appropriate interventions (Taylor and Marsden, 2014) and can be enhanced or decreased as a result of language learning environment (Ushioda, 2009). Both this relatively strong influence on language learning and its malleability make motivation a factor that is crucial for language learning policies.

In the years since research into language learning motivation started in the 1950s (Gardner and Lambert, 1959), there have been a large number of theories of language learning motivation proposed. The choice of constructs for English Impact was guided by the most up-to-date theories and research on language learning motivation. This included the L2 Motivational Self System (Dörnyei and Ushioda, 2009) which consists of three constructs – ideal L2 self, ought-to L2 self and language learning experience; international orientation (Yashima, 2009); and self-concept (Bong and Skaalvik, 2003).

The scales were measured by 32 questions addressing eight different areas of language learning motivation. These hypothesised motivational scales have strong theoretical and empirical grounding described in Table 2. Including four questions for each scale increases the value of the information gathered for each area, something that is exploited in the analytic approach. Grouped in one section of the questionnaire and in a randomised order, students were asked to give a response to each statement using a six-point Likert scale.

 Table 2: Reference table of motivational scales and associated descriptions

Motivational scale (latent variable)	Descriptive name	Motivational scale descriptions	Question reference	Question details
Ideal L2 self	Personal language	Ideal L2 self is an image of oneself as a proficient	11	I imagine myself speaking English fluently.
(IDEAL)	goals	speaker of a second language (Dörnyei, 2005). Though it relates to the future-self, <i>ideal L2 self</i> needs to be considered attainable to retain its motivational properties. English Impact	12	I imagine myself comfortably reading in English on the Internet.
			13	I imagine myself easily being able to follow what others say to me in English.
		employed the Iwaniec (2014) scale as it encompasses the four skills of language learning.	14	I imagine myself writing emails in English with ease.
Ought-to L2 self (OUGHT)Social expectationsThe ought-to L2 self is b on the external expectation placed on students and relates to the 'attributes to one believes one ought to possible negative outcom (Dörnyei, 2005, pages 105–106).	The ought-to L2 self is based on the external expectations placed on students and relates to the 'attributes that	01	I consider learning English important because the people I respect think that I should do it.	
		one believes one ought to possess in order to avoid possible negative outcomes' (Dörnyei, 2005, pages 105–106).	02	Studying English is important to me because other people will respect me more if I have knowledge of English.
			03	Studying English is important to me because an educated person is supposed to be able to speak English.
			04	Learning English is necessary because people surrounding me expect me to do so.
Language learning	Interest in learning	Language learning experience is concerned with the influence of the immediate environment on language learning (Dörnyei, 2005) and implies a strong focus on language learning attitudes.	EX1	Learning English is really great.
experience (EXPER)	English		EX2	I look forward to my English classes.
			EX3	I find learning English really interesting.
			EX4	l really enjoy learning English.

? // G W 2 0 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á , P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 2 1 3 Q J H

	Motivational scale (latent variable)	Descriptive name	Motivational scale descriptions	Question reference	Question details
	Instrumentality (INSTR) Future opportunities	Instrumentality represents motivation stemming from the	INSTR1	I need English for my future career.	
			(Gardner and Lambert, 1972). This scale measures the perceptions of usefulness of	INSTR2	The things I want to do in the future require me to use English.
		English on jol future prospe	English on job markets and future prospects.	INSTR3	I study English because it will facilitate my job hunt in the future.
				INSTR4	I study English as it will help me to earn good money.
	International orientation	Global communication	International orientation is a construct recently developed in response to the	INTOR1	Studying English will help me understand different people from other countries.
			changing role of English. It denotes an 'interest in foreign or international affairs [] readiness to interact with intercultural partners' (Yashima, 2000, page 57). The scale used, found in Iwaniec (2014), lends itself to adaptations that take into account the growth of online interaction rather than travelling abroad.	INTOR2	In the future, I would really like to communicate with people from other countries.
				INTOR3	In the future, I would really like to communicate with people from other countries online.
				INTOR4	If I could speak English well, I could get to know more people from other countries via the Internet.
En co (SE	English self- concept	Self- confidence in	Self-concept is 'a person's perception of himself'	SELF1	l usually get good marks in English.
	(SELF)	English	(Shavelson, Hubner and Stanton, 1976) and this scale relates to self-evaluation in	SELF2	Compared to other students, I'm good at English.
		the student's ability to study English. The most common	SELF3	l have always done well in English.	
		is Marsh's (Marsh, 1990) Academic Self-Description Questionnaire, adapted to language learning by Iwaniec (2014).	SELF4	Studying English comes easily to me.	

	Motivational scale (latent variable)	Descriptive name	Motivational scale descriptions	Question reference	Question details
	Parental encouragement (PAREN)	Family expectations	Like the ought-to L2 self, parental encouragement focuses on external expectation. As the participants in English Impact are as young as 15.5 years, there is a potential for their motivation to be influenced by their parents or guardians. Parents are considered to be one of the three groups of important others, together with teachers and peers (Williams and Burden, 1997).	PAR1	My parents think I need to know English to be well educated.
	(,,,,_,,,			PAR2	My parents have stressed the importance English will have for me in the future.
				PAR3	My parents feel that it is very important for me to learn English.
				PAR4	My parents encourage me to practice my English as much as possible.
	Motivated learning	Level of effort	Motivated learning behaviour attempts to	MB1	l work hard at learning English.
b (1	(MOTIV)	component of motivation, i.e. the reported amount of effort	MB2	I think I'm doing my best to learn English.	
			a student invests in English language learning.	MB3	l put a lot of effort into learning English.
				MB4	l spend lots of time studying English.

B E B 1 7 B = Z V 2 8 T Z Q U <</td> 2 5 3 5 % A 8 F W ? 9 B E B 1 7 B > ? 9 B E B 1 7 ? 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 2 3 3 Q J H 2 V 2 8 T I Q U <</td> 2 5 % A X G % L 1 U M 8 ? ? 9 B E B W 8 F W ? ? ? 8 F W ? ? ? 8 F W ? ? 8 F W ? ?</

MOTIVATIONAL SCALE ANALYSIS

The questionnaire responses were analysed using confirmatory factor analysis (CFA) in Mplus 7 (Muthén and Muthén, 1998–2017). Factor analytic techniques are of great value in motivational research as they explicitly address the requirement to gain insights into constructs that are not directly observable. These abstract concepts are termed latent variables. Examples include depression in psychology, consumer expectation in economics, and anomie in sociology. Socio-economic status is also often treated as a latent variable.

In factor analysis, a number of observed variables, or measures, are hypothesised to be indicators of the existence of an underlying latent variable. No individual observed variable is considered to give a precise measure of a hypothesised latent variable, as there will always be a certain amount of error in the measurement of any observed variable. Much as, for example, a psychiatrist would expect to see high levels of a range of indicators before diagnosing a patient, factor analysis will combine information from a number of observed variables in order to give information about the levels of a hypothesised latent variable.

In the motivational questionnaire employed in the current study, the aim was to capture information about students' motivational levels in eight areas using 32 questions (see Table 2 above for details). Each block of four questions was targeted to gain insight into a distinct underlying motivational construct. The questions were presented to the students in a randomised order, so that respondents would not immediately perceive the shared focus of a set of questions. Employing CFA techniques means that the information from all four observed variables can be combined to give a more robust insight into the motivational traits of interest than would be achieved from modelling individual observed responses.

There are two distinct stages of analysis reported here, both of which employ factor analytic techniques:

- construct validation this involves checking whether the guestionnaire functioned as expected with regard to gaining insights into different areas of motivation
- *multi-group analysis* for the purposes of the current report, three sets of comparative student grouping were examined: male/ female; urban/rural; and socio-economic status (based on parental employment status and household possessions as reported by the students).

The scope of the questionnaire analysis reported under these headings is limited to investigating the motivational scales in their own terms. The findings from this initial analysis are then taken forward to investigate the relationships between motivational scales and proficiency as measured by the Aptis test, as described in the next section.

Construct validation

The primary aim in construct validation was to establish whether it makes sense to understand observed student responses with reference to the eight hypothesised motivational scales. In CFA, the measurement model refers to this relationship between the responses given by the participants to the questions (the observed data) and the motivational scales (latent variables) that reflect each of the motivational constructs. The CFA technique can be employed to assess how much of the variation in the original observed dataset can be explained with reference to this pre-defined latent structure. This involves accounting for the shared variation, or correlations, between the observed measures.

Estimates from the measurement model (known as factor loadings) give an indication of how much variation in the observed variable is accounted for by the latent construct. Some variables will have a stronger relationship than others. If there is a close relationship between all observed variables and the associated latent variable, there is a strong internal consistency in the scale. However, it is worth noting that a latent variable will never account for all of the variation in any given observed variable – there will always be some measurement error. This reflects the principle inherent to factor analysis, in that any given observed variable is driven by an underlying trait (in this case of motivation) and will not provide a precise measure of it. Measurement error takes into account, for example, idiosyncratic responses to questions worded in a certain way.

CFA is a data reduction technique that draws upon a reduced number of variables to replicate patterns in the observed data. In order to assess whether the hypothesised measurement model achieves this successfully, a number of fit statistics are employed. These indices represent several different means of indicating how well the latent structure can be used to replicate the variation in the observed dataset. Essentially, if the structure hypothesised by the model is able to capture the patterns of question responses well, then the model is considered to be a good fit. For the fit indices reported here, it is usually expected that the comparative fit index (CFI) and Tucker-Lewis index (TLI) statistics be above 0.9 (or ideally 0.95), and the root mean square error of approximation (RMSEA) below 0.05 for good model fit¹. Where alterations are made to the initially hypothesised model, comparisons are made using the adjusted chi-square difference test (Satorra and Bentler, 2010). This is the recommended means of comparing CFA models

1 See Byrne (2012, pages 69–77) for a description of what these indices represent and the values accepted to show good fit.

2 MLR refers to 'maximum likelihood parameter estimates with standard errors and a chi-square test statistic (when applicable) that are robust to non-normality and non-independence of observation' (Muthén and Muthén, 2011, page 533).

? // G W 2 4 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á ,P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 2 5 3 Q J H

estimated using the maximum likelihood with robust standard errors (MLR) approach² employed in the current analysis (Muthén and Muthén, 2011). Model comparisons are undertaken in a systematic manner to ensure that any additional parameters included in the model bring about an overall improvement to the model fit.

The first step in the analysis reported here is thus to establish a measurement model that reflects the data well. This is carried out for all cases in the dataset together, before moving on to the group comparisons.

Multi-group analysis

There are a number of approaches that can be taken to making multi-group comparisons within a CFA modelling approach (see, e.g., Byrne, 2012, pages 193–281). For current purposes, the focus is on two areas:

- A. comparing relative *levels* of motivation expressed for each scale
- A. comparing the *relationships between* the different areas of motivation.

Under (A), the model is used to derive what are known as factor scores for each of the students on each of the motivational scales. In other words, for each of the motivational areas listed in Table 2, each participating student will be assigned a value (factor score) depending on their responses to the relevant question. This is more complex than simply averaging the responses, as it takes into account the weighted relationships estimated within the model. Factor scores are expressed on a standardised scale, which does not bear any easily perceptible relationship to the original measurement scale. However, it is the comparison between levels that are of interest here, therefore. the scale is not essential.

For (B), the relationships examined are those between the latent variables, this relationship is referred as the structural model. In CFA terminology, this is restricted to covariances, which do not presume any directionality in the relationship between variables. The value of investigating these relationships as part of this multi-group analysis is that they tell us whether the balance between the motivational scales is consistent between groups. The model is set up so that the measurement model is kept consistent across groups, but relationships between latent variables are allowed to vary where significant differences are found.

This is achieved by creating two models: one in which all estimated parameters are assumed to be the same between groups (the constrained model) and one in which all parameters are freely estimated between groups (the configural model). If there is a significant difference between these two models, as assessed by comparing model chi-squares, this indicates some group differences. The aim is then to bring the constrained model closer to the configural model by freeing up parameters of interest, in this case, covariances between latent variables. Modification indices (MIs) given by the software show where the greatest improvements in model fit can be achieved. Where there is a significant difference between groups, it will improve the model to estimate the covariances indicated by the MIs separately across groups, rather than constraining them to be the same. Parameters are freed up in a step-by-step basis, and only incorporated in the model if they lead to a significant improvement in the adjusted model chi-square.

This process will lead to a model that has the same latent structure across groups, but for which some parameters (in this case, covariances between latent variables), are estimated separately. This enables key differences between groups to be explored.

Links between motivation and proficiency

To explore the relationship between motivational variables and proficiency, two key pieces of information were used to run profile and correlation analysis in SPSS 22 (IBM Corp., 2013)³:

- scores from the Aptis test, both overall and for each test component
- factor scores (F-scores) for each participant for each motivational scale.

Aptis test outcomes are available as CEFR levels, as well scale scores for both individual components and the overall test. The CEFR levels assign participants into a broad proficiency banding, while the scale scores provide a more detailed insight into test performance. In the analysis carried out, the CEFR bands are used to set up comparative groups of students, while the scale scores are used for the more detailed correlational analysis.

The F-scores, meanwhile, were derived from the multi-group CFA described above and ascribe each questionnaire respondent with a level for each motivational scale. Essentially, once the CFA model is set up satisfactorily, a value to reflect participants' levels for each latent variable is calculated. So, for example, if a given participant gave strongly positive responses to the questions on English self-concept, he/she would have a higher F-score for this motivational scale than a respondent who provided low or mixed responses to the same questions. The exact balance of the relationship between observed responses and F-scores is determined by the factor loadings estimated in the model. The F-scores themselves run on a standardised zero-centred continuous scale (i.e. between -1 and 1).

3 P-values are derived from a comparison of the correlation coefficients following a Fisher's z-transformation. Calculations performed using this internet resource: http://vassarstats.net/rdiff.html



N 2 6 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á , P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 2 7

4. IMPLEMENTATION IN BOGOTÁ, COLOMBIA

The main objective of the sample design for English Impact is to present the most accurate possible results based on the comparison unit chosen – government-funded schools in the city of Bogotá, Colombia. In order to meet the established participation standards described in Chapter 3, a precisely defined comparison unit had to be outlined and agreed.

EDUCATIONAL CONTEXT IN COLOMBIA

Colombia is the oldest democratic state in Latin America. While its constitution, which was established in 1886, has seen a number of changes, the general structure of the government remains largely unaltered. Colombia's national government is split into a judicial branch, a legislative branch and an executive branch, which leads the republic. The president is elected for a four-year term. as are members of the Senate and the House of Representatives. Judges selected for the Supreme Court of Justice serve eight-year terms.

The most significant change to the constitution was in 1991, when the move towards today's decentralised Colombian state was made concrete. The Constitution of 1991 led to a number of reforms, the most notable of which devolved significant authority to departments, while general responsibilities remained with the national government.

Colombia is made up of 32 departments as well as Bogotá D.C., the capital district. Each department is led by a governor, who serves for four years and is elected by the people, as well as a state assembly, whose members are elected every four years. Each department comprises a number of municipalities that elect their own mayors and councils for four-year terms. Each departmental and municipal government has a budget for, and authority over, issues including but not limited to security, education, planning and development as well as the monitoring and evaluation of existing programs⁴.

The Ministry of Education (Ministerio de Educación Nacional, MEN) is responsible for national education policies. However, the unitary government structure means that local governments also have education secretariats which are responsible for education in their respective areas. Colombia's 94 education secretariats provide funding and guidelines for the schools under their jurisdiction, but they are not compelled to follow the national standards set by the central government.

The education system in Colombia consists of pre-primary, primary and lower-secondary school, followed by upper-secondary school, or high school. There are public and private institutions at each education level, with around 85 per cent of primary schools and 60 per cent of secondary schools publically funded. While nine years of education are compulsory by law, students receive 14 years of education on average.

EDUCATIONAL CONTEXT IN BOGOTÁ

Bogotá is socio-economically divided in strata 1 to 6, in which strata 1 to 3 represents the low- medium income population. Regionally, Bogotá is divided into 20 localities which, depending on the territorial arrangement plan, can be also divided in strata. The Secretary of Education has 26 official rural schools only in eight localities: Ciudad Bolívar, Suba, Sumapaz, Usme, Chapinero, Usaquén, Santa Fe and San Cristóbal. The largest number of official rural schools concentrated in the town of Sumapaz⁵.

The Secretary of Education of Bogotá is the guiding entity of preschool, basic and middle education in Bogotá. Its main functions include: ensuring the quality and coverage of education; establishing the educational policy in the district; and defining the sectorial development plans. Additionally, the Secretary organises and supervises the educational service provided by official and private entities, and approves the creation and functioning of the pre-school, basic and middle school formal education institutions in the city.

In 2016, Bogotá reached 7,980,001 inhabitants; this represents 16.28 per cent of Colombia's total population (48.7 million). Children under 15 years old comprise 22.67 per cent (1.8 million) of Bogotá DC's population; those between 15 and 64 years (the population of working age – PWA) comprise 69.54 per cent; and the remaining 7.79 per cent is over 65 years old.

4 Georgetown University Center for Latin American Studies, Political Database of the Americas: Decentralization and local governance, 2005. Retrieved from: http://pdba.georgetown.edu/Decen/decen_e.html

of bringing education to rural areas], El Tiempo. Retreived from: http://www.eltiempo.com/bogota/retos-de-la-educacion-en-zonasrurales-de-bogota-106036

of the Educational sector for 2016] retrieved from: https://www.educacionbogota.edu.co/archivos/SECTOR_EDUCATIVO/ESTADISTICAS_ EDUCATIVAS/2017/Caracterizacion_Sector_Educativo_De_Bogota_2016.pdf

G W 2 8 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á ,P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 2 9 3 Q J H

The school-age population of Bogotá DC in 2016 was 1,453,290 students, according to the projections of DANE and the District Department of Planning. In Bogotá, the population aged 15 and over has an average of 9.6 years of education, compared to the national average of 7.3 years.

In 2016, 827,615 students were at publicly-funded schools. The largest number of students enrolled in district schools was at the primary level (39.6 per cent), followed by secondary (35.3 per cent), middle school (14.6 per cent) and preschool (10.4 per cent).

The educational offer of the official sector is made up of places offered by the official educational system, as well as contracted administration and non-official institutions with contracted enrolment, modalities through which the District Capital seeks to guarantee the education rights. The types of schools were as follows:

- 361 district schools operated directly by the Secretary
- 22 schools with contracted administration: these schools have contracted out the administration to non-profit educational organisations, so that they operate autonomously
- 73 unofficial schools: run by private institutions which have an agreement with the Secretary to guarantee the right to education of children belonging to SES strata 1, 2 and 3 (in a small proportion); they are in areas of the city where the supply of educational quotas through official infrastructure is insufficient⁶.

5 Malaver, C (2017, 6 July). Pobreza e inasistencia, los retos de llevar educación al área rural [Poverty and non-attendance, the challenges

The 20 localities of Bogotá DC

6 B E N G L I S H ? I M P A C T



THE COMPARISON UNIT

- Government-funded schools in Bogotá
- 150 schools and 1,800 students were sampled for participation

 Table 3: Stratification variables chosen

As described above, implicit stratification variables were chosen to ensure a proportional sample allocation across the implicit strata. The stratification variables applied to the Bogotá sample frame are outlined in Table 3.

Stratification variable name	Variable labels
Location	Urban/ Rural
Socio-economic status	Low/ Medium/ High income



5. LANGUAGE LEARNING ENVIRONMENT

From the 1,800 students sampled for participation, 1,479 students participated in English Impact Bogotá and completed a 53-item questionnaire translated into Spanish, as

described in Chapter 3 above. Answers reported to 12 questionnaire items are presented below as a demographic profile.

Table 4: Demographic variables of participating students from Bogotá

Demographic variables	Reported by participating students in Bogotá
Gender	52.3% female; 47.7% male
Age	Mean age of 15.9 years
Language most often spoken at home	99.4% Spanish; 0.6% other
Country of birth	99.3% Colombia; 0.7% other
Attendance at pre-school	92.9% attended pre-school; 7.1% didn't attend pre-school

The gender ratio of female and male students was a near equal split, 52.3 per cent and 47.7 per cent respectively. As described in Chapter 3, the target average age of the target population was 15.5 years. The mean age reported by participating students was 15.9 years old, showing the accuracy of students sampled from the target population. The largest proportion, 99.4 per cent, indicated they spoke Spanish most often at home.

When asked to report their country of birth, over 99.3 per cent selected Colombia. A small number of students, 0.3 per cent, reported they most often spoke English at home.

Participating students were asked to report when they started learning English: 15.5 per cent said pre-school; 32.3 per cent said Grade 1: 14.8 per cent said Grade 3; and over 11 per cent said they started learning English in Grade 6

Figure 2: The grade participating students from Bogotá started learning English



Students were asked if, in their current grade, they chose to study English: 18.1 per cent said they chose to study English at school; 75.9 per cent said it was a compulsory school subject; and 6 per cent said their parents suggested they study English. When asked how long they spend





each week at school learning English, the largest number, 62.9 per cent, said they spent between three and four hours per week studying English, as shown by Figure 3. A smaller group, 9.5 per cent, spent between 5 and 6 hours studying English.

Figure 3: Time spent learning English per week at school



described their extracurricular language learning: 0.3 per cent said they learnt in private classes; 0.9 per cent learnt on their own using books and magazines; 1.3 per cent with a one-to-one tutor; 3.7 per cent learnt at an English language school; and 12.4 per cent said on their own using the Internet.

Those students who studied English outside of school were asked to indicate what activity best

Participating students were also asked to provide

information on studying English outside of school;

only 18.7 per cent said they chose to study

English outside of school.

Figure 4: Activity undertaken when studying English outside of school



The same segment of the participating student population was asked how much time they spent learning English outside of school. Almost half, 48.9 per cent, studied between one and two hours per week, in addition to their regular school hours.

A further 18.8 per cent spent three to four hours on English language learning, while 9.1 per cent studied English for five to six hours. Those who spent 30 minutes or less each week on English accounted for 23.2 per cent.





To understand more fully the possible further pathways of Bogotá school students in Grade 10, all participating students were asked what they would like to do when they finished school. Nearly two-thirds, 60.9 per cent, said they want to go to college or into further education; 23 per cent

Figure 6: Future pathway when finished compulsory schooling



The impact of media usage on language learning is often identified anecdotally by teachers and policymakers as a catalyst for accelerated proficiency, especially among the digitally native generation of millennial learners that comprise the target population of English Impact Bogotá.

Percentage

would like to continue studying to get into university; 9 per cent want to start an apprenticeship; 2 per cent want to get a job; another 2 per cent wish to start their own business; and 1 per cent would like to start a training course. The remaining 2 per cent said they did not know.

To gauge and understand their language use when interacting with various types of media, they were asked to indicate how they watch, read or listen to the following things most regularly, in English or in Spanish

Over half (55.4 per cent) of the participating students said they used English regularly to play computer games. In terms of social networks and the Internet, 15.8 per cent and 16.7 per cent respectively, indicated they interacted in English.

Social Networks

Magazines

Books

Internet

Radio

Films

ΤV

Computer Games

Type of media

15.8%

16.7%

25.0%

20

10

5.0%

4.6%

6.4%

6.4%

0

A quarter of the students watched films in English, while only 6.4 per cent each watched TV and listened to the radio in English. Just 5.0 per cent read magazines and 4.6 per cent read books in English.

Participating students were also asked what they did most often when using a computer, tablet or smart phone. The largest group, 44.4 per cent, most often use it for social media, such as Facebook, Instagram and Twitter.

Another 14.8 per cent use it most regularly to watch videos. School work was the top usage for 13.9 per cent of the respondents, while 9.1 per cent used it for their own research or collecting games for 7.3 per cent. Only 2.4 per cent said they did not have access to a computer, tablet or smart phone

43.1%

80

90

100

Figure 8: Activity most often done on computer, tablet or smart phone





data. Online communication was the most regular use for 7.8 per cent and computer

73.4%

60

Spanish

70

83.2%

94.5%

94.7%

81.9%

92.5%

50

Percentage

Figure 7: Language most regularly used to watch, read or listen to different types of media

55.4%

30

40

English



6. SAMPLING RESULTS

The sampling implementation process was carried out by a cross-organisational team of colleagues from ACER, representatives from the Bogotá Secretary of Education, and the British Council in Bogotá and the UK. With a focus specifically on the comparison unit of Bogotá, all students fulfilling the target population criteria detailed in Chapter 3 are described in Table 5.

The overall rate of school-level and within-

cent outlined in participation standard 1.3.

successfully achieved.

school exclusions was within the rate of 5 per

Therefore, a high standard of participation was

 Table 5: Target population definition in Bogotá

Grade 10 equivalent in Bogotá	Years of formal schooling	Average age at time of testing	Information about age of entry, promotion and retention
Grade 10	10	15.5	Students start at ISCED 1 in the year they turn six years old. Every child is required to be enrolled in Level 1 and Level 2 by law. In secondary education, every student is expected to undertake at least two hours a week of English instruction.

The target school population across the country was 100 per cent of all government-funded schools throughout Bogotá. There were a number of school-level exclusions (2.8 per cent) but no within-school exclusions (0 per cent).

 Table 6: Coverage and exclusions

Coverage	Notes on coverage	School-level exclusions	Within-sample exclusions / refusals	Overall exclusions / refusals
100%	All schools in Bogotá's public education system	2.8%	0.0%	2.8%

As described in Chapter 3, a process of implicit stratification was used to ensure a proportional sample allocation across all implicit strata. The stratification variables applied to the Bogotá sample frame were urban or rural location (on the outskirts of the city) and socio-economic status.

The first stage of the two-stage cluster sample process was to draw the school sample from the complete school sample frame. A total of 150 schools were drawn, as shown in Table 7.

Table 7: School sample size

Number of Schools in original sample	Number of eligible schools in original sample	Number of schools in original sample that participated	Number of replacement schools that participated	Total number of schools that participated
150	150	130	17	147 [†]

tWhile 149 schools took part overall, of the schools participating from the original sample, two had a within-school response rate lower than 50%, and so they were deducted from these final estimates of school-level participation.

The second stage of the two-stage cluster sampling process was the random selection of eligible students, from the target grade, within each participating school to take part in the assessment. Table 8 shows the total number of students in all sampled schools, the total number

N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B :: E = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B :: E = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 P B = 6 P W ? 9 B E B 1 7 P B = 6 P W ? 9 B E B 1 7 P B = 6 P W ? 9 B E B 1 7 P B = 6 P W ? 9 B E B 1 7 O A G W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W ? 9 B E B W 8 F W ?

The sample process involves the selection of two substitute schools that can be used in the event of the first sampled school not being able to participate. Seventeen substitute schools were used in Bogotá, due to non-response from the first sampled schools. There was no participation from one of the original sampled schools. (While 149 schools took part in total, two of the schools had a within-school response rate lower than 50 per cent, so they were deducted from the final estimates of school-level participation.)

withdrawn, excluded, eligible, and absent on the day of the assessment. The total number of students that took part in the assessment across the region was 1,479, of which 1,470 attended schools that achieved the required response rate

Table 8: Student sample size

Within- school student participation (weighted percentage)	Number of sampled students in participating schools	Number of students withdrawn from school	Number of students excluded	Number of eligible students	Number of students absent	Number of students assessed
80%	1,788	24	0	1,764	294	1,470 [†]
†This figure does not	include the nine partic	ipating students from th	e schools that did not r	neet the 50% participat	ion rate for estimation a	it the school level.

The results from these additional students are, however, included in the final reported results, bringing this number to 1,479 students.

A comparison between the school and student population and an estimate calculated from the draw sample is carried out as a check on the accuracy of the sampling procedure, as shown in Table 9. The table shows the actual number of schools and students in Bogotá and the number of participating sampled schools and students, and an estimate of the student population size based on the sample data. The population figures are derived from the sampling frame

used to select the sample, while the achieved sample figures refer to the number of sampled schools and students who participated in the assessments. The achieved sample figures were calculated using sampling weights and used as a check on the accuracy of the sampling procedure. As shown, the population size estimated from the sample closely matched the population size from the sampling frame.

 Table 9: Population and sample size

Bogotá population		Achieved sample population			
Schools	Students	Schools	Students	Student population size estimated from sample	
397	55,918	149	1,479	55,807	

The English Impact Bogotá response rate is an important participation standard and indication of the successful implementation of the school and student sampling procedure. Table 10 shows the weighted school and student participation rates. The weight applied to each school corresponds to the number of schools and students that they

represent in the entire population. Each student within each school had a weight equal to:

Total population of students in stratum

Total number of students participating in assessment

Table 10: Weighted school and student participation rates

School participation		Student	Overall participation		
Before replacement	After replacement	participation	Before replacement	After replacement	
86.7%	98.0%	83.9%	72.7%	82.2%	

The weight applied can vary from school to school. The weighted response rates, which take into account the weight each school has in the total sample, that is, the number of students it represents. As outlined in participation standard 1.7, the main survey response rate should reach at least 85 per cent of all sampled students across responding schools, with an overall participation rate after replacement of schools of 82.2 per cent. In Bogotá, this standard was reached when rounded to zero decimal places.

W **4 0** 1 6 B **E N G L I S H** ? **I M P A C T** B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B **B O G O T Á ,**P **C O L O M B I A** // G W B :: 1 7 B

Overall school and student participation in Bogotá met the English Impact international participation standard of at least 84 per cent of sampled students in 98 per cent of sampled schools.

7. ASSESSMENT OUTCOMES

The English language assessment taken by all participating students contained five components, testing four language learning skills: listening, reading, speaking and writing, plus grammar and vocabulary. Over two-thirds of the participating population from Bogotá (73 per

cent) achieved at A1 CEFR level in their overall English language test performance, while 21.4 per cent achieved at A2 level, 3.8 per cent at B1 level, and 1 per cent at B2 level. A0 denotes those students who did not achieve the minimum threshold for an A1 level.

 Table 11: Distribution of student population for overall test performance by CEFR level

	Overall English language test performance						
CEFR Level	Percentage (%)	Standard error (%)	95% CI				
С	0.2	0.06	[0.1, 0.3]				
B2	1.0	0.10	[0.8, 1.2]				
B1	3.8	0.28	[3.3, 4.4]				
A2	21.4	0.61	[20.2, 22.6]				
A1	73.0	1.46	[70.2, 75.9]				
AO	0.5	1.82	[0.0, 4.1]				
Missing	0.1	0.00	[0.1, 0.1]				
Total	100	-	-				

As described in Chapter 3, the 95 per cent confidence interval is in the region of \pm 1.96 Standard Error around the estimate and provides a measure of the certainty of the estimate. Table 12 shows the mean scores for each component

(a score out of 50) and overall (a score out of 200). As shown in this table, the standard error of the mean estimates for all four skills are small, indicating a high level of precision for these estimates.

Table 12: Mean population score by skills and overall achievement

Overall score	Grammar vocabulary	Listening	Reading	Speaking	Writing	Overall
Mean	12.19	21.63	14.35	3.39	7.38	53.71
SE	0.160	0.213	0.241	0.331	0.364	1.207
95% CI	(11.88 – 12.51)	(21.22 – 22.05)	(13.87–14.82)	(2.74 – 4.03)	(6.67 – 8.09)	(51.35 – 56.08)

Listening The highest mean scale score achieved 21.63

RECEPTIVE SKILLS

The receptive skills – listening and reading – were assessed using the computer-based test delivered via tablet and using individual headphones for the listening component.

The highest mean scale score was for the listening skill, 21.63, as shown in Table 12 above. With respect to CEFR levels, the estimated distribution of the listening skill performance

	Listening com	oonent		Reading component			
CEFR level	Percentage (%)	Standard error (%)	95% CI	Percentage (%)	Standard error (%)	95% CI	
С	0.6	0.21	[0.2, 1.0]	0.8	0.22	[0.3, 1.2]	
B2	1.3	0.28	[0.7, 1.8]	1.0	0.26	[0.4, 1.5]	
B1	23.3	1.27	[20.8, 25.8]	7.3	0.77	[5.8, 8.9]	
A2	71.9	1.39	[69.2, 74.6]	57.2	1.23	[54.8, 59.6]	
A1	2.5	0.47	[1.5, 3.4]	30.1	1.36	[27.4, 32.8]	
AO	0.4	0.16	[0.1, 0.7]	3.6	0.45	[2.7, 4.4]	
Missing	0.1	0.06	[0.0, 0.2]	0.1	0.06	[0.0, 0.2]	
Total	100	-	-	100	-	-	

shown in Table 13 indicates that 71.9 per cent of participating students achieved at A2 level, while nearly a quarter, 23.3 per cent, achieved at B1 level for their listening skill.

The mean achievement for the Bogotá population in reading skills is 14.35, the second highest of the four components assessed. Of the participating students, 57.2 per cent achieved at A2 level on the CEFR, while 30.1 per cent were at A1 level.

 Table 13: Estimated distribution of student population for listening and reading skills by CEFR level

PRODUCTIVE SKILLS

The productive skills – speaking and writing – were also tested using the computer-based English language assessment via tablet, with additional individual headphones, a microphone to capture speech responses, and a keyboard to enable students to type with as much ease as possible.

Participating students achieved their lowest performance scores for the productive skills. Speaking skills achieved the lowest mean scale

score, 3.39, as shown in Table 12 above. The distribution of speaking skills when referenced against the CEFR shows the largest proportion of students, 78.6 per cent, achieved at A0 level, while 12.7 per cent achieved at A2 level. See Table 14.

The third highest mean scale score was for writing skills at 7.38. The estimated distribution of writing performance when referenced against the CEFR shows that half (50.1 per cent) of all students achieved at A0 level, while 39 per cent were at A1 level.

Table 14: Estimated distribution of student population for speaking and writing skills by CEFR level

	Speaking			Writing		
CEFR level	Percentage (%)	Standard error (%)	95% CI	Percentage (%)	Standard error (%)	95% CI
С	0.1	0.00	[0.1, 0.1]	0.0	0.00	[0.0, 0.0]
B2	0	0.13	[0.0, 0.4]	0.4	0.23	[0.0, 0.9]
B1	3.4	0.58	[2.2, 4.5]	3.3	0.64	[2.1, 4.6]
A2	5.1	0.70	[3.8, 6.5]	7.1	0.77	[5.6, 8.6]
A1	12.7	1.10	[10.5, 14.8]	39.0	1.65	[35.8, 42.3]
AO	78.6	1.77	[75.1, 82.0]	50.1	2.18	[45.8, 54.3]
Missing	0.1	0.06	[0.0, 0.2]	0.1	0.06	[0.0, 0.2]
Total	100	-	-	100	-	-

Male and female students were almost equivalent on the overall scores, with males marginally higher

COMPARING ACHIEVEMENT BY GENDER

The ratio of participating female and male students was quite evenly distributed at 53.3 per cent and 54.16 per cent respectively.

Table 15: Mean performance score by gender, skills and overall achievement

Gender	Grammar and vocabulary	Listening	Reading	Speaking	Writing	Overall		
Female	12.06	21.66	14.19	3.16	7.32	53.30		
Male	12.34	21.60	14.52	3.63	7.44	54.16		
Difference (F-M)	-0.27	0.06	-0.33	-0.47	-0.12	-0.86		
Standard error	0.197	0.297	0.412	0.367	0.390	1.339		
Comparison (95% confidence)			•					
Positive difference 🔺 No difference 🕨 Negative difference 🔻								

As shown in Table 15, male students had a higher overall score and consistently achieved higher mean scale scores for all test components, apart from Listening, which was almost equivalent. However, these differences were not significant and could therefore have been the result of random variation.

COMPARING ACHIEVEMENT BY URBAN AND RURAL SCHOOL LOCATION

One of the stratification variables selected for the Bogotá sample frame was an urban or rural marker for each of the participating schools.

The rural schools were located on the outskirts of Bogotá. Analysis of the mean performance of schools classified as urban or rural highlights the different levels of performance at both types of schools.

Table 16: Mean performance score by school location, skills and overall achievement

School Location	Grammar vocabulary	Listening	Reading	Speaking	Writing	Overall
Urban	12.21	21.66	14.38	3.43	7.43	53.90
Rural	10.94	19.73	11.46	0.32	3.21	40.22
Difference (U-R)	1.27	1.93	2.92	3.11	4.22	13.67
Standard error	0.171	0.347	0.430	0.462	0.388	5.905
Comparison (95% confidence)						

Positive difference

No difference Negative difference

As shown by the row describing comparative difference in the 95 per cent confidence intervals, there was a positive and significant difference between urban and rural schools' performance, with the schools in urban areas outperforming those in rural areas in all domains. Urban schools had an overall mean score of 53.90, while the score for rural schools was 40.22.

COMPARING ACHIEVEMENT BY SOCIO-ECONOMIC STATUS

An additional implicit stratification variable used for Schools with medium socio-economic status had the Bogotá comparison unit was socio-economic a higher performance in all domains, with an status. Table 17 shows the comparative overall overall mean score of 58.91, compared to 50.44 for schools with low socio-economic status

Table 17: Comparing overall mean score performance by socio-economic status

By school socio- economic status	Grammar and vocabulary	Listening	Reading	Speaking	Writing	Overall	
Low	11.92	21.12	13.90	2.67	6.50	50.44	
Medium	12.62	22.46	15.06	4.52	8.77	58.91	
Difference (L-M)	-0.70	-1.34	-1.16	-1.85	-2.27	-8.46	
Standard error	0.332	0.452	0.551	0.733	0.000	0.000	
Comparison (95% confidence)	•	•	•	•	•	•	
Positive difference 🔺 No difference 🕨 Negative difference 🔻							

mean scale score performance for each socioeconomic status.

8. ENGLISH LANGUAGE LEARNING MOTIVATION

The analysis presented in this chapter explores two key areas:

- 1. the levels and character of language learning motivation for different groups of students (split by gender and socio-economic status)
- 2. relationships between each of the motivational scales and English language proficiency.

Questionnaire responses

1.200

1,000

800

400

The guestionnaire received a full response rate, largely because students could not proceed through the app without answering all questions. It can be seen from Figure 10 that some of the questions received overwhelmingly positive responses, whereas in other cases, a much wider spread of options was chosen.

Questions about English as a medium of communication received particularly positive responses. Items about the role of English on the job market, learners' future vision of themselves as successful language learners and parental encouragement also yielded very positive responses (see example response profile given in the left-hand image of Figure 10). However, the responses to scales on effort investment, societal expectations and evaluation of one's ability to study English, were much more balanced (see examples given in the centre and right-hand images of Figure 10).

The upcoming section accounts for nuances in the patterns of response to this part of the survey.

Figure 10: Responses to a sample of questionnaire items



g19 17c : Studying English is important to me because me more if I have knowledge of English other people will respect



CONSTRUCT VALIDATION FINDINGS

Measurement model

As described in Chapter 3, the aim of the construct validation exercise was to establish a measurement model that satisfactorily reflects the pattern of observed responses. The value of a well-fitting measurement model is that it provides a statistically sound means of estimating levels of motivation on each of the scales for all participating students. This, in turn, enables comparisons to be drawn between groups of students using further CFA modelling techniques.

The hypothesised measurement model has the structure shown in Figure 11 below. Please cross-reference the terms with Table 2 above. This initial model showed this structure to have a borderline acceptable fit to the data according to accepted thresholds for CFA models (CFI = 0.919; TLI = 0.907; RMSEA = 0.054). In order to improve fit, the scale of international orientation had to be removed, as it did not emerge in the analysis as a distinct factor. A closer look at the associated data indicated that the items elicited two distinct patterns of response. This meant that the scale did not function as a cohesive factor; it is unclear exactly why this may be, it may relate to the interpretation this group of students brought to these particular items, or that there was some additional noise in the data. However, as it stands, inclusion of this scale was not helpful to the overall model.

- 7 Decisions regarding the inclusion of the covariances mentioned in the table above were made with reference to the 'modification indices' generated by the Mplus software. However, no covariances were included that did not make sense substantively. Hence, these were restricted to covariances between errors for questions against the same motivational scale.
- 8 See Brown (2015, pages 157–162) for a fuller discussion of correlated measurement error and the implications in CFA

? // G W **4 8** 1 6 B **E N G L I S H** ? **I M P A C T** B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B **B O G O T Á .**P **C O L O M B I A** // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H **4 9** 3

Additionally, three items from other scales (mb4, ex2 and instr4 indicated in Figure 11 by crossing out) had to be removed, as they did not load on the intended factors. The possible explanation behind such a mismatch might be differences in the meaning of these items due to translation from English into the language of the survey, Spanish. Removing these observed variables resulted in a model with reasonable fit to the data according to accepted thresholds for CFA models (CFI = 0.944; TLI = 0.934; RMSEA = 0.050).

It was found that the model could be further improved by allowing some of the error terms (or residual error) for individual questions to correlate with each other. These relationships are summarised in Table 18⁷. Including these enables the model to take into account commonality between two observed variables in addition to that explained with reference to the latent variable, perhaps reflecting something in the wording of both questions that provokes a particular shared response⁸.

The final measurement model is shown in Figure 12, this included eight correlated error variances listed in Table 19 (indicated by orange arrows) and did not include a latent variable for International Orientation. The improved model provides a good fit for the data (CFI = 0.962; TLI = 0.953; RMSEA = 0.043). Final model estimates are given in Appendix B.

Table 18: Improvements in adjusted chi-square values following addition of covariances between error terms

		Most constrained	INSTR1 with	SELF1 with	PAR1 with
Constrained model:		model	INSTR2	SELF 3	PAR4
Degrees of freedom	d0	254	253	252	251
Scaling correction factor	c0	1.2686	1.2675	1.2674	1.2677
MLM chi-square value	TO	1209.069	1124.723	1072.757	1023.467
More relaxed model:		Add INSTR1 with INSTR2	Add SELF1 with SELF 3	Add PAR1 with PAR4	Add O2 with O3
Degrees of freedom	d1	253	252	251	250
Scaling correction factor	c1	1.2675	1.2674	1.2677	1.2655
MLM chi-square value	T1	1124.723	1072.757	1023.467	983.221
Calculations					
Diff test scaling correction	cd	1.5469	1.2927	1.1921	1.8177
Adjusted chi-sq value	Calc	108.2385309	65.9741807	62.1631059	53.1829404
Adjusted chi-sq value	TRd	69.971	51.036	52.146	29.258
Constrained model:	d0	02 with 03	I1 with I2	EX1 with EX3	PAR1 with PAR2
Degrees of freedom	c0	250	249	248	247
Scaling correction factor	TO	1.266	1.265	1.265	1.263
MLM chi-square value		983.221	945.03	927.659	914.298
More relaxed model:	d1	Add I1 with I2	Add EX1 with EX3	Add PAR1 with PAR2	Add O1 with O2
Degrees of freedom	c1	249	248	247	246
Scaling correction factor	T1	1.265	1.265	1.263	1.261
MLM chi-square value		945.03	927.659	914.298	904.36
Calculations					
Diff test scaling correction	cd	1.2904	1.3894	1.6601	1.7307
Adjusted chi-sq value	Calc	48.425	22.445	18.363	14.273
Adjusted chi-sq value	TRd	37.527	16.155	11.062	8.247

Figure 11: Hypothesised measurement model

Double-headed arrows represent covariances between each of the latent variables:

Circles show the latent variables



?// G W 5 0 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á ,P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 5 1 3 Q J H

Boxes show the observed variables (question responses) Figure 12: Measurement model following modifications (final model)

Double-headed arrows represent covariances between each of the latent variables:

Circles show the (motivational scales)

Boxes show the (question responses) Red double-headed arrows indicate correlated error terms



Table 19: Error covariances included in the measurement model

Covariance ref	Question ref	Question details	Reduction in adjusted chi-square value	Significance	
1	INSTR1	I need English for my future career.	69.971	p<.001	
	INSTR2	The things I want to do in the future require me to use English.			
2	SELF1	I usually get good marks in English.	51.036	p<.001	
	SELF3	I have always done well in English.			
3	PAR1	My parents think I need to know English to be well educated.	52.146	p<.001	
	PAR4	My parents encourage me to practice my English as much as possible. I work hard at learning English.			
4	02	Studying English is important to me because other people will respect me more if I have knowledge of English.	29.258	p<.001	
	03	Studying English is important to me because an educated person is supposed to be able to speak English.			
5	11	I imagine myself speaking English fluently.	37.527	p<.001	
	12	I imagine myself comfortably reading in English on the Internet.			
6	EX1	Learning English is really great.	16.155		
	EX3	I find learning English really interesting.			
7	PAR1	My parents think I need to know English to be well 11.062 educated.		p<.001	
	PAR2	My parents have stressed the importance English will have for me in the future.			
8	01	I consider learning English important because the people I respect think that I should do it.	8.247	p<.005	
	02	Studying English is important to me because other people will respect me more if I have knowledge of English.			

?// G W 5 2 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á, P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 5 3 3 Q J H

The structural model

Having established the viability of the measurement model, it is insightful to examine the association between the latent variables as estimated by the model. Standardised covariances (correlations) estimated by the model are given in Table 20. All correlations are significant and positive, indicating the interlinked nature of the motivational traits. The values given range from those classified as moderate to strong correlations (Cohen, 1988).

 Table 20: Standardised (STDYX standardisation) correlations between variables in the final model

Latent variables		Two-tailed					
Estimate		S.E.	Est. /S.E.	P-value			
IDEAL	EX	0.870	0.016	54.853	0.000		
INSTR	IDEAL	0.854	0.019	44.410	0.000		
INSTR	EX	0.808	0.020	39.679	0.000		
MB	EX	0.778	0.018	43.118	0.000		
MB	INSTR	0.652	0.022	29.165	0.000		
MB	IDEAL	0.650	0.021	30.741	0.000		
OUGHT	INSTRU	0.619	0.026	23.516	0.000		
OUGHT	IDEAL	0.547	0.027	20.508	0.000		
OUGHT	EX	0.502	0.029	17.557	0.000		
OUGHT	MB	0.435	0.028	15.561	0.000		
PAR	INSTR	0.736	0.024	31.130	0.000		
PAR	IDEAL	0.657	0.025	26.632	0.000		
PAR	OUGHT	0.646	0.024	27.483	0.000		
PAR	EX	0.590	0.027	22.229	0.000		
PAR	MB	0.511	0.025	20.325	0.000		
SELF	EX	0.751	0.018	40.660	0.000		
SELF	MB	0.710	0.020	35.231	0.000		
SELF	IDEAL	0.656	0.020	32.314	0.000		
SELF	INSTR	0.566	0.025	22.230	0.000		
SELF	PAR	0.427	0.027	15.789	0.000		
SELF	OUGHT	0.385	0.030	12.695	0.000		

Most of the variables are strongly related as they represent different aspects of language learning motivation. It can be noticed, however, that two latent variables in particular are generally less related to other areas of motivation. These are 'parental encouragement' and the 'ought-to L2 self'. This could be ascribed to the fact that, unlike other latent variables, they focus on external pressures connected with studying English. The first one refers to levels of parental encouragement and the second one relates to pressures from broader environmental by representing what is expected of the participants in terms of studying English. In this sense, they can be seen as more externalised motives to study English (Ryan and Deci, 2000). Most importantly, both parental encouragement and the ought-to L2 self scales are least closely related to motivated learning behaviour, the latent variable that aims to capture reported levels of effort invested in language learning. This finding is in line with previously reported results, which indicate that the influence of parental encouragement and the ought-to L2 self on motivated behaviour is limited (Csizér and Kormos, 2009; Iwaniec and Ullakonoja, 2016; Taguchi, Magid and Papi, 2009)

MULTI-GROUP ANALYSIS FINDINGS

Comparative analyses were conducted for two different groupings of the data:

- gender
- socio-economic status (SES).

Gender

The first group comparison is between male and female students. The final sample included 774 female learners and 705 male learners. It is prudent not to assume identical motivations drive these groups since empirical studies consistently suggest that female students tend to be more motivated to study English than their male peers (see Iwaniec, 2015 for a review). Studies point to higher achievement in language learning among female learners than male learners (Fernandez Fontecha, 2010; Jimenez Catalan, 2010). Considering that motivation is shown to affect language learning achievement (Hsieh and Kang, 2010; Kim and Kim, 2014; Marsh and Martin, 2011), understanding where the differences lie with respect to underlying motivational traits will help formulate targeted policies on foreign language learning and teaching. However, it needs to be noted that the majority of studies examining gender differences took place in Australia, Canada, Europe and, to some extent, Asia. In contrast, no such studies were conducted in South America.

A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B :: E = ? 9 W B 0 A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? 2 // G W 5 4 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á, P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 5 5 3 Q J H 3 % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 M S W ? 9 B E B 1 7 0 A G W B :: 7 = ? 9 =

Consequently, this exploration of gender differences in Bogotá is of particular importance.

Levels of motivation

- As can be seen in Table 21, significant differences across gender groups were found between the levels of motivation reported on four out of seven motivational scales. In all cases, female students reported higher levels of motivation than their male peers (significance level from t-test indicated in parenthesis):
- motivated behaviour (p=.011)
- parental encouragement (p=.017)
- instrumentality (p=.022)
- ideal L2 self (p=.032).

Table 21: T-tests to check for differences between mean factor scores calculated from the gender model

	Group comparisons				t-test statistics		
Motivational variable	Gender	N	Mean	Std. deviation	t	Degrees of freedom	Sig. (2-tailed)
Ideal L2 Self	female	774	.0453	.84004	2.146	1477	.032*
	male	705	0485	.83874			
Parental encouragement	female	774	.0508	.89667	2.380	1477	.017*
	male	705	0616	.91845			
English self-concept	female	774	0180	.79829	476	1477	.634
	male	705	.0012	.74531			
Language learning experience	female	774	.0272	.83945	1.694	1477	.091
	male	705	0455	.80822			
Instrumentality	female	774	.0531	.93521	2.284	1477	.022*
	male	705	0563	.90326			
Motivated behaviour	female	774	.0585	.92592	2.552	1477	.011*
	male	705	0624	.89181			
Ought-to L2 self	female	774	0215	.82552	-1.048	1477	.295
	male	705	.0243	.85609			

** p<.01 *p < .05

The results imply that, compared to male students, female learners reported to:

- invest more effort in language learning
- receive more positive parental encouragement to study English
- have more robust visions of themselves as successful users of English in the future
- be more likely to learn English with a view to using it to progress in their future professional careers.

These results are consistent with findings from previous empirical studies (see Iwaniec, 2015 for a review). Research in the area implies that there might be a number of reasons why female learners express higher motivation than male learners.

• General perceptions of language learning as a female domain foster the construction of femininity among females learning languages,

whereas they have the opposite effects for the perceptions of masculinity of males learning a foreign language (Carr and Pauwels, 2005).

- Language proficiency is perceived as more directly relevant for the future careers of female than male learners (Clark, 1998); the perception that is even more augmented by the perceived limited choice of professional careers for females (Norton and Pavlenko, 2004).
- Females are perceived as having a stronger preference for, and value communication more, than male learners (Chavez, 2000).
- Preferred learning styles and use of language learning strategies of female learners (visual and auditory learning) is more convergent with the requirements of language classes than in the case of male learners who are often kinaesthetic learners (Oxford, 1993).

Relationships between motivational scales

Analysis shows that freeing up one covariance This results in a model with good overall fit leads to a significant improvement in the statistics (CFI= 0.958; TLI = 0.954; RMSEA = adjusted model chi-square. This means that the 0.042). A summary of the covariance estimates for following relationship is significantly different this gender-specific model is given in Table 22 between the groups of female and male students:

Table 22: Latent covariances allowed to vary between gender groups (standardised estimates)

Covariance	Estimate	S.E.	Est. /S.E.	Two tailed P-value
IDEAL WITH EX				
Female model	0.771	0.023	33.689	.000
Male model	0.728	0.024	30.889	.000

As can be seen, the strength of the covariance is higher for female than male learners. The findings imply that having confidence in one's ability to study English is more closely related to the positive language learning experience for female than male learners.

Previous research findings show that girls perceive foreign language classes as more rewarding than boys, which might help explain this difference. Clark (1998), for example, found that language classes have more 'intrinsic appeal' to girls than boys, as girls deal better with relatively slow progress, a dose of repetition and textbook-based language tasks. Moreover, when working together, girls appear to collaborate and support each other; in contrast, boys are often competitive, which allows female learners to benefit more from communicative language learning (Oxford, 1994).

Oxford (1993) also reported that girls are more likely to prefer auditory learning styles, whereas boys were more likely to be kinaesthetic learners.

? // G W **5 6** 1 6 B **E N G L I S H** ? **I M P A C T** B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B **B O G O T Á ,**P **C O L O M B I A** // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H **5 7** 3 Q J H

• language learning experience with English self-concept (12.209 on 1 d.f., p<.01).

As girls' preferred learning styles are more likely to be useful in the language class, it is likely that their language experience is more positive.

Socio-economic status (SES)

The final group comparison is drawn between students differing in the socio-economic status (SES) indicators. Research suggests that educational attainment tends to differ according to students' SES – students with lower SES attain less than their peers who have a higher SES (OECD, 2013). Studies of language learning motivation report similar findings; namely, levels of motivation tend to fall in line with students' SES (Iwaniec, 2015; Kormos and Kiddle, 2013; Lamb, 2012, 2013). Hence, it is vital to examine the differences in the language learning motivation of students from different socio-economic backgrounds in order to address this at a policy level.

The analysis reported here is based on a measure of SES constructed from the information given by the students to the questions listed in Table 23⁹.

noted that certain categories were over-represented (general labour, for example), whereas others were under-represented. This could be due to the age of participants who might tend towards selecting the more generic job description. Instead the focus has been placed on the information regarding the general level of that job.

Table 23: Questions used in the construction of SES variable

Q48	What is the highest level of school completed by either of your parents, whoever completed the highest level of school?	 Pre-primary[1] Primary [2] Lower secondary [3] Upper secondary [4] Post-secondary non tertiary [5] Undergraduate [6] Postgraduate [7]
Q50	What level is that job [†] ?	 Lower level support worker [1] Mid-level skilled worker [2] Mid-level professional [3] Senior level manager [4]
Q51	Do you have any of these in your home?	 [Grid answer format YES or NO] A tablet computer [1] A car [2] A bedroom of my own [3] Air conditioning [4] A smart TV [5] My own smartphone [6]

¹This refers to the job given in response to Q49 ("What kind of job does the same parent (as Q48) do?")

SES scores are calculated using a data reduction technique called Principal Component Analysis (PCA), in which each of the relevant variables is used to create a single, weighted scale (after Caro and Cortés, 2012). From this, it is possible to calculate each participant's SES score. The weights accorded to each variable are determined by the data, as PCA takes into account the common variation of the variables determined to be of value in describing SES. Based on the inspection of frequencies and distribution of these SES scores, the decision was made to split the sample into two groups, later referred to as students with low and medium SES. As the distribution of data was normal, the difference in the scores on the continuous variable between

the students who are in the top of lower SES category and bottom of higher SES category is, in some cases, minimal. Thus, care needs to be taken when analysing the results. However, for the ease of the process, it was decided that the two-way split was a workable solution. It should also be noted that this individual-level SES variable is not the same as the school-level variable used to compare achievement in Chapter 6. This variable is a relative measure based on the individual responses to specified questions as detailed above. The description of the technique as 'formative' by Caro and Cortés (2012, page 17) relates to the fact that the variables that contribute towards the scale are fixed in the analysis, rather than beforehand.

Levels of motivation

As can be seen in Table 24, significant differences were found on four out of seven aspects of motivation examined. In all cases, students with higher SES have higher scores on the motivational scales than learners with lower SES (significance level from t-test indicated in parenthesis):

	Group comparisons			t-test statistics			
Motivational variable	SES	N	Mean	Std. deviation	t	Degrees of freedom	Sig. (2-tailed)
Ideal L2 self	lower	781	0486	.80342	-2.349	1477	.019*
	higher	698	.0537	.87132			
Parental encouragement	lower	781	1171	.88581	-5.006	1477	.000**
	higher	698	.1182	.92081			
English self-concept	lower	781	0561	.72161	-2.970	1477	.003**
	higher	698	.0642	.82494			
Language learning	lower	781	0242	.80789	-1.361	1477	.174
experience	higher	698	.0343	.84454			
Instrumentality	lower	781	0561	.88896	-2.471	1477	.014*
	higher	698	.0627	.95937			
Motivated behaviour	lower	781	0092	.86535	524	1477	.600
	higher	698	.0159	.96043			
Ought-to L2 self	lower	781	0306	.79510	-1.588	1477	.112
	higher	698	.0393	.88787			

** p<.01 *p < .05

Compared to learners with lower SES, therefore, students with higher SES reported to:

- have more robust visions of themselves as successful English users in the future
- be more confident in their ability to study English
- receive more encouragement and support from their parents to study English

W 5 8 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á , P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 5 9 3

- ideal L2 self (p=.019)
- English self-concept (p=.003)
- parental encouragement (p<.001)
- instrumentality (p=.014).

Table 24: T-tests to check for differences between mean factor scores calculated from the gender model

• be more likely to learn English with a view to use it to progress in their future professional career.

These findings clearly point to an advantage for students with higher SES over students with lower SES in terms of language learning motivation. The members of the higher group are not only more likely to have language learning goals involving English, but they also collectively

evaluate themselves as more capable of learning English and receive more parental support in their learning pursuits. These findings are in line with previous studies in the field of language learning motivation (Iwaniec, 2015; Kormos and Kiddle, 2013; Lamb, 2012).

Previous research shows a number of possible explanations why students with higher SES report higher levels of language learning motivation than their peers with lower SES.

- Students with higher SES have much easier access to positive role models. Their parents tend to be better educated, with higher levels of English competence and in better employment, where English is more likely to be used. This vicarious experience strengthens these students' confidence in their ability to learn English, and shows them that English is useful on the job market (see Iwaniec, 2015 for a review).
- Higher economic status enables students with higher SES to travel abroad more freely, to have access to private education, extra classes and language learning materials that their peers with lower SES are less likely to be able to access.
- University-educated parents tend to have higher aspirations for their children, engage more in activities promoting learning (Davis-Kean, 2005; Eccles, 1994) and provide a more supportive learning environment for their children (Klebanov, Brooks-Gunn and Duncan, 1994) than parents who have lower levels of education.

At the same time, there are no significant differences between the two SES groups in terms of their language learning attitudes, the perceived societal pressure to study English and the reported effort to study English. Iwaniec (2018) also reported lack of significant differences in Polish students' language learning attitudes when students in Poland were divided according to their parents' education and school location, both proxy measures of SES. Similarly, she reported that the influence of SES on self-regulation, a measure taking into account effort investment.

The motivation profiles for the two SES groups are shown Figure 13. This helps visualise where the greatest discrepancies lie between groups.





To examine the role of SES in greater depth, learners were divided into 10 groups in order of SES level, with students in the 1st percentile having the lowest SES factor scores and those in the 10th percentile having the highest SES factor scores. The results presented in Figure 14 show that students in the 9th and 10th percentile reported distinctly higher motivation in terms



of all aspects of motivation, than students from more modest backgrounds. The differences between students from the 1st to 8th percentiles are minimal. This suggests that the divisive line in terms of language learning motivation is between the top 10% to 20% of students and the remaining groups.

Figure 14: Mean factor scores on motivational variables for SES percentile groupings



Relationships between motivational scales

In order to conduct this comparative analysis, the split between higher and lower SES was used. Analysis shows that freeing up two covariances leads to a significant improvement in the adjusted model chi-square. This means that these relationships are significantly different between students from lower and higher SES groups:

- motivated behaviour with English self-concept (4.654 on 1 d.f., p<.05)
- ought-to L2 self with English self-concept (5.003 on 1 d.f., p<.05).

This results in a model with good overall fit statistics (CFI= 0.958; TLI = 0.954; RMSEA = 0.042). A summary of the covariance estimates for this gender-specific model is given in Table 25.

 Table 25: Latent covariances allowed to vary between gender groups (standardised estimates)

Covariance	Estimate	S.E.	Est./S.E.	Two tailed P-value
MB with SELF				
Lower SES	0.696	0.026	27.158	.000
Higher SES	0.720	0.026	27.639	.000
OUGHT with SELF				
Lower SES	0.359	0.037	9.832	.000
Higher SES	0.405	0.035	11.572	.000

As can be seen in the table, the covariances are stronger for learners from higher SES group than those from lower SES group. These findings imply that having confidence in one's ability to study English is more closely related to: reported effort investment for higher SES learners, than their peers from lower SES backgrounds perceiving societal pressure to learn English for higher SES learners, than their peers from lower SES backgrounds.

These results suggest that societal pressure is a greater source of English self-concept for students with higher SES. This might be because their parents, who are more highly educated, are more likely to be involved in professions where English is a clear advantage. As a result, this heightens their awareness of what is expected of them in terms of studying English.

Additionally, the level of English language selfconcept amongst the higher SES students is more strongly tied the likelihood of investing effort in studying English, which gives them the advantage over students from lower SES backgrounds.

LINKS BETWEEN MOTIVATION AND PROFICIENCY

In this section, the relationship between Bogotá students' language learning motivation and their proficiency in English is examined. While this is clearly of interest at policy level, there are important caveats to interpreting the findings of such analysis in the current context. First, the reader needs to keep in mind that language learning can be affected by a plethora of factors (Ortega, 2009), only one of which is motivation. Some of them are specific to individuals, for example, aptitude or 'gift for languages', anxiety, language learning strategies, cognition and personality traits. The amount of exposure to language and the opportunities to use this language are also crucial. In addition, it should be noted that in this study, the language learning motivation of Bogotá students was measured when they took the proficiency test. However, motivation is dynamic and changes over time. Hence, whereas the proficiency measure is a cumulative measure of what students have achieved over years of learning, the data on motivation presents a single snapshot.

N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E B // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E B // 1 7 B = 6 > ? 9 W B 8 F W ? 9 B E B 1 7 O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 O A G W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W

Finally, there appears to be relatively little variation in the English proficiency of Bogotá students, as compared to their motivational scores. Considering these points, it is expected, therefore, that motivation can explain only some variance in proficiency. Nonetheless, a focus on the comparative values leads to some interesting insights.

Presentation of the results looks first to give an overview of the motivational profiles for students, who achieve at different levels on the overall CEFR scale, followed by correlations between Aptis test scores and motivational scales for gender and SES groups. Implications of the findings are discussed.

Learner proficiency and motivational profiles

This analysis groups the students by their CEFR level and compares the motivation profiles for each group, based on the calculation of factor scores (F-scores) derived from the final CFA model described above. The vast majority of learners achieved level A1 (1,072) or A2 (323), with seven students achieving A0, 58 achieving B1, 15 with B2, and 3 with C. For this analysis, students were divided into three groups, according to their overall CEFR band, which takes into account their performance on each of the four skill areas (listening, reading, speaking and writing):

- Low scorers students in A0 band and those scoring 41 or less points classified in band A1, 520 learners (53.7% female, 46.3% male; 63.3% lower SES, 36.7% higher SES), 35.2% of the sample;
- Average scorers students in band A1 with scores over 41, 559 learners (52.6% female, 47.4% male; 53.8% lower SES, 46.2% higher SES), 37.8% of the sample
- High scorers students in bands higher than A1, 399 learners (50.4% female, 49.6% male; 37.8% lower SES, 62.2% higher SES), 27% of the sample.

Figure 15 shows the relative patterns of mean factor scores for each of the motivational scales for the three groups. It can be seen that the

factor scores for each group are consistently rising in tandem with their proficiency profiles, albeit with varying degrees of divergence.

?// G W 6 4 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á ,P C O L O M B I A // G W B

Figure 15: Motivational profiles for learners at different levels of proficiency



 Table 26: Significant differences between groups according to their proficiency



Darker colour = higher factor score

Table 26 summarises the significant differences between learners at different proficiency levels. It can be seen that the three groups differ significantly from each other on five aspects of language learning motivation. Moreover, there is a significant difference between low scorers and those who score higher in terms of effort investment and the perceived societal pressure to study English, but average and high scorers do not differ in this respect.

tivational scale							
Language learning experience	Instrumentality	Motivated behaviour	Ought-to L2 self				

Different shades = statistically significant difference

Correlations between proficiency and motivation

The results presented in Table 27 show the strength of correlations between the overall scale score¹⁰ and the factor scores for motivational scales. The correlations are listed in descending strength.

10 The total score achieved by the participants in the four skill areas, out of a possible total of 200 points (50 per component). In this

Table 27: Correlations between motivational variables and overall total score

6 6 1 6 B E N G L I S H ? I M P A C T B

Motivational variables	Total score
English self-concept	.385*
Language learning experience	.301*
Ideal L2 self	.288*
Instrumentality	.249*
Motivated behaviour	.245*
Parental encouragement	.225*
Ought-to L2 self	.104*

* p<.01

As can be seen, all motivational variables are positively related to proficiency. There are, however, large differences in correlation strength. This can be clearly observed when comparing scatterplots showing the bivariate relationship between factor scores for the scale most closely related to proficiency (English self-concept, shown in Figure 16) and least closely related to proficiency (the ought-to L2 self, shown in Figure 17). In the first case, it is visible that the scores on English self-concept tend to grow in tandem (though loosely) with the overall proficiency scores, although with a considerable amount of variation, whereas in the case of the ought-to L2 self, the points are more widely dispersed.

In terms of assessing the strength of correlations, there are two medium correlations with proficiency (English self-concept, language learning experience) and five small correlations (ideal L2 self, instrumentality, ought-to L2 self and motivated behaviour) and one small correlation (parental encouragement) (interpretation according to Cohen, 1988). The order of strength of correlation mirrors the results from the covariances between latent variables estimated in the confirmatory factor analysis described above, with variables displaying the strongest relationships with proficiency representing the more internalised motives for studying English, and those with the weaker relationships more externalised.

Figure 16: Correlation between overall proficiency and English self-concept

W8FW?4>JHY4HS83QJH?//GWB**BOGOTÁ**, P**COLOMBIA**//GWB



Figure 17: Correlation between overall proficiency and the ought-to L2 self





Overall scale

Table 28 shows correlations between motivational variables and different components of the proficiency test. Compared to the overall proficiency scores, correlations with scores on individual components are slightly weaker. This is because the measure of overall proficiency is a composite of skills, hence it is a more comprehensive scale.

There are small differences in the strength of correlation depending on the component of the English proficiency assessed. Correlations with writing tend to be strongest, whereas correlations with listening and reading tend to be lowest. There is some variation in the order of strength of correlations between different components and motivational variables.

 Table 28: Correlations between motivational variables and scale scores for individual components of the proficiency test

Motivational scales	Listening	Reading	Speaking	Writing	Grammar and vocabulary
English self- concept	.256**	.330**	.322**	.370**	.307**
Language learning experience	.218**	.240**	.240**	.297**	.233**
Instrumentality	.191**	.183**	.200**	.245**	.185**
Ideal L2 self	.212**	.218**	.236**	.294**	.216**
Motivated behaviour	.170**	.100**	.199**	.113**	.197**
Parental encouragement	.172**	.156**	.195**	.221**	.152**
Ought-to L2 self	.065*	.068**	.105**	.107**	.093**

** p<.01 *p < .05

The following analysis examines where differences lie in terms of these correlations for the different student groupings, split by gender and SES as per the analysis reported above.

Correlations by gender groups

The results of correlational analysis of the cohort split along gender lines (Table 29) show

Table 29: Correlations between overall proficiency and motivational variables, by gender

Motivational scale	Female	Male	P-value [†]
English self-concept	.340	.433	.036*
Ideal L2 self	.273	.308	.465
Language learning experience	.302	.297	.912
Instrumentality	.241	.263	.652
Motivated behaviour	.231	.267	.459
Parental encouragement	.201	.252	.303
Ought-to L2 self	.070	.135	.208

[†]Based on Fisher's z-score ** p<.01 *p < .05

This finding is in contrast to those from previous studies, which tended to report that boys' confidence was inflated in comparison to girls and was not justified in their achievement (Bandura, 1977; Skaalvik and Skaalvik, 2004). More investigation would be necessary to understand the processes explaining this finding.

A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B I A // G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B I A // G W B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 8 8 3 Q J H ? // G W B B O G O T Á, P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 6 9 3 Q J H 6 % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B H 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ?

that estimates of male students' motivation indicate a closer relationship with proficiency, than that of their female peers. However, only one of these differences is significant: the one between English self-concept and proficiency. This indicates that boys' confidence in language learning is more closely related to their proficiency than that of girls.

	Mala		D	t	
anu	motivational	variables	5, UY	yenue	I

Correlations by socio-economic status groups

The results of the correlational analysis of motivational variables and overall proficiency are presented in Table 30. The differences in the strength of correlation between motivation and proficiency are significant for all the variables, with the exception of parental encouragement, for which the significance is borderline. In all cases, motivational variables are more closely related to the proficiency scores for the higher than lower SES group. This suggests that students from the higher SES group are more likely to translate their motivation into proficiency gains than students' from the lower SES group. The less advantaged students appear to be encountering more obstacles when learning English and higher proficiency is for them more of a dream, than for more advantaged students, a finding that has been previously reported by Lamb (2013) in the context of Indonesia.

Of particular significance is the finding that motivated behaviour is more directly related to proficiency gains for higher SES group, which implies that students from lower SES learn English less efficiently.

Table 30: Correlations between overall proficiency and motivational variables, by SES

Motivational scale	Lower SES	Higher SES	P-value [†]
English self-concept	.321	.427	.018*
Ideal L2 self	.213	.339	.009**
Language learning experience	.238	.352	.016*
Instrumentality	.180	.294	.020*
Motivated behaviour	.178	.302	.011*
Parental encouragement	.157	.255	.050*
Ought-to L2 self	.035	.153	.023*

[†]Based on Fisher's z-score ** p<.01 *p < .05

SUMMARY OF ENGLISH LANGUAGE LEARNING MOTIVATION FINDINGS

The findings discussed in this chapter give a greater depth to the information about English language levels by exploring some of the underlying patterns of motivation among teenage language learners. The final measurement model included seven of the eight originally hypothesised motivational traits; the International Orientation scale did not provide information about a clearly distinguishable underlying trait.

The findings of the study also point to an advantage for students with higher SES in terms of language learning motivation. In particular, the levels of motivation of the students estimated to be within the top 10% of the SES measure were Levels of motivation were found to increase in line distinctly higher than the remainder of their peers with English language proficiency, although the in lower SES groups. In addition, motivational distinction was less marked in the case of oughtvariables are more closely related to the to self variable, which reflects social expectations, proficiency scores for the higher than lower SES and the motivated learning behaviour variable. group. This suggests that students from the Overall, however, correlations between motivation higher SES group are more likely to translate their and proficiency were not strong, with the closest motivation into proficiency gains than students links being seen with English self-concept and from the lower SES group. language learning experience.

A::: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 P 9 B E B 1 7 ? A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? Y G W 7 0 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á, P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 7 1 3 Q J H W 8 :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W ? 9 B E B W 8 F W 8 F W ?

Female learners reported generally higher levels of motivation. However, analysis revealed that male students' motivation had a closer relationship with proficiency than that of their female peers. In particular, boys' confidence in language learning is more closely related to their proficiency than that of girls.

9. IN CONCLUSION

The research aims outlined at the beginning of this report established the intention to evaluate the English language capability of students studying at government schools in Bogotá; to compare the outcomes in schools by gender, rural and urban locations, and SES group, and to understand the relationship between English language learning motivation and increased proficiency. An evaluation of capability, as described in the introductory chapter, involves the assessment of both the level of achievement reached by a defined population; and the opportunities provided to them to achieve greater proficiency via teaching and learning practice derived from a policy or national guideline. The aim of this research was not to create a single score to show the success, or otherwise, of the policy for English language teaching and learning in Bogotá. The detailed description of the research processes and outcomes presented in this report are intended to provide policymakers, teachers and researchers with a full analysis of the complex and intertwined elements that combine to influence students learning outcomes.

Three central factors came together to produce this ground-breaking research. The research design, combining ACER's sampling expertise and the British Council's knowledge and experience in English language assessment, enabled a collaboration leading to world-class research outcomes. Complex field operations within schools often provide significant challenges to the successful completion of large-scale data collection, therefore, the consistent use of off-line enabled tablets to deliver all language assessments in every classroom has contributed to the success of English Impact Bogotá. Lastly, the positive collaboration with teachers, schools and students participating in the research process has been invaluable. Without their positive and proactive cooperation, this evaluation would not have produced the strong and reliable evidence upon which further discussions and policy decisions may be based.



REFERENCES

Ajzen, I (2005) *Attitudes, personality and behaviour.* New York: Open University Press.

Bachman, LF (1990) *Fundamental considerations in language testing.* Oxford, England: Oxford University Press.

Bandura, A (1997) *Self-efficacy: The exercise of control*. New York: WH Freeman & Company.

Bong, M and Skaalvik, EM (2003) Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, 15(1), 1–40.

Brown, TA (2015) Confirmatory Factor Analysis for Applied Research (2nd Edition). New York: Guildford Press.

Byrne, B. (2012). Structural equation modeling with Mplus : Basic concepts, applications, and programming (Multivariate applications book series). New York: Routledge.

Caro, D. H., & Cortés, D. (2012). Measuring family socioeconomic status: An illustration using data from PIRLS 2006. IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments, 5, 9–33. Available online at: http://www. ierinstitute.org/fileadmin/Documents/IERI_ Monograph/IERI_Monograph_Volume_05_ Chapter 1.pdf

Carr, J and Pauwels, A (2005a) *Boys and foreign language learning.* Basingstoke: Palgrave Macmillan.

Carr, J and Pauwels, A (2005b) *Boys and foreign language learning: Real boys don't do languages.* Basingstoke: Palgrave Macmillan.

Chávez, M (2000) *Gender in the language classroom.* London: McGraw Hill.

Clark, A (1998) *Gender on the agenda: Factors motivating boys and girls in MFLs.* London: Centre for Information on Language Teaching and Research.

Cohen, J (1988) Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, teaching, assessment.* Cambridge: Cambridge University Press.

Csizér, K and Kormos, J (2008a) Modelling the role of inter-cultural contact in the motivation of learning English as a foreign language. *Applied Linguistics*, 30(2), 166–185. Available online at: http://doi.org/10.1093/applin/amn025

Csizér, K and Kormos, J (2008b) The relationship of intercultural contact and language learning motivation among Hungarian students of English and German. *Journal of Multilingual and Multicultural Development*, 29(1), 30–48. Available online at: http://doi.org/10.2167/ jmmd557.0

Csizér, K and Kormos, J (2009) 'Learning experiences, selves and motivated learning behaviour: A comparative analysis of structural models for Hungarian secondary and university learners of English' in Z Dörnyei and E Ushioda (eds), *Motivation, language identity and the L2 self.* Bristol: Multilingual Matters, 98–119.

Davis-Kean, PE (2005) The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2), 294–304. Available online at: http://doi.org/10.1037/0893-3200.19.2.294 Dörnyei, Z and Ushioda, E (eds) (2009) *Motivation, language identity and the L2 self.* Bristol: Multilingual Matters.

Eccles, JS (1993) 'School and family effects on the ontogeny of children's interests, self-perceptions, and activity choice' in J Jacobs (ed), *Nebraska Symposium on Motivation: Vol. 40. Developmental perspectives on motivation.* Lincoln University, 145–208.

Eccles, JS (1994) Understanding women's educational and occupational choices. Applying the Eccles et al. model of achievement-related choices. *Psychology of Women Quarterly*, 18, 585–609. Available online at: http://doi. org/10.1111/j.1471-6402.1994.tb01049.x

Edele, A, Seuring, J, Kristen, C and Stanat, P (2015) Why bother with testing? The validity of immigrants' self-assessed language proficiency. *Social Science Research*, 52, 99–123. Available online at: http://doi.org/10.1016/j. ssresearch.2014.12.017

Erling, EJ and Seargeant, P (2013) *English and development: Policy, pedagogy and globalization.* Bristol: Multilingual Matters.

Fernandez Fontecha, A (2010) 'Gender and motivation in EFL vocabulary production' in RM Jimenez Catalan (ed), *Gender perspectives on vocabulary in foreign and second languages*. Basingstoke: Palgrave Macmillan, 93–116.

Ford, M (1992) *Motivating humans: Goals, emotions, and personal agency beliefs.* London: Sage Publications.

Gardner, RC and Lambert, WE (1959) Motivational variables in second language acquisition. *Canadian Journal of Behavioural Psychology*, 13, 266–272.

Gardner, RC and Lambert, WE (1972) 'Motivational variables in second-language acquisition' in RC Gardner and WE Lambert (eds), *Attitudes and motivation in second-language learning.* Rowley, Mass.: Newsbury House.

:: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B V ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B V ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B V ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B V ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G W B V ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B 9 8 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? // G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? // G W 7 4 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á, P C O L O M B I A // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 7 5 3 Q J H % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8</td> W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W 8 F W 8 F W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W ? 9 B E B W 8 F W 8 F W ?

Graddol, D (2006) *English next: Why Global English may mean the end of 'English as a foreign language'.* London: British Council.

Henry, A and Cliffordson, C (2013) Motivation, gender, and possible selves. *Language Learning*, 63(2), 271–295. Available online at: http://doi. org/10.1111/lang.12009

Hsieh, PP-H and Kang, H-S (2010) Attribution and self-efficacy and their interrelationship in the Korean EFL context. *Language Learning*, 60(3), 606–627. Available online at: http://doi. org/10.1111/j.1467-9922.2010.00570.x

Hsieh, PP-H and Schallert, DL (2008) Implications from self-efficacy and attribution theories for an understanding of undergraduates' motivation in a foreign language course. *Contemporary Educational Psychology*, 33(4), 513–532. Available online at: http://doi.org/10.1016/j.cedpsych.2008.01.003

IBM Corp. (2013) IBM SPSS Statistics for Windows. Armonk; NY: IBM Corp.

Iwaniec, J (2014) Motivation of pupils from southern Poland to learn English. *System*, 45(1), 67–78. Available online at: http://doi. org/10.1016/j.system.2014.05.003

Iwaniec, J (2015) *Motivation to learn English of Polish gymnasium pupils*. Lancaster University. Available online at: http://opus.bath.ac.uk/49838/

Iwaniec, J and Ullakonoja, R (2016) Polish and Finnish teenagers' motivation to learn English: The role of context. *European Journal of Applied Linguistics*, 4/2, 277–300. Available online at: http://doi.org/10.1515/eujal-2015-0034

Jimenez Catalan, RM (2010) 'Gender tendencies in lexical acquisition and use' in RM Jimenez Catalan (ed), *Gender perspectives on vocabulary in foreign and second languages*. Basingstoke: Palgrave Macmillan, 117–138.

Kim, T-Y and Kim, Y-K (2014) 'EFL students' L2 motivational self system and self-regulation: Focusing on elementary and junior high school students in Korea' in K Csizér and M Magid (eds), *The impact of self-concept on language learning*. Bristol: Multilingual Matters, 117–138.

Klebanov, PK, Brooks-Gunn, J and Duncan, GJ (1994) Does neighborhood and family poverty affect mothers' parenting, mental health, and social support. *Journal of Marriage and Family*, 56(2), 441–455. Available online at: http://www. jstor.org/stable/353111

Kormos, J and Kiddle, T (2013) The role of socio-economic factors in motivation to learn English as a foreign language: The case of Chile. System, 41(2), 399–412. Available online at: http://doi.org/10.1016/j.system.2013.03.006

Kormos, J, Kiddle, T and Csizér, K (2011) Systems of goals, attitudes, and self-related beliefs in second-language-learning motivation. *Applied Linguistics*, 32(5), 495–516. Available online at: http://doi.org/10.1093/applin/amr019

Kyriacou, C and Zhu, D (2008) Shanghai pupils' motivation towards learning English and the perceived influence of important others. *Educational Studies*, 34(2), 97–104. Available online at: http://doi.org/10.1080/03055690701811099

Lamb, M (2012) A self system perspective on young adolescents' motivation to learn English in urban and rural settings. *Language Learning*, 62(4), 997–1023. Available online at: http://doi. org/10.1111/j.1467-9922.2012.00719.x

Lamb, M (2013) 'Your mum and dad can't teach you!': constraints on agency among rural learners of English in the developing world. *Journal of Multilingual and Multicultural Development*, 34(1), 1–16. Available online at: http://doi.org/10.1080/01434632.2012.697467

Marsh, HW and Martin, AJ (2011) Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, 81, 59–77. Available online at: http://doi.org/10.1348/000709910X503501 Mills, N, Pajares, F and Herron, C (2007) Self-efficacy of college intermediate French students: Relation to achievement and motivation. *Language Learning*, 57(3), 417–442. Available online at: http://doi. org/10.1111/j.1467-9922.2007.00421.x

Moskovsky, C, Racheva, S, Assulaimani, T and Harkins, J (2016) The L2 Motivational Self System and L2 Achievement: A Study of Saudi EFL Learners. *The Modern Language Journal*, 16, 1–14. Available online at: http://doi.org/10.1111/ modl.12340

Muthén, LK and Muthén, BO (2011) *Mplus user's guide.* (6th ed.). Los Angeles, CA: Muthén and Muthén.

Norton, B and Pavlenko, A (2004) 'Gender and English language learners: Challenges and possibilities' in B Norton and A Pavlenko (eds), *Gender and English language learners*. Virginia: TESOL, 1–12.

O'Sullivan, B (2011) 'Language testing' in J Simpson (ed), *Routledge handbook of applied linguistics*. Oxford: Routledge.

O'Sullivan, B (2015) Aptis test development approach. Aptis Technical Report, TR/2015/001. London: British Council. Available online at: https://www.britishcouncil.org/sites/default/ files/tech_001_barry_osullivan_aptis_test_-_ v5_0.pdf

O'Sullivan, B and Dunlea, J (2015) *Aptis General Technical Manual* V 1.0. Aptis Technical Report, TR/2015/005. London: British Council. Available online at: https://www.britishcouncil.org/sites/ default/files/aptis_general_technical_ manual_v-1.0.pdf

O'Sullivan, B and Weir, CJ (2011) 'Language testing and validation' in B O'Sullivan (ed) *Language testing: theory & practice.* Oxford: Palgrave, 13–32. OECD (2013) PISA 2012 results: *Excellence through equity: Giving every student the chance to succeed* (Volume II). OECD Publishing. Available online at: http://doi.org/http://dx.doi. org/10.1787/9789264201132-en

Ortega, L (2009) *Understanding second language acquisition*. London: Hodder Education.

Oxford, R (1993) Instructional implications of gender differences in second/foreign language (L2) learning styles and strategies. *Applied Language Learning*, 5, 65–94.

Oxford, R (1994) 'La difference continue... Gender differences in second/foreign language learning styles and strategies' in J Sunderland (ed), *Exploring gender: Questions and implications for English language education.* Prentice Hall, 140–147.

Pfenninger, SE and Singleton, D (2016) Affect trumps age: A person-in-context relational view of age and motivation in SLA. Available online at: http://doi.org/10.1177/0267658315624476

Ryan, RM and Deci, EL (2000) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. Available online at: http://doi.org/10.1037110003-066X.55.1.68

Saito, M (2003) Amartya Sen's Capability Approach to education: a critical exploration. *Journal of Philosophy of, Education.* 37 (1), 17 – 33.

Satorra, A and Bentler, PM (2010) Ensuring positiveness of the scaled difference chi-square test statistics. *Psychometrika*, 75, 243. Available online at: http://doi.org/doi:10.1007/s11336-009-9135-y

Shavelson, RJ, Hubner, JJ and Stanton, GC (1976) Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46/3: 407–441.

:: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E // 1 7 B = 6 P W 8 B F W ? 9 B E B 1 7 B = 2 V 2 8 T Z Q U - < 2 5 3 5 % A A : F N O A G // X C K P W ? 9 B E // 1 7 B = 6 P W 8 B // 9 B E B 1 7 B = 2 P W 8 B // 9 B E B 1 7 P B = 2 P W 8 B // 9 B E B 1 7 P B = 2 P W 8 B // 9 B E B 1 7 P B = 0 P W 8 P W ? 9 B E B 1 7 P B = 0 P W 8 P W ? 9 B E B 1 7 P B = 0 P W 8 P W ? 9 B E B // 1 7 B = 0 P W 8 P W ? 9 B E B // 1 7 B = 0 P W 8 P W ? 9 B E B // 1 7 B = 0 P W ? 9 B E B // 1 7 B = 0 P W ? 9 B E B // 1 7 D A G W B :: 7 = ? 9 = W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B // 1 7 B = 0 P W ? 9 B E B // 1 0 A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B // 1 0 A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B // 1 0 A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B // 1 0 A G W B :: 7 = ? 9 = W

Shohamy, E (1998) Critical language testing and beyond. *Studies in Educational Evaluation*. 24 (4). 331–345.

Skaalvik, S and Skaalvik, EM (2004) Gender differences in math and verbal self-concept, performance expectations, and motivation. *Sex Roles*, 50(3/4), 241–252. Available online at: http:// doi.org/10.1023/B:SERS.0000015555.40976.e6

Taguchi, T, Magid, M and Papi, M (2009) 'The L2 motivational self-system among Japanese, Chinese and Iranian learners of English: A comparative study' in Z Dörnyei and E Ushioda (eds), *Motivation, language identity and the L2 self.* Bristol: Multilingual Matters, 66–98.

Taylor, F and Marsden, EJ (2014) Choosing to study foreign languages in England: An experimental intervention. *The Modern Language Journal*, 98(4), 902–920. Available online at: http://doi.org/10.1111/modl.12146

Ushioda, E (2009) 'A person-in-context relational view of emergent motivation, self and identity' in Z Dörnyei and E Ushioda (eds), *Motivation, language identity and the L2 self.* Bristol: Multilingual Matters, 215–228.

Weir, CJ (2005) *Language Testing and Validation: an evidenced-based approach*. UK: Palgrave Macmillan.

Williams, M and Burden, RL (1997) Psychology for language teachers: *A social constructivist approach*. Cambridge, UK: Cambridge University Press.

Yashima, T (2009) 'International posture and the ideal L2 self in the Japanese EFL context' in Z Dörnyei and E Ushioda (eds), *Motivation, language identity and the L2 self.* Bristol: Multilingual Matters, 215–228.

APPENDIX A – OVERVIEW OF THE APTIS FOR TEENS TEST STRUCTURE

Test	Part	Skill focus	ltems/ Part	Level	Tasks/ level	ltems/ Task	Task focus	Task description	Response format	
	1	Grammar	25	A1	5	1	Syntax and word usage	Sentence completion: select the best word to	3-option multiple choice	
				A2	5-7	1		complete a sentence based on syntactic		
				B1	5-7	1		appropriacy.		
				B2	5-7	1				
	2	Vocabulary	25	A1	1	5	Synonym (vocabulary breadth)	Word matching: match two words which have the same or very similar meanings.	5 target words. Select the best match for each from a bank of 10 options.	
Core					A2	1	5	Meaning in context (vocabulary breadth)	Sentence completion: select the best word to fill a gap in a short sentence. Understanding meaning from context.	5 sentences, each with a 1-word gap. Select the best word to complete each from a bank of 10 options.
50 items						B1	1	5	Meaning in context (vocabulary breadth)	Sentence completion: select the best word to fill a gap in a short sentence. Understanding meaning from context.
					1	5	Definition (vocabulary breadth)	Matching words to definitions.	5 definitions. Select the word defined from a bank of 10 options.	
				B2	1	5	Collocation (vocabulary depth)	Word matching; match the word which is most commonly used with a word targeted from the appropriate vocabulary level.	5 target words. Select the best match for each from a bank of 10 options.	

Test	Skill focus	ltem/ Part	Level	Task/ level	ltems/ Task	Format	Task description	Response format
Listening 25 items	Lexical recognition	10	A1	10	1	Monologues	Q&A about listening text. Listen to short monologues (recorded messages) to identify specific pieces of information (numbers, names, places, times, etc.).	3-option multiple choice. Only the target is mentioned in the text.
	Identifying specific, factual information	5	A2	5	1	Monologues and dialogues	Q&A about listening text. Listen to short monologues and conversations to identify specific pieces of information (numbers, names, places, times, etc.).	3-option multiple choice. Lexical overlap between distractors and words in the input text.
	Identifying specific, factual information	5	B1	5	1	Monologues and dialogues	Q&A about listening text. Listen to short monologues and conversations to identify propositions. The information targeted is concrete and of a factual/literal nature. Requires integration of information over more than one part of the input text.	3-option multiple choice. Distractors should have some overlap with information and ideas in the text. Target and distractors (where possible) are paraphrased.
	Meaning representation / inference	5	B2	5	1	Monologues and dialogues	Q&A about listening text. Listen to monologues and conversations to identify a speaker's attitude, opinion or intention. The information targeted will require the integration of propositions across the input text to identify the correct answer.	3-option multiple choice. Both target and distractors are (where possible) paraphrased, and distractors refer to important information and concepts in the text that are not possible answers to the question.

Structure of the Teens Listening component

Structure of the Teens Core component

B E B 1 7 B = Z V 2 8 T Z Q U <</td> 2 5 3 5 % A 8 F W ? 9 B E B 1 7 B > ? 9 B E B 1 7 ? 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 7 9 3 Q J H 2 V 2 8 T I Q I 2 5 3 5 % A X G % L 1 U M 8 ? ? 9 B E B W 8 F W ? ? 9 B E B W 8 F W ? ? 8 S Y ? 8 F</

A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A :: F N O A G // X C K P W ? 9 B E // 1 7 B = 6 > ? 9 B E B 1 7 B = 6 > ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B ? 9 B E B 1 7 B = 6 > ? 9 W B ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? N O A G W B N 7 = ? 9 W N O A G W B :: E = ? 9 W B ? 9 B E B 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á P C O L O M B I A // G W B :: E = ? 9 W B ? 9 B E B 1 7 B = 6 P W ? 9 B E B 1 7 B = 6 P W ? 9 B E B W 8 F W ? 4 > J H Y 4 H S 8 3 Q J H ? // G W B B O G O T Á P C O L O M B I A // G W B :: G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 < Z V 2 8 T I Q U < 2 5 3 5 % A X G % L 1 U M 8 G % 2 A X G M L 1 9 8 1 ^ 9 # D R M < L % A 9 M 5 F W ? 9 B E B W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7 = ? 9 = W 8 F W 8 F W ? 9 B E B W 8 F W ? F W 8 F W ? F W B :: 1 7 B = 6 P W ? 9 B E B 1 7 O A G W B :: 7

Test	Part	Skill focus	ltems/ Part	Level	Tasks/ level	ltems/ Task	Task focus	Task description	Response format
	1	Sentence level meaning	5	A1	1	5	Sentence level meaning (careful local reading)	Gap fill. A short text with 5 gaps. Filling each gap only requires comprehension of the sentence containing the gap. Text-level comprehension is not required.	3-option multiple choice for each gap.
	2	Inter-sentence cohesion	6	Α2	1	6	Inter-sentence cohesion (careful global reading)	Re-order jumbled sentences to form a cohesive text.	Re-order 6 jumbled sentences. All sentences must be used to complete the story.
Reading 25 items	3	Text-level comprehension of short texts	7	B1 1 7 Text-level comprehension of short texts Candidates match 4 short paragraphs giving information about 4 people's opinions on different topics and identify which of the four people could 7 gaps in a short text. Select the best word to fill each gap from a bank of 9 options.					
	4	Text-level comprehension of long text	7	B2	1	7	Text-level comprehension of longer text (Global reading, both careful and expeditious)	Matching the most appropriate headings to paragraphs. Requires integration of micro- and macro-propositions within and across paragraphs, and comprehension of the discourse structure of more complex and	7 paragraphs forming a long text. Select the most appropriate heading for each paragraph from a bank of 8 options.
tructure of	f the Tee	ns Reading compo	onent					abstract texts.	

Structure of the Teens speaking component

B E B 1 7 B = Z V 2 8 T Z Q U <</td> 2 5 3 5 % A B F W ? 9 B E B 1 7 B > ? 9 B E B 1 7 ? 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 8 1 3 Q J H Z V 2 8 T I Q V 2 5 3 5 % A X G % L 1 U M 8 ? ? 9 B E B W 8 F W ? ? ? 8 F W ? ? ? 8 F W ? ? ? 8 E W</

Channel of input / Expected Test Part Skill focus Level Task description prompts output Rating criteria Writing at the A1 The candidate Form with 9 clearly 9 short gaps word level completes a form by marked categories filled by 1-2 Simple personal filling in some basic (name, date of birth, word responses information personal information. etc.) There are 9 on a form. All responses are at the gaps in the form to word/phrase level, such be filled. as name, birthdate, etc. Α2 2 Short written The candidate continues Written. The rubric 20-30 words description of filling in information on a presents the context, concrete, form. The task setting followed by a short Separate task-based personal and topic are related to question asking for holistic scales are information at the same purpose as the information from the used for each task. the sentence form used in part 1. candidate related to Performance level. The candidate must the context. descriptors describe write a short response the expected using sentence-level performance at each writing to provide score band. The personal information in following aspects response to a single of performance are written question. addressed (not all aspects are assessed B1 30-40 words in for each task): Writing 3 Interactive The candidate responds Written. The rubric writing. interactively to 3 presents the context response to each 1) task completion (discussion forum, separate questions. Responding to question a series of Each response requires social media, etc.). 2) grammatical written a short paragraph-level Each question is range and accuracy questions with response. The questions displayed in a 3) lexical range sequence following short are presented as if the paragraphcandidate is writing on the completion of and accuracy an internet forum or level the response to the 4) cohesion responses. social network site. The previous question. and coherence task setting and topic are related to the same 5) punctuation purpose/ activity used and spelling. in parts 1 and 2. Continuous The candidate writes an Written. The rubric Β2 220-250 words 4 paragraphargumentative essay on presents the level essay a topical issue the context in the form writing. candidate is likely to of an advert giving encounter in public or basic information educational domains. about an essav competition.

Structure of the Teens Writing component

	vv									vv										u	vv			
	۱۸/	۱۸/		F	14/	2		R	F					Δ									14/	
	4	H	S	8	3	2	%	A	X	G	X								8	1			8	F \
	2		Q		8	A			N	0	A	G	11	X	С	K	0		?	9	В	E	B	1
	В	Е	В			8	N		A															
		1	X				8																	
	W	W	8	F	W	?			E			/	9	A	G	W	B	0			2	9	W	
		H	0	8	3	Q	%		Ň															
Ŀ.	G	M	G	1	9				9	4												7		
	B	E	В						A	G							9	W	N	0	A	d	W	В
		1	Х						Т	T					N	Э	A	G	W			7	-	? !
	4	Н	S	8		Q			X				8		A.	9	46							FA
	G	M	L	1		8			9		D				Ē									
	R	F	R		0	A	N					5		A	-									
		W			W						1	7												
		Н	S	8		Q					X	9												
	В	Е			%		N				W	В												
	W	W		F	W	?			E		1													
	2		Q	1		A					Ê	6												
						0				# G	ł.	9							 8			9	8	F
				I						Ö	A	G	11						2	9			B	1
		1								11		÷/	B					G	W					2
			В							G							9					G		
	4	H		8	3																	9		FA
	G	M			9				9								A				<			4
	R	F	R		%	8	N	0		G							ď.		í.	9	A	G	W	R.
	W	W	8	E	W	?	9	B		B								0	7		2			*
		н	S	8	3	Q	%	A	X									9	8				8	E A
	В	E	В	11		8	Ν	0	А	G									N	0		G	Vé	
	W	W	8	F	W	?		В	E													9		
	2		Q	0	0	A	1		N	4											Б	1		
	4	H	S	8		0	%	A	X	G	X									4	~	9		
		~	Q		8	A		F	N	0	A									9	В	E	B	
		1	Х	9		9		>	Т	11	G										Т	7		
	B	E.	В				Ν	0	A	G	W									Ŷ.	А	G	2	8-
		H.	S	8	3	Q	%	A	X	G	8											9		
				÷	9	A		F	N	#* ()	7													
		E	В		%	8	Ν		A	G	W	B								5	A			в.
	W	W	8	F	W	?	9		Е	В	1	7	0								2			
	4	H	S	8	3	Q	%	A		G		9	8									9	8	
		E	В	2	%	8	N	0	A	G	W	B		6									VV.	
	2	VV	0	F	W Q	í A	,	B	N	B	Å		77	Â	6	K								
	G	M		1	9	8																		
	4	H	S	8	3	0																		
R			Q	ł		A																		
H	8		X	9																				
H	2	E	B			1																		
		M	-	1	9	0																		
	Đ		0	i.	8	Ă																		
	B		В	1	%	8																		
	W	W	8		N	2																		
	4 P	H	S	8	3	0																		
	W	W	8	F	/o W	2																		

 ? 9 > ?
 ?
 // G W B :: 1 G B = 6 % A X G 1 2 % 1 ^

 B = 6 > ?
 9 B E B V 7 B U Z V 2 8 T Z Q U ? <</td>

 E C ? 9 W B 8 F W ? 9 B E B 1 7 B 9 8 1 % > ?

 W N O A G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 F W 8 F W ? 9 B E B W 8 F W 8 F W ? 9 B E B ::

 1 9 2 7 // G W B :: E = ? 9 W B 8 F W ? 9 B E B ::

 1 9 2 7 // G W B :: E = ? 9 W B 8 F W 8 F W ? 9 B E B ::

 1 1 0 0 < 2 5 3 3 % A X G % L 1 0 M 8 :: 2 0 0 0 ?</td>

 1 1 0 0 < 2 5 3 3 % A X G % L 1 0 M 8 :: 2 0 0 0 ?</td>

 1 2 9 > ? // G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 Y N O A G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 Y N O A G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 Y N O A G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 Y N O A G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 Y N O A G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 Y N O A G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 Y N O A G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 Y N O A G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 Y N O A G W B :: E = ? 9 W B 8 F W ? 9 B 7 B 1

 Y A X G % L 1 U M 8 :: 2 A 0

 Y A Y G W B F W ? 9 B F B 0 Z V 2 8 T Z 0 0 7 7 1

 Y A Y G W B F W ? 9 B 8 F W ? 9 B 7 B 1

 Y A Y G W B 8 F W ? 9 B 7 B 1

 Y A Y G W B 8 F W ? 9 B 7 B 1

 Y A Y G W B 8 F W ? 9 B 7 B 1

APPENDIX B - FINAL CFA MEASUREMENT MODEL ESTIMATES

STANDARDISED MODEL RESULTS

STDYX Standardisation

	Estimate	S.E	Est./S.E	Two-Tailed P-Value
PAR BY				
PAR1	0.717	0.020	35.987	0.000
PAR2	0.749	0.016	47.077	0.000
PAR3	0.843	0.013	66.400	0.000
PAR4	0.718	0.017	41.890	0.000
SELF BY				
SELF1	0.720	0.017	43.058	0.000
SELF2	0.818	0.013	63.426	0.000
SELF3	0.794	0.014	57.845	0.000
SELF4	0.780	0.015	51.188	0.000
OUGHT BY				
01	0.688	0.020	34.959	0.000
02	0.648	0.023	28.291	0.000
03	0.672	0.021	32.491	0.000
04	0.820	0.015	54.335	0.000
EX BY				
EX1	0.738	0.016	45.747	0.000
EX3	0.783	0.015	52.606	0.000
EX4	0.839	0.010	80.450	0.000
MB BY				
MB1	0.807	0.014	58.845	0.000
MB2	0.728	0.015	47.594	0.000
MB3	0.894	0.010	87.035	0.000
INSTR BY				
INSTR2	0.782	0.016	49.667	0.000
INSTR3	0.781	0.014	54.125	0.000
INSTR1	0.701	0.021	32.893	0.000

	Estimate	S.E
IDEAL BY		
11	0.705	0.017
12	0.796	0.013
13	0.742	0.016
14	0.756	0.016
SELF WITH		
PAR	0.427	0.027
OUGHT WITH		
PAR	0.646	0.024
SELF	0.385	0.030
EX WITH		
PAR	0.590	0.027
SELF	0.751	0.018
OUGHT	0.502	0.029
MB WITH		
PAR	0.511	0.025
SELF	0.710	0.020
OUGHT	0.435	0.028
EX	0.778	0.018
INSTR WITH		
PAR	0.736	0.024
SELF	0.566	0.025
OUGHT	0.619	0.026
EX	0.808	0.020
MB	0.652	0.022
IDEAL WITH		
PAR	0.657	0.025
SELF	0.656	0.020
OUGHT	0.547	0.027
EX	0.870	0.016
MB	0.650	0.021
INSTR	0.854	0.019

B E B 1 7 B = Z V 2 8 T Z Q U <</td> 2 5 3 5 % A 3 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 8 5 3 Q J H Z V 2 8 T I Q U <</td> 2 5 3 5 % A X G % L 1 U M 8 = ? 9 B E B W 8 F W ? 8 F W ? 8 F W ? 8 F</t

Est./S.E	Two-Tailed P-Value
42.142	0.000
60.399	0.000
47.309	0.000
48.709	0.000
15.789	0.000
27.483	0.000
12.695	0.000
22.229	0.000
40.660	0.000
17.557	0.000
20.325	0.000
35.231	0.000
15.561	0.000
43.118	0.000
31.130	0.000
22.230	0.000
23.516	0.000
39.679	0.000
29.165	0.000
26.632	0.000
32.314	0.000
20.508	0.000
54.853	0.000
30.741	0.000
44.410	0.000

	Estimate	S.E	Est./S.E	Two-Tailed P-Value
INSTR1 WITH				
INSTR2	0.394	0.036	10.808	0.000
SELF1 WITH				
SELF3	0.280	0.033	8.526	0.000
PAR1 WITH				
PAR4	-0.211	0.035	-6.066	0.000
PAR2	0.161	0.042	3.849	0.000
O2 WITH				
03	0.274	0.037	7.457	0.000
I1 WITH				
12	0.239	0.034	7.114	0.000
EX1 WITH				
EX3	0.163	0.035	4.684	0.000
O1 WITH				
02	0.117	0.038	3.065	0.002
Intercepts				
11	3.751	0.086	43.520	0.000
12	4.006	0.101	39.717	0.000
13	3.396	0.066	51.516	0.000
14	3.552	0.079	45.012	0.000
PAR1	3.397	0.075	45.436	0.000
PAR2	3.133	0.065	47.955	0.000
PAR3	3.666	0.083	44.317	0.000
PAR4	2.893	0.056	52.077	0.000
SELF1	3.457	0.069	50.256	0.000
SELF2	2.975	0.056	52.972	0.000
SELF3	3.003	0.056	53.935	0.000
SELF4	2.856	0.052	54.927	0.000
01	2.796	0.053	52.532	0.000
02	2.586	0.047	54.505	0.000
03	2.929	0.057	51.076	0.000
04	2.770	0.054	51.364	0.000

	Estimate	S.E	Est./S.E	Two-Tailed P-Value
EX1	4.082	0.099	41.415	0.000
EX3	3.945	0.095	41.673	0.000
EX4	3.453	0.070	49.416	0.000
MB1	3.144	0.060	52.553	0.000
MB2	3.398	0.067	50.418	0.000
MB3	3.416	0.065	52.352	0.000
INSTR1	3.734	0.094	39.686	0.000
INSTR2	3.787	0.096	39.590	0.000
INSTR3	3.377	0.074	45.728	0.000
Variances				
PAR	1.000	0.000	999.000	999.000
SELF	1.000	0.000	999.000	999.000
OUGHT	1.000	0.000	999.000	999.000
EX	1.000	0.000	999.000	999.000
MB	1.000	0.000	999.000	999.000
INSTR	1.000	0.000	999.000	999.000
IDEAL	1.000	0.000	999.000	999.000
Residual Variances				
11	0.502	0.024	21.266	0.000
12	0.367	0.021	17.494	0.000
13	0.449	0.023	19.264	0.000
14	0.428	0.023	18.205	0.000
PAR1	0.487	0.029	17.053	0.000
PAR2	0.439	0.024	18.440	0.000
PAR3	0.289	0.021	13.517	0.000
PAR4	0.485	0.025	19.716	0.000
SELF1	0.481	0.024	19.957	0.000
SELF2	0.330	0.021	15.648	0.000
SELF3	0.370	0.022	16.960	0.000
SELF4	0.392	0.024	16.472	0.000

B E B 1 7 B = Z V 2 8 T Z Q U <</td> 2 5 3 5 % A 3 F W ? 9 B E B 1 7 B 9 8 1 % > ? 9 B E B 1 7 ? 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 8 7 3 Q J H Z V 2 8 T I Q U <</td> 2 5 3 5 % A X G % L 1 U M 8 = ? 9 B E B W 8 F W ? 8 F W ? 8 F W ? 8 F</t

	Estimate	S.E	Est./S.E	Two-Tailed P-Value
Residual Variances				
01	0.527	0.027	19.472	0.000
02	0.580	0.030	19.527	0.000
03	0.548	0.028	19.677	0.000
04	0.328	0.025	13.240	0.000
EX1	0.456	0.024	19.136	0.000
EX3	0.387	0.023	16.616	0.000
EX4	0.296	0.017	16.945	0.000
MB1	0.349	0.022	15.780	0.000
MB2	0.470	0.022	21.145	0.000
MB3	0.201	0.018	10.980	0.000
INSTR1	0.509	0.030	17.065	0.000
INSTR2	0.388	0.025	15.744	0.000
INSTR3	0.391	0.023	17.355	0.000

© British Council 2017