



A :: FNOAG//XCKPW?9BE//17B=6>?93EB17B=ZV28TZQU-2535%AXG
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?9W
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
G%2
FWB
83Q
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?9W
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2
A::FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU-<2535%AXG
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
83QJHY4HS83Q%AXGX981^9#98>T//GJB::83QJHY4HS83Q%AXGX98
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
C::FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU-<2535%AXG
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?9W
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
BWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?9W
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2
A::FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU-<2535%AXG
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2
83QJHY4HS83Q%AXGX981^9#98>T//GJB::83QJHY4HS83Q%AXGX98
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
C::FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU-<2535%AXG
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?9W
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
BWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?9W
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2
A::F**ENGLISH**K**IMPACT**//17B=**AN?EVALUATION2OF**ZQU-<2535%AXG
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
G%2AXGML1981^9#DRM<L%A**ENGLISH8LANGUAGE5CAPABILITY**P2
83QJHY4HS83Q%AXGX981^9#98>T//GJB::83QJHY4HS83Q%AXGX98
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
?//GWB::16B=6PW8FW?9>?//G**SRI1LANKA**W8FW?4>JHY4HS83QJHY4
C::FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU-<2535%AXG
G%2AXGML1981^9#DRM<L%A**FEBRUARYT2018**2535%AXG%L1UM8P2
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?9W
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
BWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?9W
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2
A::FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU-<2535%AXG
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
G%2AXGML1981^9#DRM<L%A**ELIZABETHISHEPHERDAAND**L1UM8P2
83QJHY4HS83Q%AXGX981^9#98>T//GJB::83QJHY4HS83Q%AXGX98
FWB::17B=6PW?9BEB17OAG**VICTORIA8AINS**WORTHBW8FW8FW?9B
?7GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
C::FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU-<2535%AXG
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?9W
?//GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
BWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?9W
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2
A::FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU-<2535%AXG
?>GWB::16B=6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
9%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2
83QJHY4HS83Q%AXGX981^9#98>T//GJB::83QJHY4HS83Q%AXGX98
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?9B
?//GWB::16**=**6PW8FW?9>?//GWB::17B=6PW8FW?4>JHY4HS83QJHY4
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8P2

FOREWORD	02
CONTRIBUTORS	05
1. EXECUTIVE SUMMARY	08
2. INTRODUCTION	10
3. RESEARCH METHODOLOGY	12
The sample design: a stratified two-stage cluster sample design	12
Method of delivery	16
The English language assessment	16
The student context questionnaire	19
Motivational scale analysis	24
4. IMPLEMENTATION IN SRI LANKA	28
Sri Lanka educational context	28
Teacher information	32
The comparison unit	33
5. LANGUAGE LEARNING ENVIRONMENT	34
6. SAMPLING RESULTS	40
7. ASSESSMENT OUTCOMES	44
Receptive skills	46
Productive skills	47
Comparing achievement by gender	48
Comparing achievement by urban and rural school location	49
Comparing achievement by Sri Lankan province	50
Comparing achievement by medium of instruction within schools	51
8. ENGLISH LANGUAGE LEARNING MOTIVATION	54
Construct validation findings	55
Multi-group analysis findings	60
Links between motivation and proficiency	67
Summary of English language learning motivation findings	74
9. IN CONCLUSION	76
REFERENCES	78
APPENDIX A – OVERVIEW OF THE APTIS FOR TEENS TEST STRUCTURE	83
APPENDIX B –FINAL CFA MEASUREMENT MODEL ESTIMATES	88

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535% A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535% A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW O 2 1 6 B E N G L I S H ? I M P A C T B :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWB S R I B L A N K A FW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H O 3 3 Q J H
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
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

LIST OF TABLES

Table 1: Level of localisation for Aptis for Teens (from O’Sullivan and Dunlea, 2015)	18
Table 2: Reference table of motivational scales and associated descriptions	21
Table 3: Types of schools and statistics, 2016	30
Table 4: Types of schools by province	30
Table 5: Schools by medium of instruction	31
Table 6: Students by medium of study and province	31
Table 7: Teacher categories by province	32
Table 8: Stratification variables chosen	33
Table 9: Demographic variables of participating students from Sri Lanka	34
Table 10: Target population definition in Sri Lanka	40
Table 11: Coverage and exclusions	40
Table 12: School sample size	41
Table 13: Student sample size	41
Table 14: Population and sample size	42
Table 15: Weighted school and student participation rates	42
Table 16: Distribution of student population for overall test performance by CEFR level	44
Table 17: Mean population score and CEFR levels by skills and overall achievement	44
Table 18: Estimated distribution of student population for listening and reading skills by CEFR level	46
Table 19: Estimated distribution of student population for speaking and writing skills by CEFR level	47
Table 20: Mean performance score by gender, skills and overall achievement	48
Table 21: Mean performance score by school location, skills and overall achievement	49
Table 22: Comparing overall mean score performance with nine Sri Lankan provinces	50
Table 23: Mean performance score by medium of instruction	51
Table 24: Improvements in adjusted chi-square values following addition of covariances between error terms	56
Table 25: Error covariances included in the measurement model	58
Table 26: Standardised (STDYX standardisation) correlations between variables in the final model	59
Table 27: T-tests to check for differences between mean factor scores calculated from the gender model	61
Table 28: Latent covariances allowed to vary between gender groups (standardised estimates)	62

Table 29: Questions used in the construction of SES variable	63
Table 30: T-tests to check for differences between mean factor scores calculated from the gender model	64
Table 31: Significant differences between groups according to their proficiency	69
Table 32: Correlations between motivational variables and overall total score	70
Table 33: Correlations between motivational variables and scale scores for individual components of the proficiency test	72
Table 34: Correlations between overall proficiency and motivational variables, by gender	73
Table 35: Correlations between overall proficiency and motivational variables, by SES	74

LIST OF FIGURES

Figure 1: The socio-cognitive model for test design and validation	17
Figure 2: The grade participating students from Sri Lanka started learning English	35
Figure 3: Time spent learning English per week at schoo	36
Figure 4: Activity undertaken when studying English outside of school	36
Figure 5: Time spent learning English outside of school	37
Figure 6: Future pathway when finished compulsory schooling	37
Figure 7: Language most regularly used to watch, read or listen to different types of media	38
Figure 8: Activity most often done on computer, tablet or smart phone	39
Figure 9: Responses to the questionnaire items	54
Figure 10: Measurement model	57
Figure 11: Mean factor scores on motivational variables for higher and lower SES groups	66
Figure 12: Mean factor scores on motivational variables for SES percentile groupings	67
Figure 13: Motivational profiles for learners at different levels of proficiency	68
Figure 14: Correlation between overall proficiency and English self-concept	71
Figure 15: Correlation between overall proficiency and the ought-to L2 self	71

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
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 ?
? // G W O 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H O 5 3 Q J H
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 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 ? 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 8 F W ?

FOREWORD

Gill Caldicott

Country Director, British Council, Sri Lanka

The English Impact Survey, which evaluates English language capability at the secondary school level, makes a significant contribution to the current debate on the state of English proficiency in Sri Lanka. The Government is rightly concerned about the levels of English of students leaving the school system, whether to enter further and higher education or to join the workforce. This research has shown that the sampled students, aged just over 15 years, are performing in English at the relatively low levels of A1 and A2 CEFR, reinforcing the findings of the Government’s own research and analysis. In the medium to longer term, if not addressed, this could adversely affect Sri Lanka’s success in developing the economy, and its ability to respond to international trade and tourism opportunities.

Fortunately, the time is right and the opportunities for reform and improvement at a systemic level are forthcoming. To be effective, this should include: the development of the teaching of English, of curriculum, methodology and examining of English; providing opportunities for teachers to improve their own English as well as use relevant and up-to-date methodology; the engagement of Sri Lankan industry and commerce in determining the level and nature of language proficiency that is needed; and, not least, the ability to tap into the innate motivation that students have shown for learning English – a very positive outcome of the research.

Sri Lanka is not alone in facing these challenges. In the global context, issues like these are becoming increasingly important as English as a lingua franca continues to grow across the world. The British Council remains committed to supporting the growth of English as a medium for international communication and a route to better life opportunities, respecting always the local culture and traditions, and the role of local languages. Responding to our Memorandum of Understanding (MoU) with the Ministry of Education in Sri Lanka, we are working through the National Colleges of Education to set up a cadre of skilled master trainers who will work with teachers across the country to improve their pedagogical approaches. In addition, I am pleased to note both private and corporate funding for improving the language level and the methodology of English teachers. It is through such initiatives, and a sustained effort over the medium term on the part of the education community in Sri Lanka, that significant change can be achieved.

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 forward.

We were incredibly lucky to work with exceptional local British Council and Ministry teams in Sri Lanka 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 Sri Lanka. 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 Ministry to continue to conceive and pursue successful policies for many years to come.

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
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 ?
? // G W **O 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 **S R I B L A N K A** F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H **O 7** 3 Q J H
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 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 ? 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 8 F W ?

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 the 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 Ministry of Education of Sri Lanka 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 Sri Lanka’s experiences in the increasingly important field of English language teaching and learning.

We would like to acknowledge the invaluable contributions made by the British Council staff in Sri Lanka 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 Sri Lanka would not have been possible.

British Council, Sri Lanka for its implantation of English Impact:

- Hasna Dole, Assistant Manager
- Lesley Dick, Senior Consultant
- Martin Froggett, Training Consultant.

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.

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	

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW 0816BENGLISH ? IMPACTB :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWBS RIBLANKAFW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H093QJH
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

1. EXECUTIVE SUMMARY

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

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

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

This research is underpinned by the British Council’s Royal Charter and charitable objective to develop a wider knowledge of the English language and it 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, it was delivered/ implemented 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 Sri Lanka

- English Impact employs a two-stage cluster sample design used by other recognised large-scale international surveys, sampling schools at the first stage and students at the second stage.
- 150 government-funded schools and 1,734 students were sampled for English Impact 2016–17, with 148 schools and 1,437 students participating following exclusions, student withdrawal from school or absence.
- Students were sampled from Grade 11 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, with a mean age at the time of testing of at least 15 years and six months.
- 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 in 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 Tamil and Sinhalese, gathered opinions and information 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 Sri Lanka met the English Impact international participation standard of at least 85 per cent of sampled students in 85 per cent of sampled schools.
- Of the participating students, 49 per cent started learning English in pre-school or first grade.
- In their current grade, 52 per cent of participating students chose to study English at school, while it was a compulsory school subject for 41 per cent. More than half of participating students, 51 per cent, spent at least three hours a week studying English.
- Of the students who interact with the Internet and computer games, around 87 per cent reported this to be primarily using an English-language medium.
- Over half of the participating population, 58 per cent, achieved A1 CEFR level in the English language assessment, with almost 30 per cent achieving A2 level.
- The skill of listening achieved the highest mean scale score of 23.6. Almost 40 per cent of students achieved B1 level, while 48 per cent achieved at A2 level.
- Reading achieved the second highest mean scale score of 14.92. Just over 56 per cent of students achieved A2 level on the CEFR.
- Participating students achieved their lowest

performance scores for the productive skills. Speaking skills had the lowest mean scale score with 4.91, followed by writing skills, with 8.34.

- Female students performed better than male students across all skills.
- Female students reported a higher level of confidence in their ability to learn English and more motivation to learn English, than male students. However, positive relationships between proficiency and all motivational variables were more pronounced among male students than female students.
- There are clear effects of socio-economic status on language learning motivation, with students from more advantaged backgrounds reporting higher motivation in all examined areas than their peers from less advantaged families.



2. INTRODUCTION

English language learning now plays a significant role in many national and regional education systems, with 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 decades of experience in designing the processes to sample and implement large-scale research of this kind.

This global best practice in research, together with 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 implication, 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 involve functionings that are explained in relation to achievements, and capabilities, as people within societies possessing the opportunity to achieve (Saito, 2003).

Following on from Sen’s concept of capability as encapsulating functionings and capabilities, achievement and opportunity, 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.

The evaluation of English language capability, reflecting Sen’s capabilities approach, is therefore not limited to the measurement of English language proficiency as captured by a test. Other data were captured and combined to provide a broader 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 further insight into students’ assessment outcomes, and indicating the impact of the English language policy in Sri Lanka.

The research aims outlined and investigated were to:

- evaluate the English language capability of students studying at public schools in Sri Lanka
- compare the capability of students attending urban and rural schools, and schools across the nine Sri Lankan provinces
- 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.

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
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 ?
? // G W 1 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 1 3 3 Q J H
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 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 ? 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 8 F W ?

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 Sri Lanka. 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 target 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, two major sources of surveys considered, including: a) 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); and b) 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 Sri Lanka 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 Sri Lanka worked directly with the Ministry 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 the established international surveys described above. The use of these standards enabled precision around the major estimates of the research, namely English Language capability, in order 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 school-level 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 Sri Lanka 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 Sri Lanka from the comparison unit.

Standard 1.7 The English Impact Sri Lanka 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 Sri Lanka 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 years and six months, 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.

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535 % A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535 % A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981 % > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981 % > ? 9BEB 17 ?
? // GW 1 4 1 6 B E N G L I S H ? I M P A C T B :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWB S R I B L A N K A FW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H 1 5 3 Q J H
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
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

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’.

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 Sri Lankan students meeting this definition and studying at least 90 minutes of formal English learning per week, rather than for the entire student population of Sri Lanka.

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 to present this measurement in the form of standard errors and/or confidence intervals around survey estimates. This protocol will be followed in the 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, such as 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 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, meeting the predefined criteria.

To ensure comparability and maximum coverage of the eligible population, the standards for English Impact require that school and within-school 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.

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 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 large-scale surveys can be limited with respect to exploring aspects of educational provision that can vary within participating countries.

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 sub-national comparison units to participate within the research project.

METHOD OF DELIVERY

To carry out the assessments in every sampled school in Sri Lanka, a pioneering digital method of delivery was developed. The English language assessments and student questionnaires were all completed by students via a completely offline-enabled tablet. While other large-scale assessments, such as PISA and TIMSS, have made initial steps towards computer-based assessment, English Impact has pioneered the completion of a large-scale assessment using 100 per cent computer-based 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 gather 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 the British Council’s Aptis for Teens test assessing the 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 Sri Lanka, all five components were taken.

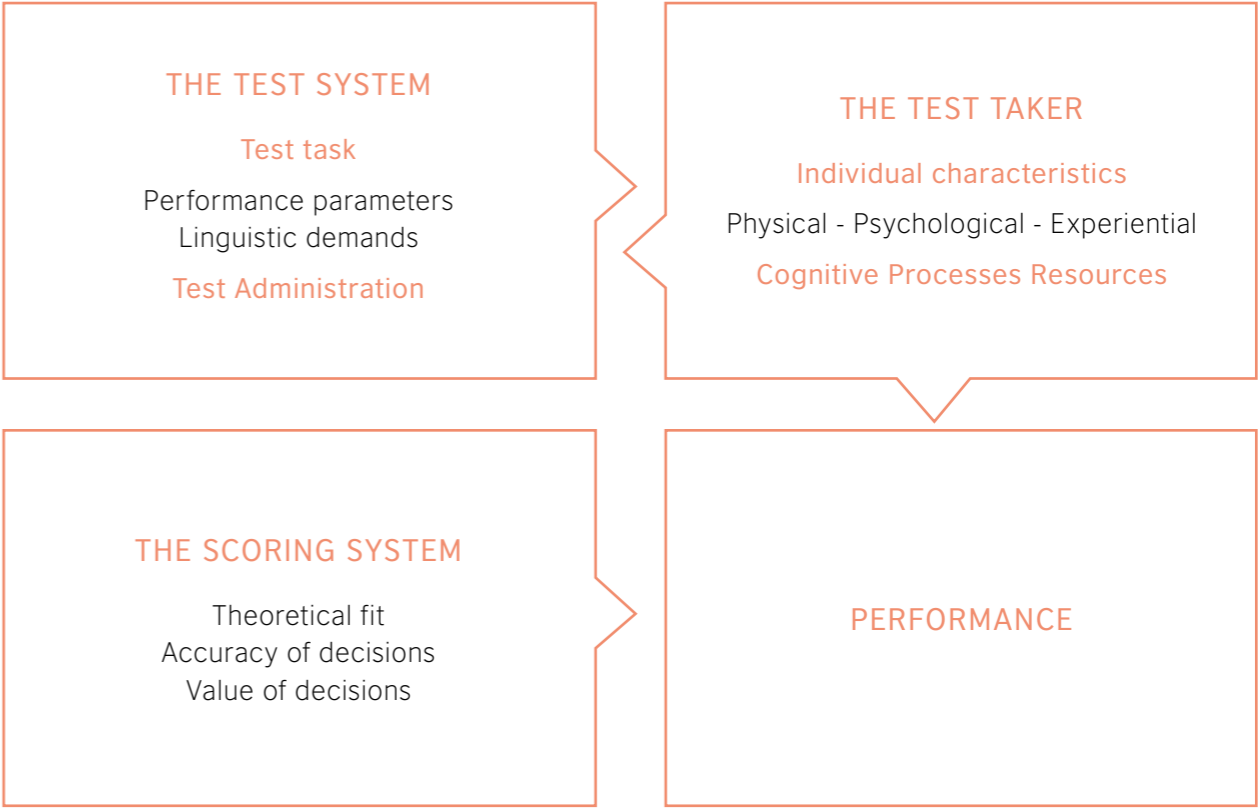
Theoretical model underpinning the test system

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 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.

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



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
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 ?
? // G W 1 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 1 9 3 Q J H
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 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 ? 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 8 F W ?

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 technical 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 opposite in Table 1.

Table 1: Level of localisation for Aptis for Teens (from O’Sullivan and Dunlea, 2015)

Level 4	Partial re-definition of target construct from existing variants. Will involve developing different task types to elicit different aspects of performance.	Developing new task types that are more relevant for a specific population of test takers, while remaining within the overall framework of 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 English language assessment comprising four skill areas plus a grammar and vocabulary component, students completed a background questionnaire to gather contextual information to support the English language capability data. The questionnaire was translated into both Tamil and Sinhala.

The student questionnaire comprises 53 items in three sections.

Demographic background, including socio-economic 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 socio-economic 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.

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
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 ?
? // 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 2 1 3 Q J H
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 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 ? 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 8 F W ?

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 second-language 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 quickly 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. 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 (IDEAL)	<i>Personal language goals</i>	Ideal L2 self is an image of oneself as a proficient 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 employed the Iwaniec (2014) scale as it encompasses the four skills of language learning.	I1	I imagine myself speaking English fluently.
			I2	I imagine myself comfortably reading in English on the Internet.
			I3	I imagine myself easily being able to follow what others say to me in English.
			I4	I imagine myself writing emails in English with ease.
Ought-to L2 self (OUGHT)	<i>Social expectations</i>	The ought-to L2 self is based on the external expectations placed on students and relates to the ‘attributes that one believes one ought to possess in order to avoid possible negative outcomes’ (Dörnyei, 2005, pages 105–106).	O1	I consider learning English important because the people I respect think that I should do it.
			O2	Studying English is important to me because other people will respect me more if I have knowledge of English.
			O3	Studying English is important to me because an educated person is supposed to be able to speak English.
			O4	Learning English is necessary because people surrounding me expect me to do so.
Language learning experience (EXPER)	<i>Interest in learning English</i>	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.
			EX2	I look forward to my English classes.
			EX3	I find learning English really interesting.
			EX4	I really enjoy learning English.

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535%AA FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B =ZV28TZQU – < 2535%ANOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? ? //GW 2216BENGLISH ? IMPACTB :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? //GWBSRIBLANKAFW ? 9 > ? //GWB :: 17B = 6PW8FW ? 4 > JHY4H233QJHG%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

Motivational scale (latent variable)	Descriptive name	Motivational scale descriptions	Question reference	Question details
Instrumentality (INSTR)	<i>Future opportunities</i>	Instrumentality represents motivation stemming from the practical benefits of language (Gardner and Lambert, 1972). This scale measures the perceptions of usefulness of English on job markets and future prospects.	INSTR1	I need English for my future career.
			INSTR2	The things I want to do in the future require me to use English.
			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 (INTER)	<i>Global communication</i>	International orientation is a construct recently developed in response to the 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.	INTOR1	Studying English will help me understand different people from other countries.
			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.
English self-concept (SELF)	<i>Self-confidence in English</i>	Self-concept is ‘a person’s perception of himself’ (Shavelson, Hubner and Stanton, 1976) and this scale relates to self-evaluation in the student’s ability to study English. The most common measurement of self-concept is Marsh’s (Marsh, 1990) Academic Self-Description Questionnaire, adapted to language learning by Iwaniec (2014).	SELF1	I usually get good marks in English.
			SELF2	Compared to other students, I’m good at English.
			SELF3	I have always done well in English.
			SELF4	Studying English comes easily to me.

Motivational scale (latent variable)	Descriptive name	Motivational scale descriptions	Question reference	Question details
Parental encouragement (PAREN)	<i>Family expectations</i>	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 behaviour (MOTIV)	<i>Level of effort</i>	Motivated learning behaviour attempts to measure the behavioural component of motivation, i.e. the reported amount of effort a student invests in English language learning.	MB1	I work hard at learning English.
			MB2	I think I’m doing my best to learn English.
			MB3	I put a lot of effort into learning English.
			MB4	I spend lots of time studying English.

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 gather 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 as respondents they 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 questionnaire 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; bilingual/non-bilingual school attendance; 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 in 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 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
- B. 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. As the comparison between levels is of key interest here, the scale is not essential.

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).

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV 28 T ZQU - < 2 5 3 5 % A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV 28 T ZQU - < 2 5 3 5 % A
NOAGWBN 7 = ? 9WNOAGWB :: E = ? 9WB 8FW ? 9BEB 17B 98 1 % > ? 9BEB 17 ? NOAGWBN 7 = ? 9WNOAGWB :: E = ? 9WB 8FW ? 9BEB 17B 98 1 % > ? 9BEB 17 ?
? // GW 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 2 7 3 Q J H
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 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 ? 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 8 F W ?

Under (B), the relationships examined are those between the latent variables, in the part of the model 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 as 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.

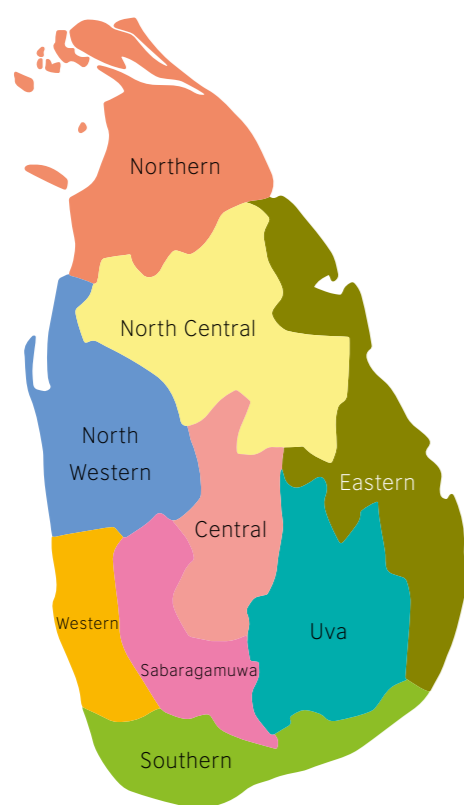
The F-scores, meanwhile, were derived from the multi-group CFA described above and ascribe each questionnaire respondent with a level for each motivational scale. Essentially, once the CFA model is set up satisfactorily, a value to reflect participants’ levels for each latent variable is calculated. 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>

4. IMPLEMENTATION IN SRI LANKA

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 Sri Lanka. In order to meet the established participation standards described in Chapter 3, a precisely defined comparison unit had to be outlined and agreed.



The provinces of Sri Lanka

SRI LANKA EDUCATIONAL CONTEXT

Sri Lanka has moved from a low-income to a middle-income country. Growth averaged 6.3 per cent in 2012–13 and reached 7.8 per cent in 2014. Per capita income rose from GDP USD\$859 in 2000 to USD\$3,256 in 2013. However, this prosperity is not evenly spread. While unemployment is low at 4.4 per cent nationally, it is at or above 20 per cent in the four conflict-affected districts. There is also significant under-employment, and youth unemployment is a thorny political issue, with 76 per cent of the unemployed below 29 years old. Other Human Development Indices are impressive by regional and lower middle-income standards. Sri Lanka has surpassed most Millennium Development Goals targets, outperforming its neighbouring countries on most of them. It has been stated that ‘improving the quality of human capital through effective education and skills development is central to Sri Lanka’s economic growth and competitiveness’. (<http://www.worldbank.org/en/country/srilanka/overview>)

Over the decade to 2015, Sri Lanka undertook a number of education initiatives. From 2012, the 1,000 Schools initiative led to the building of science, maths and language laboratories in 1,000 schools in underprivileged, mostly rural areas. A second initiative in 2013 saw the introduction of a fifth stream at A Level – the technological stream – as the beginning of an increase in technological/vocational education, with English, maths and science subjects receiving the most attention from the government as well as the public. A third as yet unimplemented, initiative under the previous government was the planned introduction of assessment of English speaking

and listening at O level, in an attempt to make education more closely aligned with the needs of the job market. Sri Lanka took part in the World Bank ‘Skills Toward Employment and Productivity’ survey, which revealed that about 18 per cent of its youth had either been on a training course of at least 30 hours in the previous 12 months or had completed an apprenticeship. In 2013, of the 20 per cent of young people enrolled in higher education, successful course completion rates were around 7 per cent (UNESCO GEM, 2016). During that time, the spent on education was 1.8 per cent of GDP, representing below 10 per cent of total public spend.

Since 2015, education has been identified as one of the five pillars for reform, with a commitment to increase spent on education to 6 per cent, although present spend is around 4 per cent. As yet, specifics are lacking and these are likely to take some time to emerge, with the Government’s early priorities being the UNHCR’s report on human rights in the civil war, and improving the economy (working with the IMF).

As Sri Lanka emerges from years of conflict, many recognise that education will be key to the country realising its full potential. Education is receiving a lot of media interest and the Sri Lankan Government and President are committed to improving the quality of education provision and demonstrating the impact of reforms. In 2017, the Ministry of Education - simultaneously introduced reforms in three areas: (1) the establishment of a school inspectorate; (2) the lengthening of the school year by two years to 13 years; and (3) the introduction of School Boards. There is also a keen interest in improving English language teacher education.

Several development partners, such as the ADB, World Bank, KOICA and JICA, have engaged in education to support the Government in achieving its plan of gradually increasing education expenditure to 6 per cent of GDP, in line with prioritising human capital development to ensure equity of the social market economy.

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535%AA FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535%AN
OAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW 3 0 1 6 B E N G L I S H ? I M P A C T B :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWB S R I B L A N K A FW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H 3 1 3 Q J H
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
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

The education system in Sri Lanka

The estimated size of the market in Sri Lanka is 4.1 million learners of English (MoE Department of English, 2013) and nearly 22,000 teachers of English spread across nine provinces, each with its own provincial, district and zonal structure. According to the School Census Preliminary

Reports of 2016 (Ministry of Education), Sri Lanka has a total of 10,162 schools and 4,143,330 pupils.

Schools in Sri Lanka are classified by type: 1AB, 1C, and Type 2 and Type 3 schools. The following tables explain the different types of schools and provide a breakdown for 2016.

Table 3: Types of schools and statistics, 2016

Type of school and definition	No. of schools	No. of pupils
1AB Schools with Advanced Level Science stream classes	1,016	1,626,565
1C Schools with Advanced Level Arts and/or Commerce streams but no Science stream	1,805	1,034,743
Type 2 Schools with classes only up to Grade 11	3,408	826,255
Type 3 Schools with classes only up to Grade 5	3,933	655,767

The spread of type of school across provinces is significant. The highest number of 1AB schools is in the Western Province. Of the total number of schools, 38.7 per cent have classes only up to Grade 5 or Grade 8.

Table 4: Types of schools by province

Province	1AB	1C	Type 2	Type 3	Total
Western	199	246	486	429	1,360
Central	115	316	467	623	1,521
Southern	150	208	365	391	1,114
Northern	104	109	302	456	971
Eastern	94	181	351	480	1,106
North Western	107	255	460	437	1,259
North Central	61	130	231	383	805
Uva	83	177	311	324	895
Sabaragamuwa	103	183	435	410	1,131
Total	1,016	1,805	3,408	3,933	10,162

The medium of instruction can be Sinhala, Tamil or English. The predominant medium of instruction is Sinhala. Some schools are bilingual, with Sinhala and Tamil streams, or Sinhala and

English streams, or Tamil and English streams. A minority have all three languages as medium of instruction. Table 5 shows all schools by medium of instruction.

Table 5: Schools by medium of instruction

Medium of instruction	No. of schools
Sinhala only	6,338
Tamil only	2,989
Sinhala and Tamil	66
Sinhala and English	554
Tamil and English	168
Sinhala, Tamil and English	47

Of the total number of students in Sri Lanka, two per cent are studying in the English medium, with about 30 per cent of them at schools in the Western Province.

Table 6: Students by medium of study and province

Province	Medium of study			Total
	Sinhala	Tamil	English	
Western	864,228	80,045	24,559	968,832
Central	350,796	180,825	12,587	544,208
Southern	502,745	14,784	8,741	526,270
Northern	5,127	234,781	3,406	243,314
Eastern	79,859	310,429	3,792	394,080
North Western	418,936	74,385	11,299	504,620
North Central	254,556	28,879	3,489	286,924
Uva	231,067	52,114	5,242	288,423
Sabaragamuwa	326,991	51,790	7,878	386,659
Total	3,034,305	1,028,032	80,993	4,143,330
Total by %	73.2	24.8	2.0	100.0

TEACHER INFORMATION

Sri Lanka has five different categories of teachers: graduate; trained; trainee; untrained; and other. Graduate teachers include graduates, postgraduate degree or diploma holders and graduate trainees. Trained teachers include those who have completed qualifications at the Teacher Training Colleges, National Colleges of Education and also those who have a two-year diploma in science and mathematics. Trainee teachers are those who are not yet absorbed

into the Sri Lanka Teacher Service, while untrained teachers and two-to-three year diploma holders are absorbed into the service. The final category, ‘other’, includes teachers paid by other government institutes and teachers paid by non-government institutes or organisations.

About 97.6 per cent of teachers in government schools are either graduates or trained teachers. Table 7 shows a breakdown by category of teacher and province.

Table 7: Teacher categories by province

Province	Graduate	Trained	Trainee	Untrained	Other	Total
Western	23,931	20,096	50	198	86	44,361
Central	12,841	19,489	496	615	28	33,469
Southern	12,791	15,922	109	185	20	29,027
Northern	6,183	9,661	384	159	172	16,559
Eastern	7,192	13,271	71	267	53	20,854
North Western	11,798	15,564	57	354	32	27,805
North Central	6,971	9,508	56	108	56	16,669
Uva	7,562	11,219	443	370	207	19,801
Sabaragamuwa	10,455	13,127	221	170	7	24,180
Total	99,724	12,7857	1,887	2,426	661	23,2555
Total by %	41.6	56	1	1.2	0.3	100

The comparison unit

- Sri Lankan government-funded schools.
- 150 schools and 1,734 students were sampled for participation.

As described above, implicit stratification variables were chosen to ensure a proportional sample allocation across the implicit strata. The stratification variables applied to the Sri Lanka sample frame are outlined in Table 8.

Table 8: Stratification variables chosen

Stratification variable name	Variable labels
Administrative province	Western/ Central/ North Western/ North Central/ Northern/ Sabaragamuwa/ Uva/ Eastern/ Southern
Location	Urban/ Rural
Medium of instruction	Sinhala (S) / Tamil (T) / Sinhala/Tamil(S/T) / Sinhala/English (S/E) / Tamil/English (T/E) / Sinhala/Tamil/English (S/T/E)

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
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 ?
? // G W 3 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 3 5 3 Q J H
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 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 ? 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 8 F W ?

5. LANGUAGE LEARNING ENVIRONMENT

From the 1,734 students sampled for participation, 1,437 students participated in English Impact Sri Lanka and completed a 53-item questionnaire translated into both Tamil

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

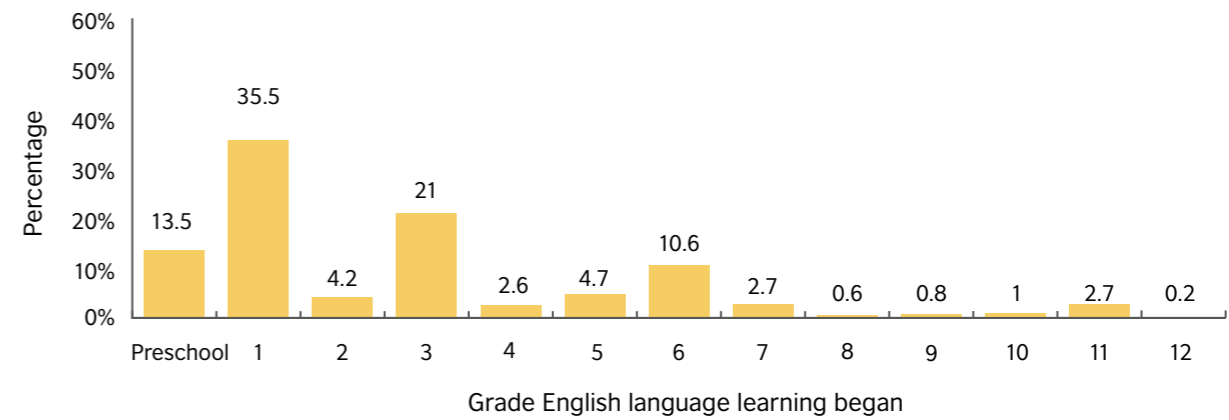
Table 9: Demographic variables of participating students from Sri Lanka

Demographic variables	Reported by participating students in Sri Lanka
Gender	54.4% female; 45.6% male
Age	Mean age of 15.7 years
Language most often spoken at home	73.1% Sinhala; 24.7% Tamil; 2% English; 0.2% other
Country of birth	99.1% Sri Lanka
Attendance at pre-school	93.3% attended pre-school; 6.7% didn't attend pre-school

The gender ratio of female and male students was a near equal split, 54.4 per cent and 45.6 per cent respectively. As described in Chapter 3, the target average age of the target population was 15 years and six months. The mean age reported by participating students was 15.7 years old, showing the accuracy of students sampled from the target population. The largest proportion, 73.1 per cent, indicated they spoke Sinhala most often at home. When asked to report their country of birth, over 99 per cent selected Sri Lanka. A small number of students, 2 per cent, reported they most often spoke English at home.

Participating students were asked to report when they started learning English: 13.5 per cent said pre-school; 35.5 per cent said first grade; 21 per cent said third grade; and over 10 per cent said they started learning English in sixth grade.

Figure 2: The grade participating students from Sri Lanka started learning English

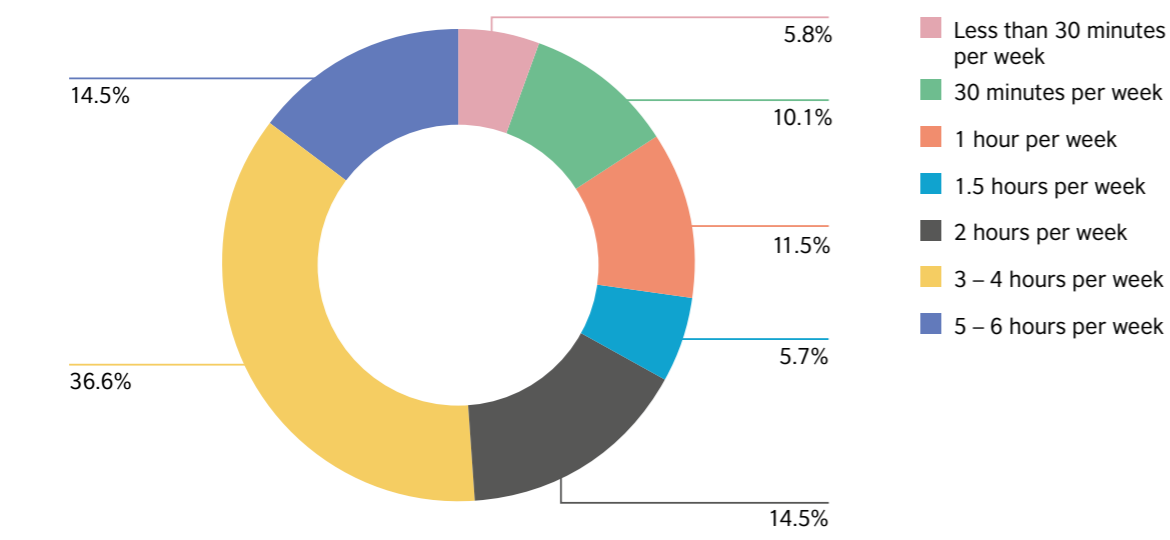


Students were asked how long they spend each week at school learning English, the largest number, 36.6 per cent, said they spent between

three and four hours per week studying English, as shown by Figure 3. A smaller group, 14.5 per cent, spent between 5 and 6 hours studying English.

A :: FNOAG // XCKPW ? 9 BE // 1 7 B = 6 > ? 9 BE B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A FNOAG // XCKPW ? 9 BE // 1 7 B = 6 > ? 9 BE B 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A
NOAG WBN 7 = ? 9 WNOAGWB :: E = ? 9 WB 8 FW ? 9 BE B 1 7 B 9 8 1 % > ? 9 BE B 1 7 ? NOAG WBN 7 = ? 9 WNOAGWB :: E = ? 9 WB 8 FW ? 9 BE B 1 7 B 9 8 1 % > ? 9 BE B 1 7 ?
? // GW 3 6 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 PW 8 FW ? 4 > J H Y 4 H S 8 3 Q J H ? // GW B S R I B L A N K A F W ? 9 > ? // GW B :: 1 7 B = 6 PW 8 FW ? 4 > J H Y 4 H 3 7 3 Q J H
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
FWB :: 1 7 B = 6 PW ? 9 BE B 1 7 O A G W B :: 7 = ? 9 = W 8 FW 8 FW ? 9 BE B W 8 FW 8 FW ? F W B :: 1 7 B = 6 PW ? 9 BE B 1 7 O A G W B :: 7 = ? 9 = W 8 FW 8 FW ? 9 BE B W 8 FW 8 FW ?

Figure 3: Time spent learning English per week at school

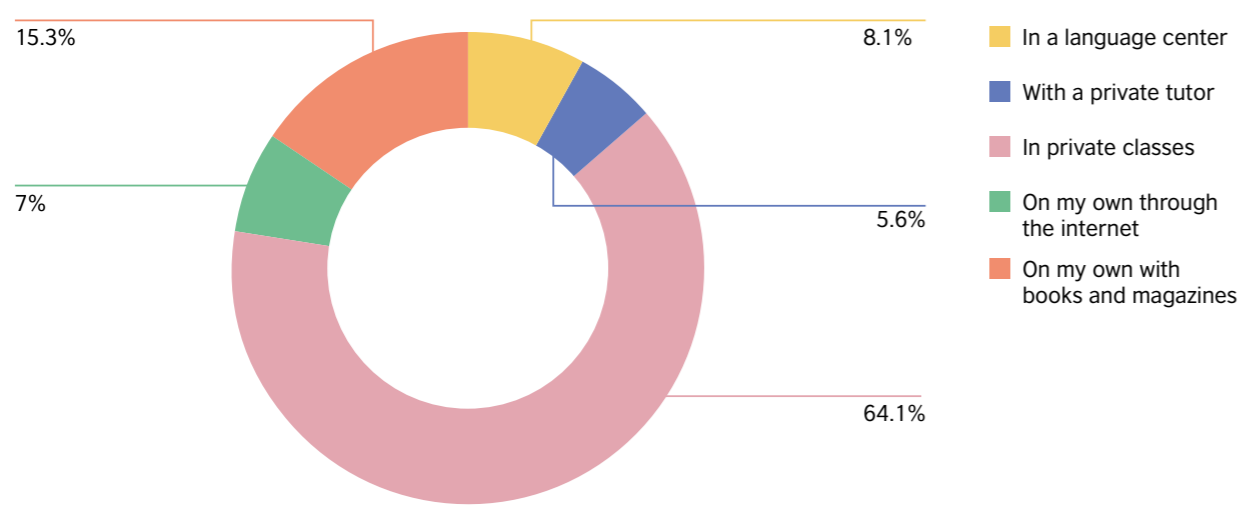


Participating students were also asked to provide information on studying English outside of school; a large proportion, 74.8 per cent, said they chose to study English outside of school.

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

described their extracurricular language learning: 64.1 per cent said they learnt in private classes; 15.3 per cent said on their own using books and magazines; 8.1 per cent said they learnt at an English language school; 7 per cent said on their own using the Internet; and 5.6 per cent with a one-to-one tutor.

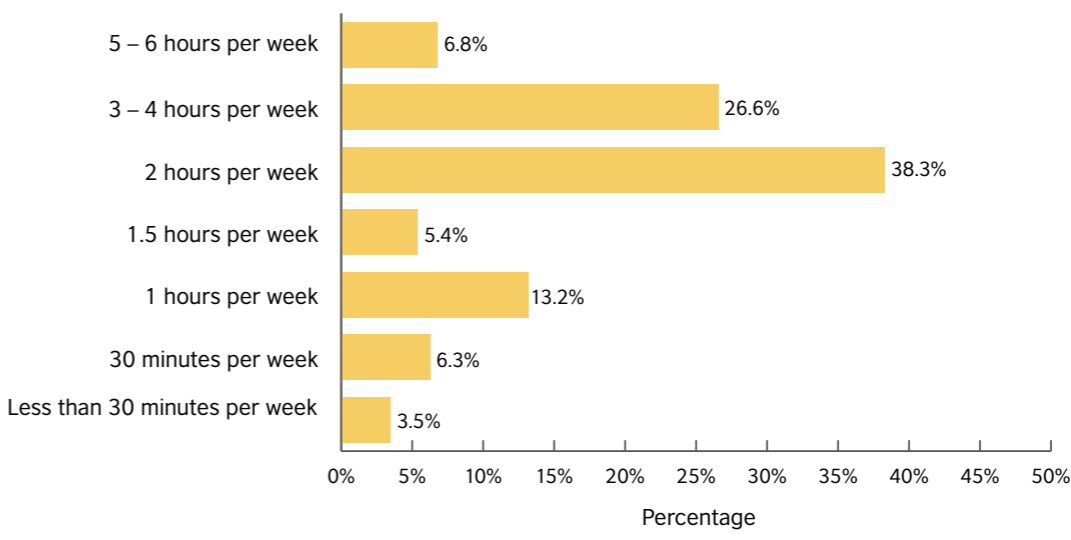
Figure 4: Activity undertaken when studying English outside of school



The same segment of the participating student population was asked how much time they spent learning English outside of school. On average, students spent between two and three to four

hours a week on English language learning, in addition to regular school hours. The largest proportion, 38.3 per cent, indicated they studied for two additional hours each week.

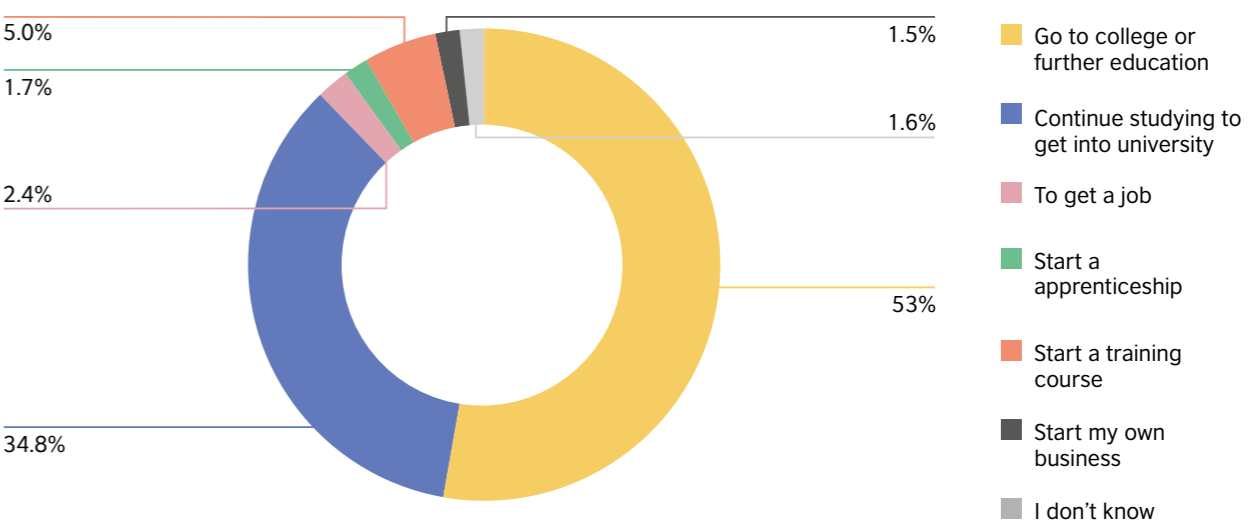
Figure 5: Time spent learning English outside of school



To understand more fully the possible further pathways of Sri Lankan school students in Grade 11, all participating students were asked what they would like to do when they finished school. Over half, 53 per cent, said they wanted to go to college or into further education; 34.8 per cent

wanted to continue studying to get into university; 5 per cent said they would like to start a training course; 2.4 per cent wanted to get a job; 1.5 per cent wished to start their own business; 1.7 per cent wanted to start an apprenticeship; and 1.6 per cent said they did not know.

Figure 6: Future pathway when finished compulsory schooling

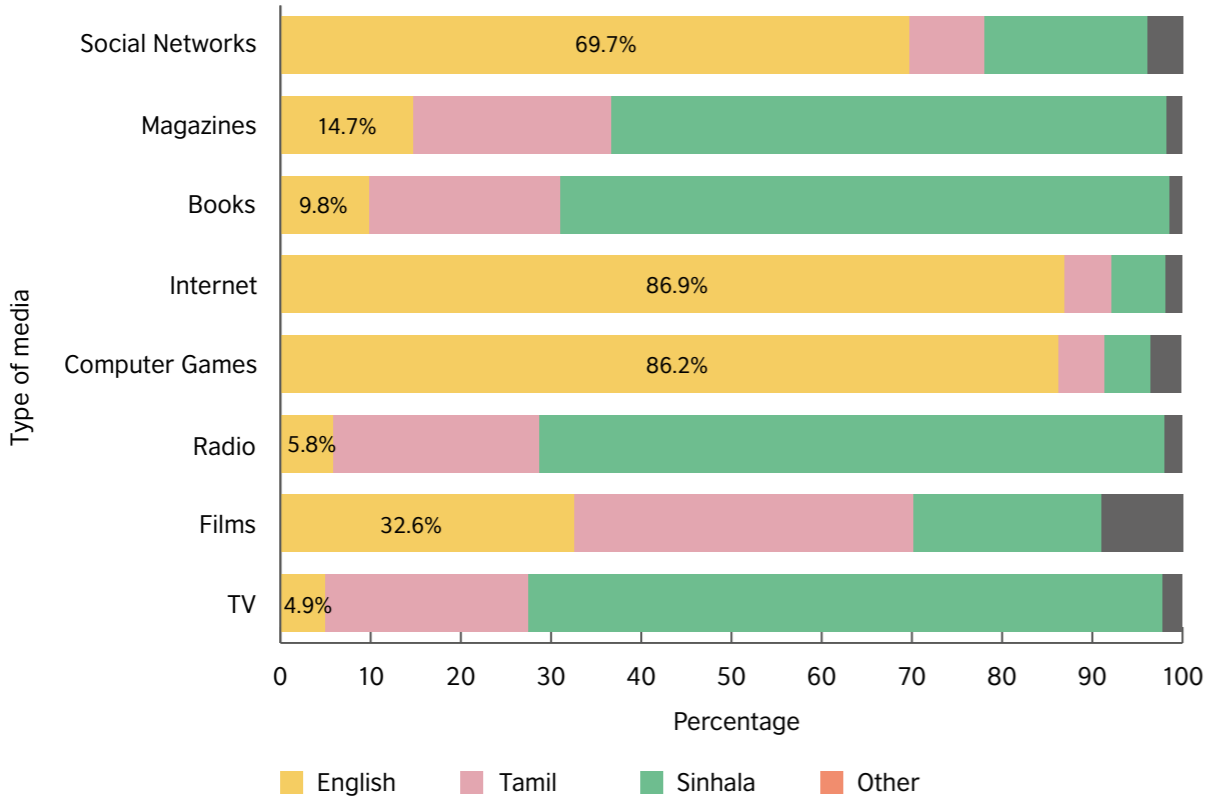


The impact of media usage on language learning is often identified anecdotally by teachers and policymakers as a catalyst for accelerated proficiency, especially among the digitally native generation of millennial learners that comprise the target population of English Impact Sri Lanka.

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, Tamil, Sinhala or another language.

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
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 ?
? // G W 3 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 3 9 3 Q J H
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 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 ? 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 8 F W ?

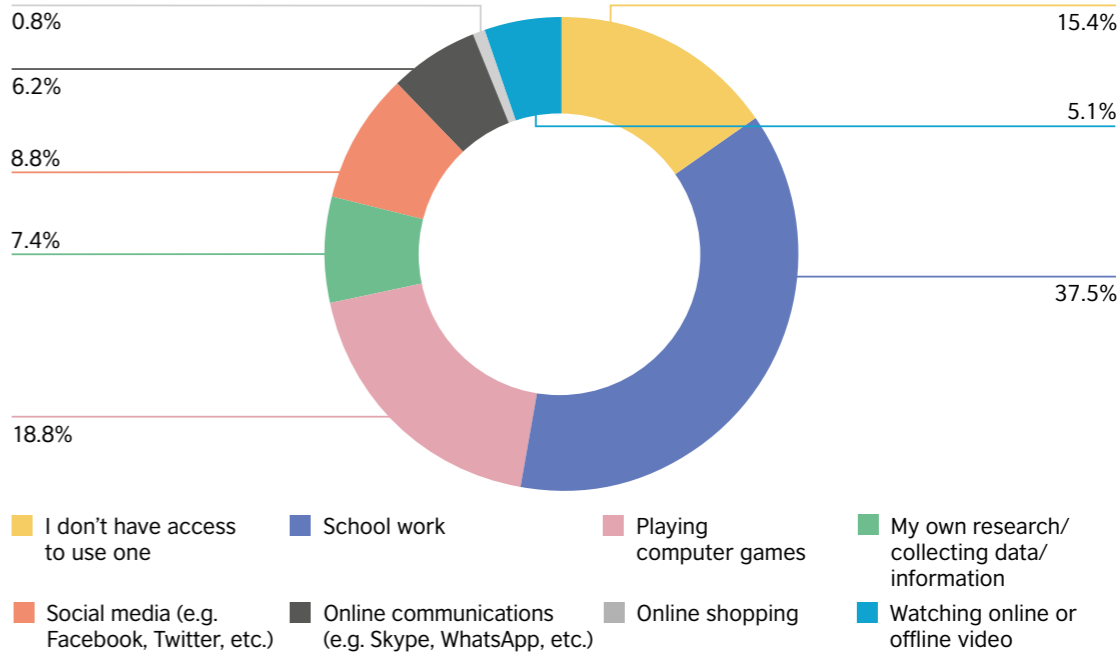
Figure 7: Language most regularly used to watch, read or listen to different types of media



Most notably 86.9 per cent and 86.2 per cent of participating students suggested they interact with the Internet and computer games, respectively, in English. This represents a significant proportion of the participating Sri Lankan population. In terms of social networks, 69.7 per cent indicated they interacted in English. Each of these aggregate responses highlights the relationship between use of technology and English language use for this demographic.

Participating students were also asked what they did most often when using a computer, tablet or smart phone: 37.5 per cent most often do their school work; and 18.8 per cent play computer games. However, 15.4 per cent said they did not have access to a computer, tablet or smart phone.

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



To more fully understand common patterns of usage of technology, students were asked how long they spend per week, on average, using a computer, tablet or smart phone.

Correlating with the proportion of students in Figure 8 who said they do not have access to a digital device, 17.1 per cent said they do not spend any time using them each week.

- 26 per cent spend between 30 minutes and 1 hour using their digital device
- 20 per cent spend 20 minutes maximum
- 19.4 per cent spend between 1 and 2 hours
- 9.9 per cent spend between 2 and 4 hours
- 7.6 per cent spend over 4 hours in total per week.

6. SAMPLING RESULTS

The sampling implementation process was carried out by a cross-organisational team of colleagues from ACER, representatives from the Sri Lankan Ministry of Education, and the British

Council in Sri Lanka and the UK. Focusing specifically on the comparison unit of Sri Lanka, all students fulfilling the target population criteria detailed in Chapter 3 are described in Table 10.

Table 10: Target population definition in Sri Lanka

Grade 10 equivalent in Sri Lanka	Years of formal schooling	Average age at time of testing	Information about age of entry, promotion and retention
Grade 11	11	15.7	<p>Students start at ISCED 1 in the year they turn six, unless their birthday is in January when they start in the previous year.</p> <p>There is no policy for promotion or retention, except in Grade 11 (target grade) when students who have not gained satisfactory O-Level scores will repeat the year and subsequent examination. There is no limit to how many times this may happen, although it is very rare for students to repeat more than twice.</p>

The target school population across the country was 100 per cent of all government-funded schools throughout Sri Lanka. No school level exclusions were made. A small number of students were classified as having a functional or intellectual disability and, therefore, exempt

from taking part in the assessment, as shown in Table 11. The overall rate of school level and within school exclusions was within the rate of 5 per cent outlined in participation standard 1.3. Therefore, a high standard of participation was successfully achieved.

Table 11: Coverage and exclusions

Coverage	Notes on coverage	School-level exclusions	Within-sample exclusions / refusals	Overall exclusions / refusals
100%	All schools in Sri Lanka's public education system	0%	2.7%	2.7%

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 Sri Lanka sample frame were administrative province, urban or rural location, and medium of instruction.

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 12. 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. Two substitute schools were used in Sri Lanka, due to non-response from two first sampled schools. There was no participation from two of the original sampled schools.

Table 12: 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	146	2	148

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 13 shows the total number of

students in all sampled schools, the total number 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 1,437.

Table 13: 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
83%	1,734	0	4	1,730	293	1,437

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 14. The table shows the actual number of schools and students in Sri Lanka 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.

A :: FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU- < 2535%A AFNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU- < 2535%
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?
?//GW4216BENGLISH?IMPACTB::17B=6PW8FW?4>JHY4HS83QJH?//GWBSRIBLANKAFW?9>?//GWB::17B=6PW8FW?4>JHY4H433QJH
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?

Table 14: Population and sample size

Sri Lanka population		Achieved sample population		
Schools	Students	Schools	Students	Student population size estimated from sample
5,791	285, 654	148	1,437	285,654

The English Impact Sri Lanka response rate is an important participation standard and indication of the successful implementation of the school and student sampling procedure. Table 15 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 15: Weighted school and student participation rates

School participation		Student participation	Overall participation	
Before replacement	After replacement		Before replacement	After replacement
97.3%	98.7%	86.1%	83.8%	84.9%

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 84.9 per cent. In Sri Lanka this standard was reached when rounded to zero decimal places.



7. ASSESSMENT OUTCOMES

The English language assessment taken by all participating students contained five components including grammar and vocabulary, testing four language learning skills: listening, reading, speaking and writing.

Over half of the participating population from Sri Lanka (58.2 per cent) achieved A1 CEFR Level in their overall English language test performance, while 29.2 per cent achieved A2 level, 6.9 per cent at B1 level, and 2 per cent at B2 level. A0 denotes those students who did not achieve the minimum threshold for an A1 level.

Table 16: Distribution of student population for overall test performance by CEFR level

CEFR Level	Overall English language test performance		
	Percentage (%)	Standard error (%)	95% CI
C	0.1	0.09	[0.0, 0.3]
B2	2.0	0.60	[0.8, 3.2]
B1	6.9	0.85	[5.2, 8.5]
A2	29.2	1.68	[25.9, 32.5]
A1	58.2	1.90	[54.5, 62.0]
A0	2.3	0.70	[0.9, 3.6]
Missing	1.3	0.30	[0.7, 1.9]
Total	100	-	-

As described in Chapter 3, the 95 per cent confidence interval is in the region of ± 1.96 Standard Error around the estimate and provides a measure of the certainty of the estimate.

As shown in Table 17 the standard errors of the mean estimates for all four skills are small, indicating a high level of precision for these estimates.

Table 17: Mean population score by skills and overall achievement

Overall score	Grammar vocabulary	Listening	Reading	Speaking	Writing	Overall
Mean	13.81	23.60	14.92	4.91	8.43	51.97
SE	0.308	0.385	0.333	0.428	0.441	1.453
95% CI	(13.20 – 14.41)	(22.85 – 24.36)	(14.27 – 15.58)	(4.07 – 5.75)	(7.56 – 9.29)	(49.12 – 54.82)



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
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 ?
? // G W 4 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 4 7 3 Q J H
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 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 ? 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 8 F W ?

Listening

The highest mean scale score achieved 23.6

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, 23.6, as shown in Table 17 above. With respect to CEFR levels, the estimated distribution of the listening skill performance

shown in Table 18 indicates that 39.6 per cent of participating students achieved B1 level, while a high proportion, 47.7 per cent, achieved A2 level for their listening skill.

The mean achievement for the Sri Lanka population in reading skills is 14.92, the second highest of the four components assessed. 56.1 per cent of participating students achieved at A2 level on the CEFR.

Table 18: Estimated distribution of student population for listening and reading skills by CEFR level

CEFR level	Listening component			Reading component		
	Percentage (%)	Standard error (%)	95% CI	Percentage (%)	Standard error (%)	95% CI
C	1.4	0.34	[0.8, 2.1]	1.2	0.36	[0.5, 1.9]
B2	3.2	0.73	[1.8, 4.7]	1.7	0.42	[0.9, 2.5]
B1	39.6	1.55	[36.5, 42.6]	8.7	0.85	[7.1, 10.4]
A2	47.7	1.67	[44.5, 51.0]	56.1	1.72	[52.7, 59.5]
A1	5.3	0.64	[4.0, 6.5]	25.8	1.22	[23.4, 28.2]
A0	2.5	0.82	[0.9, 4.1]	6.0	0.93	[4.1, 7.8]
Missing	0.3	0.14	[0.0, 0.6]	0.5	0.21	[0.1, 0.9]
Total	100			100		

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, 4.91, as shown in Table 17 above.

The distribution of speaking skills when referenced against the CEFR shows the largest proportion of students, 68.9 per cent, achieved at A0 level. See Table 19.

The third highest mean scale score was achieved for writing skills, 8.34. The estimated distribution of writing performance when referenced against the CEFR shows that 32.4 per cent of all students achieved at A1 level. The estimated distribution shows a proportion of students, 17.2 per cent, achieved above A1 level, at A2, B1 or B2 level, in writing skills.

Table 19: Estimated distribution of student population for speaking and writing skills by CEFR level

CEFR level	Speaking			Writing		
	Percentage (%)	Standard error (%)	95% CI	Percentage (%)	Standard error (%)	95% CI
C	0.0	0.00	[0.0, 0.0]	0.0	0.00	[0.0, 0.0]
B2	0.6	0.21	[0.2, 1.0]	1.1	0.33	[0.4, 1.7]
B1	5.5	1.03	[3.5, 7.5]	6.1	0.85	[4.5, 7.8]
A2	5.6	0.65	[4.3, 6.8]	10.0	1.11	[7.8, 12.2]
A1	18.7	1.34	[16.0, 21.3]	32.4	1.61	[29.2, 35.5]
A0	68.9	1.96	[65.1, 72.7]	50.0	1.96	[46.2, 53.9]
Missing	0.7	0.23	[0.3, 1.2]	0.4	0.17	[0.1, 0.7]
Total	100			100		

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW 4816BENGLISH ? IMPACTB :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWBSRIBLANKAFW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H493QJH
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

Female students achieved higher mean scale scores for all skills than males

Comparing achievement by gender

The ratio of participating female and male students was quite evenly distributed at 54.4 per cent and 45.6 per cent respectively.

As shown in Table 20, female students consistently achieved higher mean scale scores for all test components, skills and when calculating an overall score.

Table 20: Mean performance score by gender, skills and overall achievement

Gender	Grammar and vocabulary	Listening	Reading	Speaking	Writing	Overall
Female	14.31	24.07	15.43	5.51	9.26	54.40
Male	13.21	23.05	14.32	4.20	7.44	49.11
Difference (F-M)	1.10	1.02	1.11	1.32	1.82	5.29
Standard error	0.635	0.762	0.702	0.844	0.812	2.917
Comparison (95% confidence)						

Positive difference No difference Negative difference

Comparing achievement by urban and rural school location

One of the stratification variables selected for the Sri Lanka sample frame was an urban or

rural marker for each of the participating schools. Analysis of the mean performance of schools classified as urban or rural highlights the different levels of performance at both types of schools.

Table 21: Mean performance score by school location, skills and overall achievement

School Location	Grammar vocabulary	Listening	Reading	Speaking	Writing	Overall
Urban	14.89	26.38	17.10	6.30	11.11	60.88
Rural	13.48	22.76	14.26	4.48	7.61	49.24
Difference (U-R)	1.41	3.63	2.84	1.82	3.50	11.64
Standard error	0.789	0.998	1.022	1.271	1.255	4.344
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 listening, reading and writing skills, and overall scores.

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535% A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535% A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW 5 0 1 6 B E N G L I S H ? I M P A C T B :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWB S R I B L A N K A FW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H 5 1 3 Q J H
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
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

Comparing achievement by
Sri Lankan province

An additional implicit stratification variable used for the Sri Lankan comparison unit was province. The nine provinces are listed below. The overall mean scale score, as shown in Table 22, is 51.97 per cent. The table shows the comparative overall mean scale score performance for each province compared to the overall mean. The Western province had the highest overall mean scale score, followed by the Southern province. The Eastern and Sabaragamuwa provinces were equal to the overall mean, while the other five provinces were below, with the Northern Province the lowest.

Table 22: Comparing overall mean score performance with nine Sri Lankan provinces

Province	Overall mean performance	SE	In comparison to overall mean score
Central	44.3	(5.11)	▼
Eastern	51.0	(4.05)	►
North Central	47.0	(2.24)	▼
North Western	49.8	(2.82)	▼
Northern	38.2	(2.98)	▼
Sabaragamuwa	51.3	(1.81)	►
Southern	59.5	(3.94)	▲
Uva	47.0	(5.07)	▼
Western	61.0	(3.99)	▲

Below ▼ No difference ► Above ▲

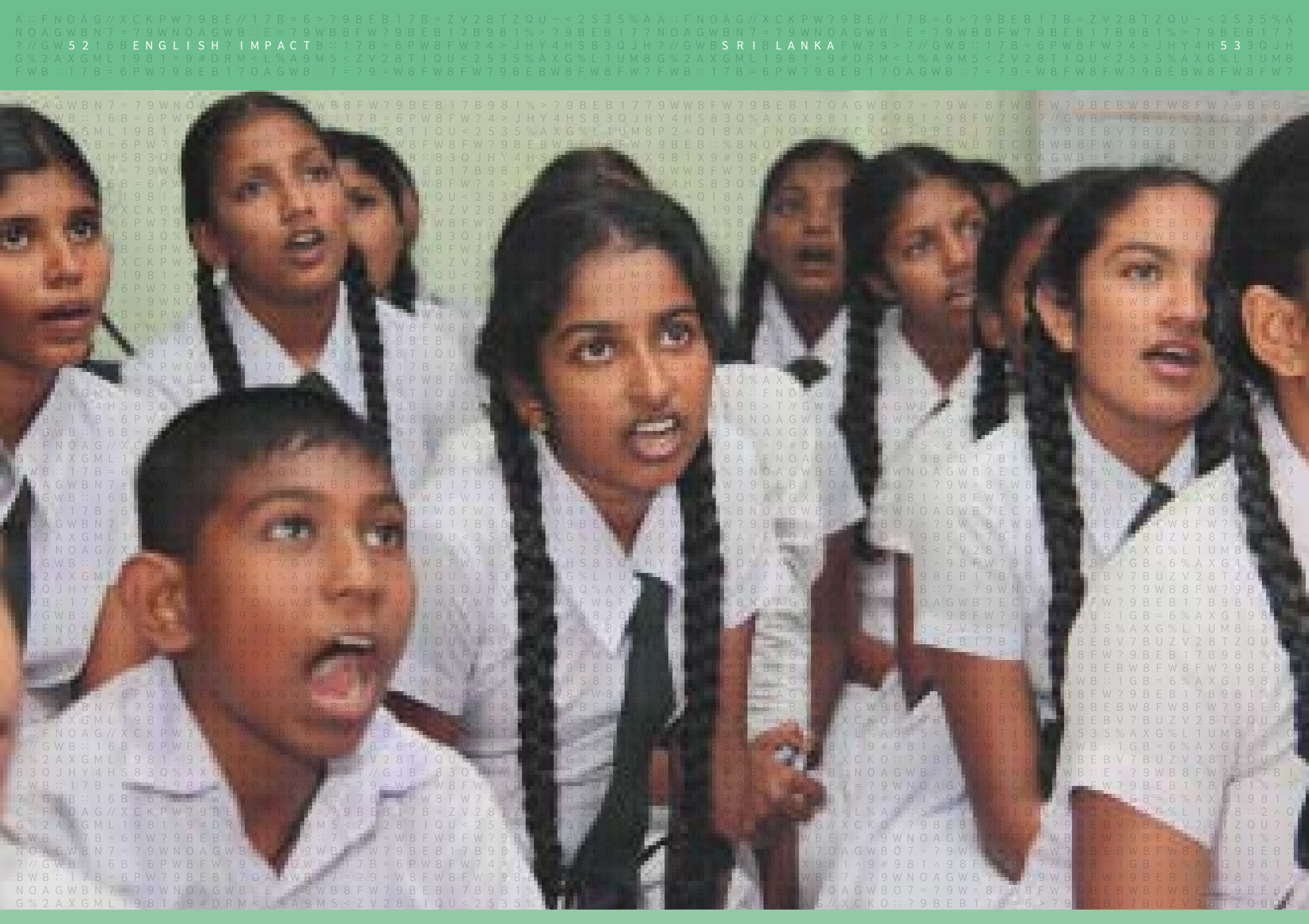
Comparing achievement by medium of instruction within schools

The final stratification variable for the Sri Lankan sample frame was medium of instruction. As shown in Table 23, the mean scale score achieved by students in bilingual and trilingual schools was significantly above the overall mean scale score, of 51.97 per cent.

Table 23: Mean performance score by medium of instruction

Medium of instruction	Mean	SE	In comparison to overall mean score
Sinhalese	40.2	(1.17)	▼
Sinhalese/English	68.1	(3.34)	▲
Sinhalese/Tamil/English	80.1	(26.33)	▲
Tamil	41.8	(2.56)	▼
Tamil/English	61.4	(6.24)	▲

Below ▼ Above ▲



A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW 5 4 1 6 B E N G L I S H ? I M P A C T B :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWB S R I B L A N K A FW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H 5 5 3 Q J H
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
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

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.

The student questionnaire received a full response rate, owing to the fact that students could not proceed through the app without

answering all questions. As can be seen from Figure 9, the responses were overwhelmingly positive. In the case of a number of items, there appears to be a ceiling effect, with the largest number of students opting for the higher response (see q19_15c). In other cases, the range of responses was more varied (see q22_11d). However, as a majority of answers were positive, the variation in the data is limited, which means that the observed effects are smaller.

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

CONSTRUCT VALIDATION FINDINGS

Measurement model

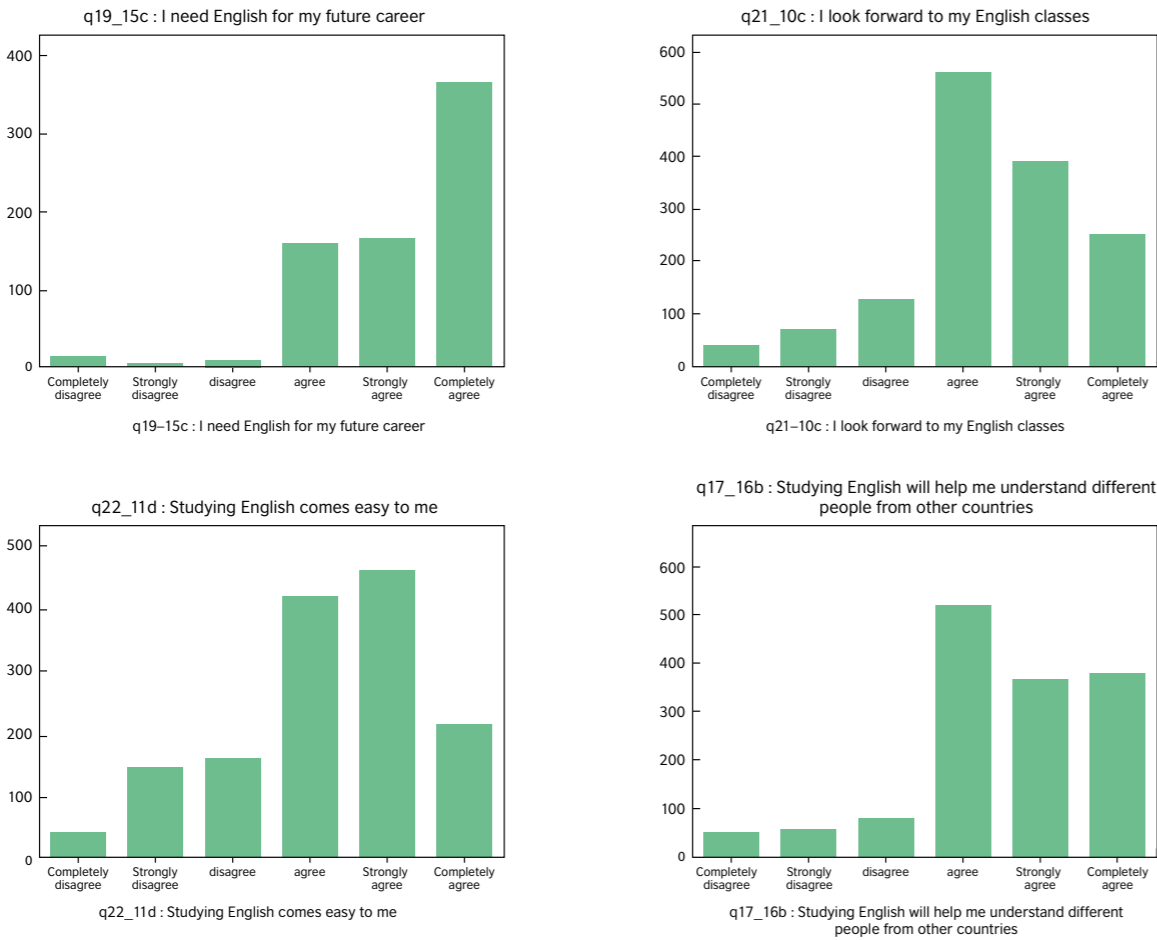
As described in Chapter 3, the aim of the construct validation exercise was to establish a measurement model that satisfactorily reflects the pattern of observed responses. The value of a well-fitting measurement model is that it provides a statistically sound means of estimating levels of motivation 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 blue in Table 24. (Please cross-reference terms with Table 2.) The initial model showed this structure to have a borderline acceptable fit to the data according to accepted thresholds for CFA models (CFI = 0.907; TLI = 0.894; RMSEA = 0.042). In order to improve model fit, some items (instr4, o3, ex1, and mb4, indicated in Figure 10 by being crossed 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 languages of the survey, Tamil and Sinhala. Additionally, as the respondents could choose the language of the survey, they might have been answering slightly different versions of the same question. Removing these observed variables resulted in a

model with reasonable fit to the data according to accepted thresholds for CFA models (CFI = 0.946; TLI = 0.936; RMSEA = 0.035).

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 indicated in red on Figure 10⁴ and summarised in Table 24. 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 response⁵. The final measurement model, therefore, included two correlated error variances listed in Table 25 (these are also indicated by orange arrows in Figure 10). The improved model provides a good fit for the data (CFI = 0.955; TLI = 0.947; RMSEA = 0.032). Final model estimates are given in Appendix B.

Figure 9: Responses to the questionnaire items



4 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.

5 See Brown (2015, pages 157–162) for a fuller discussion of correlated measurement error and the implications in CFA.

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW 5 6 1 6 B E N G L I S H ? I M P A C T B :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWBS R I B L A N K A FW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H 5 7 3 Q J H
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

Table 24: Improvements in adjusted chi-square values following addition of covariances between error terms

Constrained model		Most constrained model	MB1 with MB3
Degrees of freedom	d0	322	321
Scaling correction factor	c0	1.385	1.385
MLM chi-square value	T0	882.021	830.669
More relaxed model		Add MB1 WITH MB3	Add PAR1 WITH PAR2
Degrees of freedom	d1	321	320
Scaling correction factor	c1	1.385	1.385
MLM chi-square value	T1	830.669	787.936
Calculations		Difference	Difference
Diff test scaling correction	cd	1.3854	1.609
Adjusted chi-sq value	Calc	71.143	59.7539
Adjusted chi-sq value	TRd	51.352	37.128

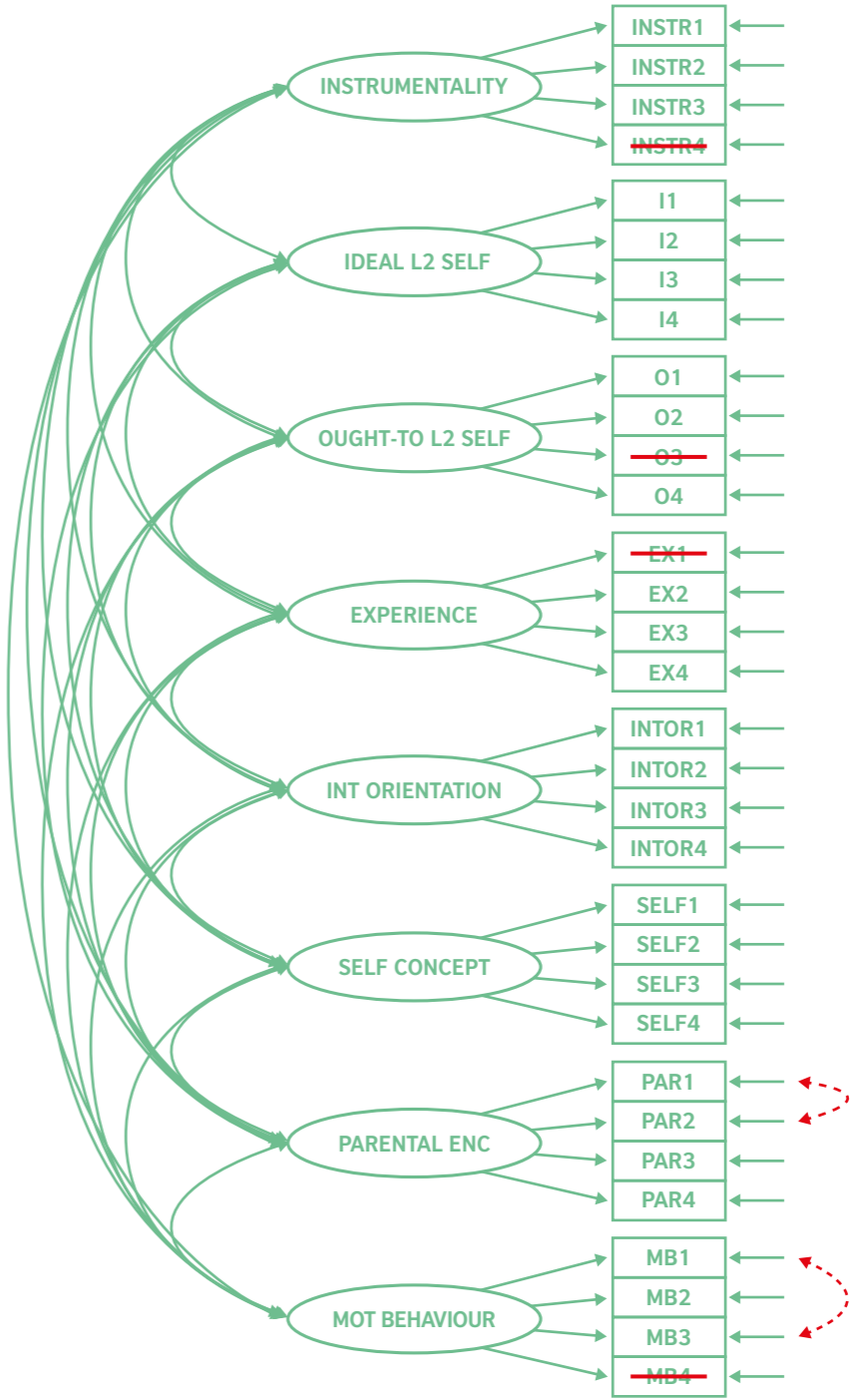
Figure 10: Measurement model

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

Boxes show the observed variables (question responses)

Circles show the latent variables (motivational scales)

Red double-headed arrows indicate correlated error terms



A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW 5 8 1 6 B E N G L I S H ? I M P A C T B :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWB S R I B L A N K A FW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H 5 9 3 Q J H
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
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

Table 25: Error covariances included in the measurement model

Covariance reference	Question reference	Question details	Reduction in adjusted chi-square value	Significance
1	MB1	I work hard at learning English.	51.352	p<.001
	MB3	I put a lot of effort into learning English.		
2	PAR1	My parents think I need to know English to be well educated.	37.128	p<.001
	PAR2	My parents have stressed the importance English will have for me in the future.		

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 26. 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 26: Standardised (STDYX standardisation) correlations between variables in the final model

Latent variables Estimate		Two-tailed			
		S.E.	Est. /S.E.	P-value	
EX	INTOR	0.645	0.031	21.118	0.000
IDEAL	EX	0.763	0.029	26.197	0.000
IDEAL	INTOR	0.714	0.027	26.273	0.000
INSTR	INTOR	0.819	0.023	36.395	0.000
INSTR	EX	0.638	0.030	21.577	0.000
INSTR	IDEAL	0.576	0.031	18.323	0.000
MB	EX	0.922	0.029	31.968	0.000
MB	INSTR	0.774	0.034	22.891	0.000
MB	IDEAL	0.734	0.033	22.450	0.000
MB	INTOR	0.708	0.036	19.792	0.000
OUGHT	INTOR	0.814	0.029	27.986	0.000
OUGHT	INSTR	0.799	0.033	24.320	0.000
OUGHT	MB	0.760	0.040	19.199	0.000
OUGHT	IDEAL	0.734	0.033	22.450	0.000
OUGHT	EX	0.649	0.035	18.515	0.000
PAR	OUGHT	0.853	0.029	29.261	0.000
PAR	INSTR	0.852	0.025	34.173	0.000
PAR	MB	0.791	0.033	24.176	0.000
PAR	INTOR	0.740	0.026	28.433	0.000
PAR	EX	0.665	0.030	22.199	0.000
PAR	IDEAL	0.640	0.030	21.645	0.000
SELF	EX	0.775	0.023	33.026	0.000
SELF	IDEAL	0.700	0.028	24.707	0.000
SELF	MB	0.612	0.033	18.490	0.000
SELF	OUGHT	0.486	0.035	13.849	0.000
SELF	INTOR	0.463	.031	14.897	0.000
SELF	INSTR	0.429	0.031	13.868	0.000
SELF	PAR	0.402	.031	12.886	0.000

Some of the strongest correlations are between the ought-to L2 self and the other motivational traits. This suggests that the Sri Lankan learners draw upon their understanding of social expectations to inform their levels of motivation.

The high correlations between parental expectations and the ought-to L2 self, and parental expectations and instrumentality, indicate that there is a consistency between what parents and what society expects of them.

A strong correlation can be observed between instrumentality and international orientation, suggesting that for Sri Lankan learners, the goals of learning English for communication with the wider world and for career opportunities are related.

In this questionnaire, the motivated behaviour scale aims to capture reported levels of effort invested in language learning. The high correlations suggest that all other motivational scales are clearly related to effort investment, with the scale of language learning experience being most strongly correlated with motivated behaviour. Similar findings have been previously reported in other studies (Csizér and Kormos, 2009; Iwaniec and Ullakonoja, 2016; Iwaniec, 2014). This means that positive language learning attitudes developed as a result of positive language learning experiences may lead to increased effort investment.

One latent variable in particular is a little less related to other areas of motivation, namely English self-concept. This construct captures learners’ confidence in their English ability. Hence it reflects the cognitive aspects of motivation, in contrast to other variables, which are affective or a blend of affective-cognitive in nature.

MULTI-GROUP ANALYSIS FINDINGS

The final sample included 1,439 learners. 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 783 female learners and 656 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 Iwaneic, 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 gender differences in language learning motivation have received very little attention outside of Europe, in South Asia particularly.

Levels of motivation

As can be seen in Table 27, significant differences across gender groups were found between the levels of motivation reported on four out of eight motivational scales. In all cases, female students reported higher levels of motivation than their male peers (significance level from *t*-test indicated in parenthesis):

- English self-concept (*p*<.001)
- language learning experience (*p*<.001)
- motivated behaviour (p=.002)
- instrumentality (*p*=.014).

Table 27: T-tests to check for differences between mean factor scores calculated from the gender model

Motivational variable	Group comparisons				t-test statistics		
	Gender	N	Mean	Std. deviation	t	Degrees of freedom	Sig. (2-tailed)
International orientation	female	783	.0083	.50300	.792	1437	.429
	male	656	-.0137	.54781			
Ideal L2 self	female	783	.0135	.57120	1.265	1437	.206
	male	656	-.0255	.59597			
Parental encouragement	female	783	.0243	.59177	1.667	1437	.096
	male	656	-.0300	.64215			
English self-concept	female	783	.0598	.66657	3.668	1437	.000**
	male	656	-.0725	.69891			
Language learning experience	female	783	.0591	.62913	3.606	1437	.000**
	male	656	-.0656	.68067			
Instrumentality	female	783	.0471	.71378	2.463	1437	.014*
	male	656	-.0495	.77305			
Motivated behaviour	female	783	.0573	.66419	3.170	1437	.002**
	male	656	-.0540	.66200			
Ought-to L2 self	female	783	.0167	.63064	1.093	1437	.274
	male	656	-.0213	.68476			

** p<.01 *p < .05

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

- have more confidence in their language learning ability
- have more positive language learning experiences, conceptualised as language learning attitudes
- invest more effort in language learning
- 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 future careers of female than male learners (Clark, 1998) – a 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 two covariances leads to a significant improvement in the adjusted model chi-square. This means that these relationships are significantly different between the groups of female and male students:

- ideal L2 self with language learning experience (7.505 on 1 d.f., p<.01)
 - instrumentality with motivated learning behaviour (4.786 on 1 *d.f.*, *p*<.05).
- This results in a model with good overall fit statistics (CFI= 0.947; TLI = 0.945; RMSEA = 0.033). A summary of the covariance estimates for this gender-specific model are given in Table 28.

Table 28: Latent covariances allowed varying between gender groups (standardised estimates)

Covariance		Estimate	S.E.	Est. /S.E.	Two tailed P-value
IDEAL WITH EX					
	Female model	0.702	0.037	18.923	.000
	Male model	0.839	0.037	22.691	.000
INSTR WITH MB					
	Female model	0.771	0.039	19.651	.000
	Male model	0.783	0.043	18.051	.000

As can be seen in Table 28, in all cases of covariances that were found to be significantly different, the strength of the covariance is higher for male than female learners.

The findings imply that:

- having a vision of oneself as a successful language learner is more closely related to positive language learning experiences for male than female learners
- having professional plans involving English is more closely related to effort investment for male than female learners.

Socio-economic status (SES)

The other group comparison is made 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 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.

This analysis is based on a measure of SES constructed from the information given by the students to the questions listed in Table 29⁶.

Table 29: 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?	<ul style="list-style-type: none">– 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?	<ul style="list-style-type: none">– 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?	<ul style="list-style-type: none">– [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]

⁶This 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 lower and higher SES. It needs to be noted that, 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.

6 The decision was taken not to use the information describing the actual job held by the parent in question (from Q49) in the analysis. It was 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.

Additionally, it should be noted that this approach means that no absolute measure of SES is employed. This is considered to be an advantage as the approach allows development of a measure that most accurately captures SES in a given context. 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 30, significant differences were found on all latent variables. 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):

- international orientation (*p*<.001)
- ideal L2 self (*p*<.001)
- English self-concept (*p*<.001)
- language learning experience (p<.001)
- motivated learning behaviour (p<.001)
- parental encouragement (*p*<.001)
- ought-to L2 self (*p*<.001)
- instrumentality (*p*<.001).

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

- be more strongly oriented to learn English for the purpose of communication with other speakers of this global language
- have more robust visions of themselves as successful English users in the future
- be more confident in their ability to study English
- have more positive language learning experiences, conceptualised as language learning attitudes
- invest more effort in language learning
- receive more encouragement and support from their parents to study English
- feel more pressure that they should know English
- be more likely to learn English with a view to using it to progress in their future professional career.

These uniform 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 also to invest more effort in language learning. They also collectively evaluate themselves as more capable of learning English and have more positive language learning attitudes. 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 that there might a number of 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.

The motivation profiles for each of the SES groups are shown in Figure 11.

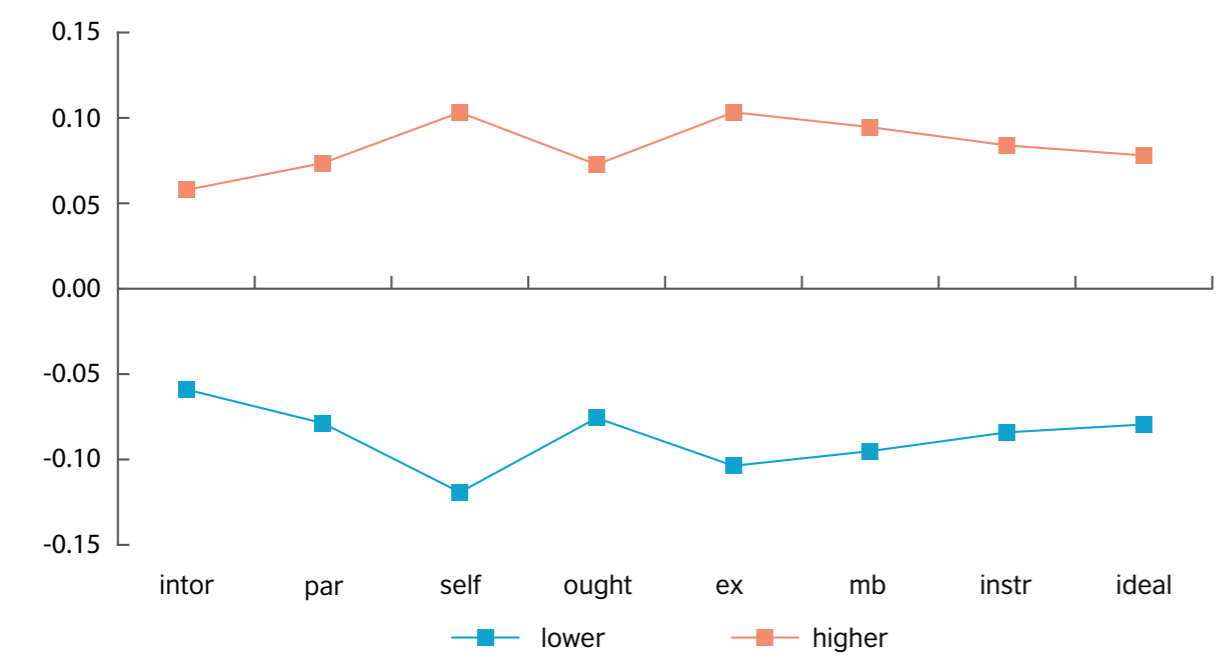
Table 30: T-tests to check for differences between mean factor scores calculated from the gender model

Motivational variable	Group comparisons				t-test statistics		
	SES	N	Mean	Std. deviation	t	Degrees of freedom	Sig. (2-tailed)
International orientation	lower	718	-.0590	.53190	-4.250	1437	.000**
	higher	721	.0578	.51012			
Ideal L2 self	lower	718	-.0795	.56354	-5.114	1437	.000**
	higher	721	.0780	.60337			
Parental encouragement	lower	718	-.0788	.63303	-4.729	1437	.000**
	higher	721	.0736	.58818			
English self-concept	lower	718	-.1192	.70200	-6.243	1437	.000**
	higher	721	.1032	.64857			
Language learning experience	lower	718	-.1037	.65531	-6.063	1437	.000**
	higher	721	.1033	.64026			
Instrumentality	lower	718	-.0841	.75251	-4.315	1437	.000**
	higher	721	.0839	.72356			
Motivated behaviour	lower	718	-.0951	.66716	-5.519	1437	.000**
	higher	721	.0946	.63640			
Ought-to L2 self	lower	718	-.0757	.64921	-4.319	1437	.000**
	higher	721	.0727	.65454			

** p<.01 *p < .05

A :: FNOAG // XCKPW ? 9BE // 1 7B = 6 > ? 9BEB 1 7B = ZV 2 8T ZQU - < 2 5 3 5 % A A FNOAG // XCKPW ? 9BE // 1 7B = 6 > ? 9BEB 1 7B = ZV 2 8T ZQU - < 2 5 3 5 % A
NOAGWBN 7 = ? 9WNOAGWB :: E = ? 9WB 8FW ? 9BEB 1 7B 9 8 1 % > ? 9BEB 1 7 ? NOAGWBN 7 = ? 9WNOAGWB :: E = ? 9WB 8FW ? 9BEB 1 7B 9 8 1 % > ? 9BEB 1 7 ?
? // GW 6 6 1 6 B E N G L I S H ? I M P A C T B :: 1 7B = 6 PW 8FW ? 4 > JHY 4HS 8 3 Q JH ? // GWB S R I B L A N K A FW ? 9 > ? // GWB :: 1 7B = 6 PW 8FW ? 4 > JHY 4 H 6 7 3 Q JH
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
FWB :: 1 7B = 6 PW ? 9BEB 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 7B = 6 PW ? 9BEB 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 ?

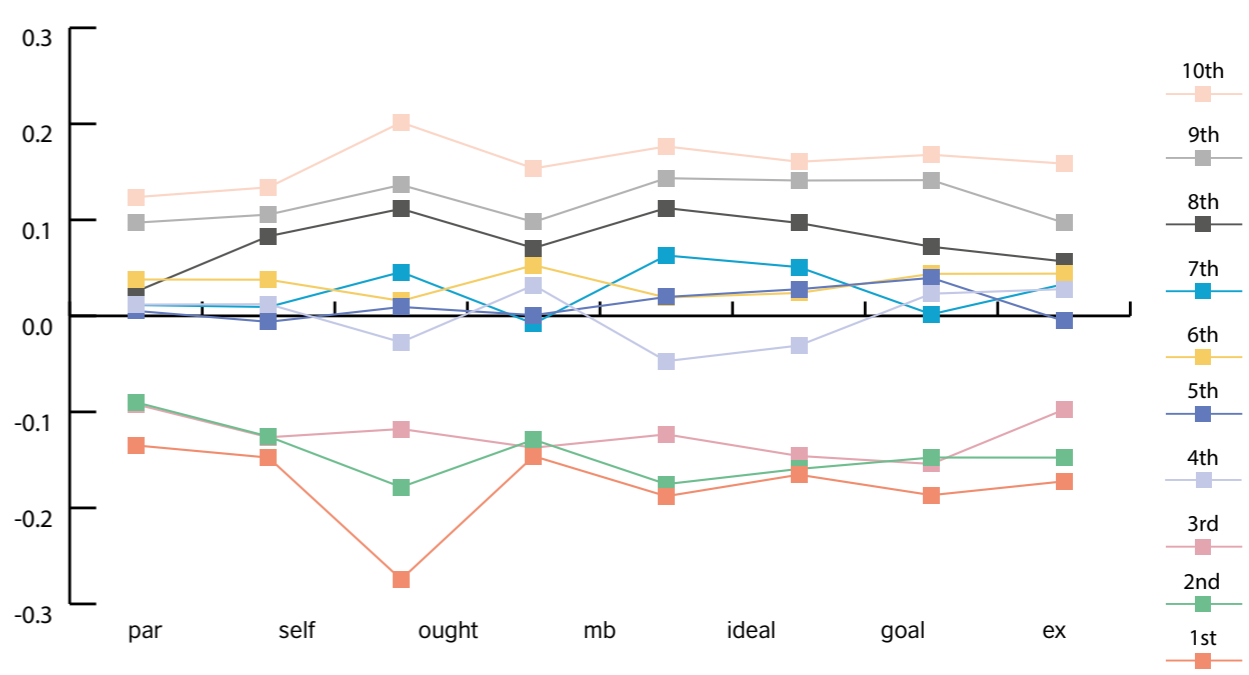
Figure 11: Mean factor scores on motivational variables for higher and lower SES groups



To examine the role of SES in greater depth, learners were divided into 10 groups, 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 12 show that the lines largely run in parallel, and that the biggest discrepancy is between students from the most advantaged and the least advantaged backgrounds. It can be seen that the motivational scores generally fall in line with students' SES status. The exception to this is students in the 7th percentile, who scored higher than students in the 9th and 8th percentiles. Students from the 10th, 9th and 7th percentile are most strongly motivated. The motivational profiles of students in 4th, 5th, 6th and 8th are similar, with those students generally reporting moderate motivation. The lowest motivational scores were reported by students in the 1st, 2nd and 3rd percentile who form a distinct group with

a similar motivational profile, the exception being English self-concept, which is considerably lower for those students who fall into the first percentile. Indeed, the gap is particularly large on the scale of self-concept, showing that students from the lowest SES group are much less confident in their ability to study and learn English, than those from the higher SES groups.

Figure 12: Mean factor scores on motivational variables for SES percentile groupings



Relationships between motivational scales

The results suggest that there are no significant differences between the higher and lower SES groups in terms of relationships between individual motivational scales. This means that, while the levels of motivation are affected by students' SES, the relationships between different motivational variables are not.

LINKS BETWEEN MOTIVATION AND PROFICIENCY

In this section, the relationship between Sri Lankan 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, and 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 Sri Lankan 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. Finally, there appears to be relatively little variation in the English proficiency of Sri Lankan 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.

The following presentation of results first gives an overview of the motivational profiles for students who achieve at different levels on the overall CEFR scale, and is followed by correlations between Aptis test scores and motivational scales for gender and SES groups. Implications of the findings will also be 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 CFA described above. The vast majority of learners achieved A1 (58.2%) or A2 (29.2%) levels, with only 9% of students scoring higher and 2.3% scoring lower. Scale scores were missing for 1.3 per cent (20 learners) of the sample. These learners are excluded from the analysis.

For this analysis, students were divided into four 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).

- 1. Low scorers: students in band A0 and those scoring 34 or less points classified in band A1; 438 learners (47.3 per cent female, 52.7 per cent male; 56.6 per cent lower SES, 43.4 per cent higher SES); 30.4 per cent of the total sample
- 2. Average scorers: students in band A1 with 35 and more points; 435 learners (52.9 per cent female, 47.1 per cent male; 54.3 per cent lower SES, 45.7 per cent higher SES); 30.2 per cent of the total sample
- 3. Above average scorers: students in band A2; 418 learners (60.5 per cent female, 39.5 per cent male; 46.2 per cent lower SES, 53.8 per cent higher SES); 29 per cent of the total sample
- 4. High scorers: students in bands higher than A2; 128 learners (64.1 per cent female, 35.9 per cent male; 25.8 per cent lower SES, 74.2 per cent higher SES); 8.9 per cent of the total sample.

Figure 13: Motivational profiles for learners at different levels of proficiency

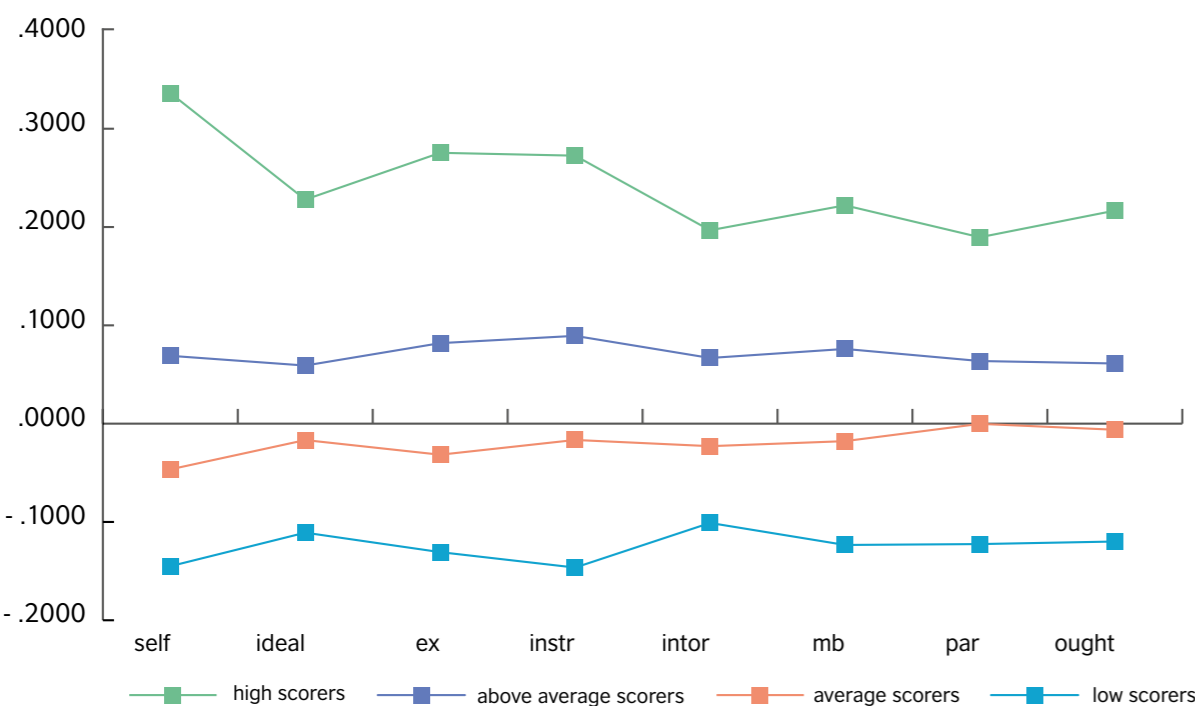


Figure 13 shows the relative patterns of mean factor scores for each of the motivational scales for the four groups. It can be seen that the factor scores for each group are consistently rising in tandem with their proficiency profiles.

Table 31: Significant differences between groups according to their proficiency

Proficiency	Proficiency	Motivational scale							
		English self-concept	Ideal L2 self	Language learning experience	Instrumentality	International orientation	Motivated behaviour	Parental encouragement	Ought-to L2 self
Low	Average								√
Low scorers	Above average	√	√	√	√	√	√	√	√
Low	High	√	√	√	√	√	√	√	√
Average	Above average								
Average	High	√	√	√	√	√	√	√	√
Above average	High	√	√	√	√				

Table 31 summarises the significant differences between learners at different proficiency levels. It can be seen that there are no significant differences between low scorers and average scorers, with the exception of the ought-to L2 self, or societal pressure. In contrast, low scorers report significantly lower levels of motivation than above average and high scorers on all aspects of motivation.

group reporting higher levels of confidence in their ability to learn English, more robust visions of themselves as future successful language users and more positive language learning attitudes. Also they report greater likelihood of pursuing English language learning as it might be useful in their future professional careers than the average learners.

Average and above average scorers do not differ significantly in any aspect of motivation examined. At the same time, average scorers reported significantly lower levels of motivation in all aspects than high scorers.

Correlations between proficiency and motivation
The results presented in Table 32 show the strength of correlations between the overall scale score⁷ and the factor scores for motivational scales. The correlations are listed in descending strength.

There are some significant differences between above average and high scorers, with the latter

7 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 dataset, total scale scores ranged between 0–175.

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW7016BENGLISH ? IMPACTB :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWBSRIBLANKAFW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H713QJH
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

Table 32: Correlations between motivational variables and overall total score

Motivational variables	Total score
English self-concept	.183**
International orientation	.162**
Language learning experience	.158**
Instrumentality	.157**
Ideal L2 self	.149**
Motivated behaviour	.135**
Parental encouragement	.135**
Ought-to L2 self	.133**

** p<.01 *p < .05

As can be seen, all motivational variables are positively related to proficiency, although the correlations are generally weak. The differences in the strength of correlations are small. This means that motivation can only explain variance in the proficiency scores to a certain extent. This is particularly visible when examining Figures 14 and 15, which provide a visualisation of two correlations of proficiency and the English self-concept and the ought-to L2 self. In both cases, most of the learners are positioned in the low right-hand corner. There are two

reasons for this. First, most students reported similar levels of motivation, as seen when examining responses to the ought-to L2 self-scale (Figure 15), whereas there is clearly more spread in the case of self-concept (Figure 14). Second, there was very little variance in the proficiency scores themselves as a vast majority of learners were classified as A1 and A2 learners. The results of the correlations suggest that, while motivation remains a factor that influences proficiency in Sri Lanka, other factors also play a substantial role.

Figure 14: Correlation between overall proficiency and English self-concept

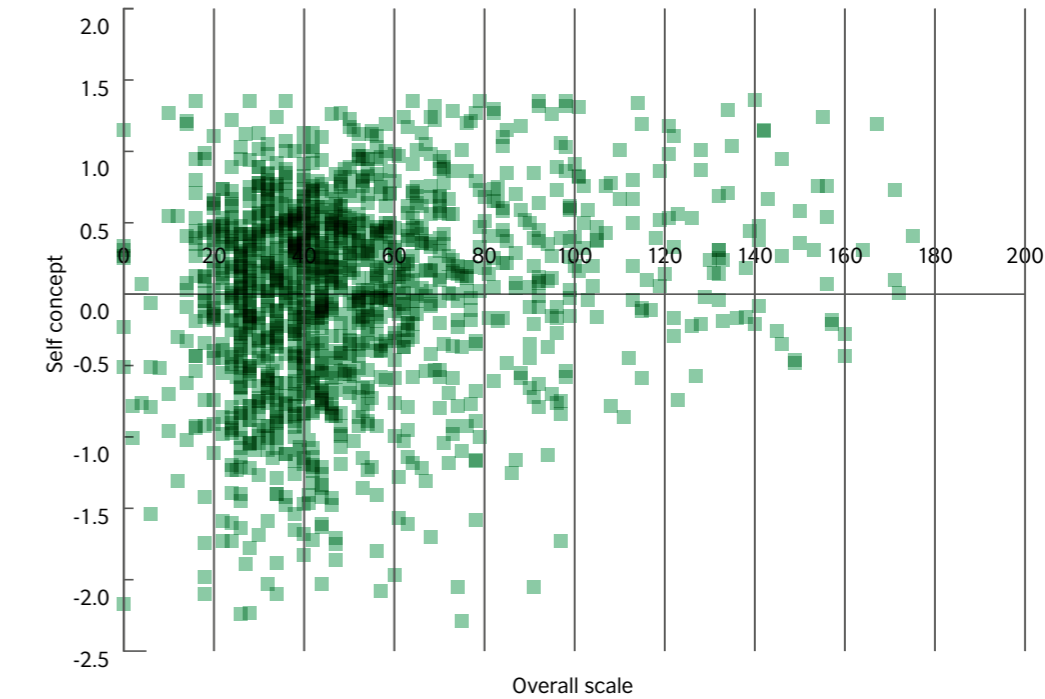
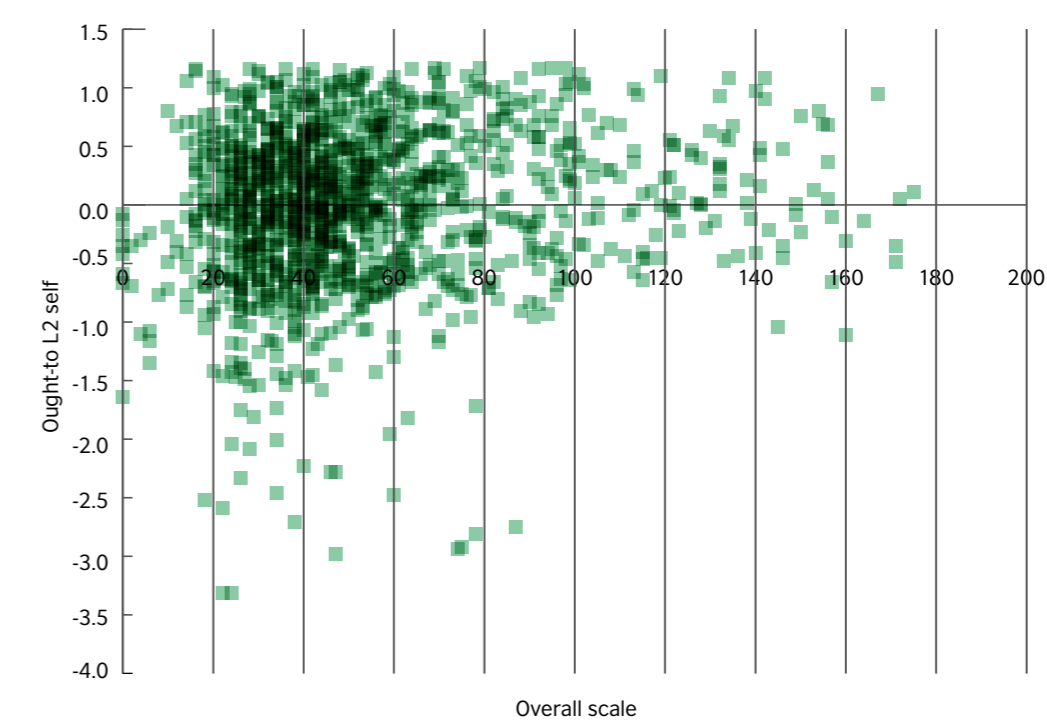


Figure 15: Correlation between overall proficiency and the ought-to L2 self



A :: FNOAG // XCKPW ? 9BE // 1 7 B = 6 > ? 9BEB 1 7 B = Z V 2 8 T Z Q U – < 2 5 3 5 % A A FNOAG // XCKPW ? 9BE // 1 7 B = 6 > ? 9BEB 1 7 B =Z V 2 8 T Z Q U – < 2 5 3 5 % A
NOAGWBN 7 = ? 9WNOAGWB :: E = ? 9WB 8 FW ? 9BEB 1 7 B 9 8 1 % > ? 9BEB 1 7 ? NOAGWBN 7 = ? 9WNOAGWB :: E = ? 9WB 8 FW ? 9BEB 1 7 B 9 8 1 % > ? 9BEB 1 7 ?
? // GW 7 2 1 6 B E N G L I S H ? I M P A C T B :: 1 7 B = 6 PW 8 FW ? 4 > JHY 4 HS 8 3 Q J H ? // GWB S R I B L A N K A FW ? 9 > ? // GWB :: 1 7 B = 6 PW 8 FW ? 4 > JHY 4 H 7 3 3 Q J H
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
FWB :: 1 7 B = 6 PW ? 9BEB 1 7 OAGWB :: 7 = ? 9 = W 8 FW 8 FW ? 9BEBW 8 FW 8 FW ? FWB :: 1 7 B = 6 PW ? 9BEB 1 7 OAGWB :: 7 = ? 9 = W 8 FW 8 FW ? 9BEBW 8 FW 8 FW ?

Table 33 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 listening and grammar and vocabulary tend to be strongest, whereas correlations with reading and speaking tend to be lowest. There is some small variation in the order of strength of correlations between different components and motivational variables.

Table 33: 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	.157**	.120**	.155**	.176**	.170**
International orientation	.158**	.122**	.122**	.133**	.160**
Language learning experience	.143**	.112**	.126**	.143**	.156**
Instrumentality	.159**	.129**	.114**	.123**	.158**
Ideal L2 self	.143**	.099**	.113**	.135**	.142**
Motivated behaviour	.131**	.100**	.101**	.113**	.139**
Parental encouragement	.144**	.106**	.095**	.105**	.132**
Ought-to L2 self	.136**	.099**	.101**	.106**	.134**

** 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 show that there are some

significant differences in the strength of relationships between motivational variables and the overall results on the proficiency test on three motivational scales: ideal L2 self, motivated behaviour, and the ought-to L2 self. In all cases, the correlations are stronger in the case of male students.

Table 34: Correlations between overall proficiency and motivational variables, by gender

Motivational scale	Female	Male	P-value [†]
English self-concept	.147	.214	.096
Ideal L2 self	.095	.209	.014*
Language learning experience	.123	.190	.098
Instrumentality	.113	.196	.056
International orientation	.136	.193	.136
Motivated behaviour	.072	.191	.012*
Parental encouragement	.100	.169	.093
Ought-to L2 self	.093	.180	.049*

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

These findings suggest that motivation is more strongly related to proficiency for male than female students. In detail, male students’ visions of themselves as successful users of English are more likely to lead them to higher proficiency gains, than they are for female students. Similarly, their reported effort investment and societal pressure to learn English lead to higher proficiency, than for female students.

This points to the conclusion that male students find it easier to realise their motivations. This is worth highlighting in a policy context, as it may reflect issues of resource allocation or cultural expectation.

Correlations by socio-economic status groups

The results of the correlational analysis of motivational variables and overall proficiency are presented in Table 35. The differences in the strength of correlation between motivation and proficiency are not significant. This suggests that, in terms of the relationship with proficiency, the same motivational processes are in place for members of both SES groups.

A::FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU- <2535%A AFNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU- <2535%
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?
?//GW7416BENGLISH?IMPACTB::17B=6PW8FW?4>JHY4HS83QJH?//GWBSRIBLANKAFW?9>?//GWB::17B=6PW8FW?4>JHY4H753QJH
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?

Table 35: Correlations between overall proficiency and motivational variables, by SES

Motivational scale	Lower SES	Higher SES	P-value†
English self-concept	.170	.156	.393
Ideal L2 self	.154	.107	.184
Language learning experience	.158	.121	.238
Instrumentality	.128	.153	.315
International orientation	.132	.156	.322
Motivated behaviour	.138	.092	.189
Parental encouragement	.131	.109	.337
Ought-to L2 self	.118	.116	.484

†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 something of the underlying patterns of motivation among teenage language learners. Female learners reported being more confident in their ability to learn English, have more positive language learning attitudes, invest more effort in language learning and are more motivated to learn English due to their future professional plans than male students. However, the positive relationship between proficiency and all motivational variables was more pronounced among the male than the female learners, suggesting that boys might have more opportunities to translate their motivation into effective learning than girls.

There are also clear effects of SES on language learning motivation, with students from more advantaged backgrounds consistently reporting higher motivation in all examined areas than their peers from less advantaged families.



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 throughout the government-funded school sector in Sri Lanka; to compare the outcomes in schools across Sri Lankan provinces; 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 Sri Lanka. 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 across Sri Lanka. Therefore, the consistent use of offline-enabled tablets to deliver all language assessments in every classroom has contributed to the success

of English Impact. 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.



A :: FNOAG // XCKPW ? 9BE // 1 7 B = 6 > ? 9BEB 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A A FNOAG // XCKPW ? 9BE // 1 7 B = 6 > ? 9BEB 1 7 B = Z V 2 8 T Z Q U - < 2 5 3 5 % A
NOAGWBN 7 = ? 9WNOAGWB :: E = ? 9WB 8 FW ? 9BEB 1 7 B 9 8 1 % > ? 9BEB 1 7 ? NOAGWBN 7 = ? 9WNOAGWB :: E = ? 9WB 8 FW ? 9BEB 1 7 B 9 8 1 % > ? 9BEB 1 7 ?
? // GW 7 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 7 9 3 Q J H
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 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 ? 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 8 F W ?

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, NE.: 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.

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.

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
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 ?
? // G W 8 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 8 1 3 Q J H
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 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 ? 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 8 F W ?

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, 117–138.

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.

Pfenniger, SE and Singleton, D (2016) Affect trumps age: A person-in-context relational view of age and motivation in SLA. *Second Language Research*, 32/3: 311–345. 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.

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>

UNICEF. (2010). Women and girls in Bangladesh. Available online at: https://www.unicef.org/bangladesh/Women_and_girls_in_Bangladesh.pdf

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.

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.

Test	Part	Skill focus	Items/ Part	Level	Tasks/ level	Items/ Task	Task focus	Task description	Response format
Core 50 items	1	Grammar	25	A1	5	1	Syntax and word usage	Sentence completion: select the best word to complete a sentence based on syntactic appropriacy.	3-option multiple choice
				A2	5-7	1			
				B1	5-7	1			
				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.
				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.
				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.	5 sentences, each with a 1-word gap. Select the best word to complete each from a bank of 10 options.
					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.				

Structure of the Teens Core component

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV 28TZQU – < 2535 % A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV 28TZQU – < 2535 % A
NOAGWBN 7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B 981 % > ? 9BEB 17 ? NOAGWBN 7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B 981 % > ? 9BEB 17 ?
? // GW 8 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 S R I B L A N K A F W ? 9 > ? // G W B :: 1 7 B = 6 P W 8 F W ? 4 > J H Y 4 H 8 5 3 Q J H
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 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 ? 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 8 F W ?

Test	Skill focus	Item/Part	Level	Task/level	Items/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

Test	Part	Skill focus	Items/Part	Level	Tasks/level	Items/Task	Task focus	Task description	Response format
Reading 25 items	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	A2	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.
	3	Text-level comprehension of short texts	7	B1	1	7	Text-level comprehension of short texts (careful global reading)	Candidates match 4 short paragraphs giving information about 4 people's opinions on different topics and identify which of the four people could say certain statements.	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 abstract texts.	7 paragraphs forming a long text. Select the most appropriate heading for each paragraph from a bank of 8 options.

Structure of the Teens Reading component

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535% A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU – < 2535% A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? //GW 8616BENGLISH ? IMPACTB :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? //GWBS R I B L A N K A FW ? 9 > ? //GW B :: 17B = 6PW8FW ? 4 > JHY4H 873QJH
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

Test	Part	Skill focus	Level	Task description	Channel of input / prompts	Time to plan	Time for response	Rating criteria
Speaking	1	Giving personal information	A1/A2	Candidate responds to 3 questions on personal topics. The candidate records his/her response before the next question is presented.	Questions presented in both written and oral form (pre-recorded). Questions presented in a sequence (e.g. Q2 is presented after the response to Q1).	No	30 seconds to respond to each question	Separate task-based holistic scales are used for each task. Performance descriptors describe the expected performance at each score band. The following aspects of performance are addressed: 1) grammatical range and accuracy 2) lexical range and accuracy 3) pronunciation 4) fluency 5) cohesion and coherence.
	2	Describing, expressing opinions, providing reasons and explanations	B1	The candidate responds to 3 questions. The first asks the candidate to describe a photograph. The next two are on a concrete and familiar topic related to the photo.	1) Questions presented in both written and oral form (pre-recorded). Questions presented in a sequence (e.g. Q2 is presented after the response to Q1). 2) A single photo of a scene related to the topic and familiar to A2/B1 candidates on screen.	No	45 seconds to respond to each question	
	3	Describing, comparing and contrasting, providing reasons and explanations	B1	The candidate responds to 2 questions / prompts and is asked to describe, contrast and compare two photographs on a topic familiar to B1 candidates. The candidate gives opinions, and provides reasons and explanations.	1) Questions presented in both written and oral form (pre-recorded). Questions presented in a sequence (e.g. Q2 is presented after the response to Q1). 2) Two photographs showing different aspects of a topic are presented on screen.	No	45 seconds to respond to each question	
	4	Integrating ideas on a topic into a long turn presentation. Giving and justifying opinions, advantages and disadvantages	B2	The candidate plans a longer turn presentation integrating information given to them and adding their own opinion/ knowledge of the subject.	The candidate is presented with a poster which they are told they have prepared and must present to their class.	90 seconds	2 minutes for the entire response	

Structure of the Teens speaking component

Test	Part	Skill focus	Level	Task description	Channel of input / prompts	Expected output	Rating criteria
Writing	1	Writing at the word level. Simple personal information on a form.	A1	The candidate completes a form by filling in some basic personal information. All responses are at the word/phrase level, such as name, birthdate, etc.	Form with 9 clearly marked categories (name, date of birth, etc.). There are 9 gaps in the form to be filled.	9 short gaps filled by 1–2 word responses	Separate task-based holistic scales are used for each task. Performance descriptors describe the expected performance at each score band. The following aspects of performance are addressed (not all aspects are assessed for each task): 1) task completion 2) grammatical range and accuracy 3) lexical range and accuracy 4) cohesion and coherence 5) punctuation and spelling.
	2	Short written description of concrete, personal information at the sentence level.	A2	The candidate continues filling in information on a form. The task setting and topic are related to the same purpose as the form used in part 1. The candidate must write a short response using sentence-level writing to provide personal information in response to a single written question.	Written. The rubric presents the context, followed by a short question asking for information from the candidate related to the context.	20–30 words	
	3	Interactive writing. Responding to a series of written questions with short paragraph-level responses.	B1	The candidate responds interactively to 3 separate questions. Each response requires a short paragraph-level response. The questions are presented as if the candidate is writing on an internet forum or social network site. The task setting and topic are related to the same purpose/ activity used in parts 1 and 2.	Written. The rubric presents the context (discussion forum, social media, etc.). Each question is displayed in a sequence following the completion of the response to the previous question.	30–40 words in response to each question	
	4	Continuous paragraph-level essay writing.	B2	The candidate writes an argumentative essay on a topical issue the candidate is likely to encounter in public or educational domains.	Written. The rubric presents the context in the form of an advert giving basic information about an essay competition.	220–250 words	

Structure of the Teens Writing component

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A A FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535% A
NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW 8816BENGLISH ? IMPACTB :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWBSRIBLANKAFW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H893QJH
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

APPENDIX B – FINAL CFA MEASUREMENT MODEL ESTIMATES

STANDARDISED MODEL RESULTS

STDYX Standardisation

	Estimate	S.E	Est./S.E	Two-Tailed P-Value
INTOR BY				
INTOR1	0.470	0.028	17.011	0.000
INTOR2	0.712	0.019	36.906	0.000
INTOR3	0.685	0.022	30.613	0.000
INTOR4	0.666	0.022	30.322	0.000
PAR BY				
PAR1	0.563	0.025	22.499	0.000
PAR2	0.547	0.026	20.749	0.000
PAR3	0.751	0.020	36.751	0.000
PAR4	0.734	0.020	36.278	0.000
SELF BY				
SELF1	0.591	0.024	24.735	0.000
SELF2	0.705	0.022	32.260	0.000
SELF3	0.768	0.017	46.299	0.000
SELF4	0.693	0.021	33.569	0.000
OUGHT BY				
O1	0.608	0.023	25.995	0.000
O2	0.611	0.027	22.902	0.000
O4	0.487	0.028	17.195	0.000
EX BY				
EX2	0.609	0.021	28.439	0.000
EX3	0.717	0.021	34.573	0.000
EX4	0.701	0.019	36.363	0.000
MB BY				
MB1	0.617	0.025	25.033	0.000
MB2	0.636	0.024	26.212	0.000
MB3	0.622	0.024	26.202	0.000

	Estimate	S.E	Est./S.E	Two-Tailed P-Value
INSTR BY				
INSTR1	0.748	0.020	37.549	0.000
INSTR2	0.713	0.020	35.849	0.000
INSTR3	0.706	0.023	31.098	0.000
IDEAL BY				
I1	0.502	0.026	18.984	0.000
I2	0.585	0.023	25.100	0.000
I3	0.669	0.023	29.179	0.000
I4	0.667	0.020	33.142	0.000
PAR WITH				
INTOR	0.740	0.026	28.433	0.000
SELF WITH				
INTOR	0.463	0.031	14.987	0.000
PAR	0.402	0.031	12.886	0.000
OUGHT WITH				
INTOR	0.814	0.029	27.986	0.000
PAR	0.853	0.029	29.261	0.000
SELF	0.486	0.035	13.849	0.000
EX WITH				
INTOR	0.645	0.031	21.118	0.000
PAR	0.665	0.030	22.199	0.000
SELF	0.775	0.023	33.026	0.000
OUGHT	0.649	0.035	18.515	0.000
MB WITH				
INTOR	0.708	0.036	19.792	0.000
PAR	0.791	0.033	24.176	0.000
SELF	0.612	0.033	18.490	0.000
OUGHT	0.760	0.040	19.199	0.000
EX	0.922	0.029	31.968	0.000

A :: FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535%AA FNOAG // XCKPW ? 9BE // 17B = 6 > ? 9BEB 17B = ZV28TZQU - < 2535%AN
OAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ? NOAGWBN7 = ? 9WNOAGWB :: E = ? 9WB8FW ? 9BEB 17B981% > ? 9BEB 17 ?
? // GW 9016BENGLISH ? IMPACTB :: 17B = 6PW8FW ? 4 > JHY4HS83QJH ? // GWBSRIBLANKAFW ? 9 > ? // GWB :: 17B = 6PW8FW ? 4 > JHY4H913QJH
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8
FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ? FWB :: 17B = 6PW ? 9BEB 17OAGWB :: 7 = ? 9 = W8FW8FW ? 9BEBW8FW8FW ?

	Estimate	S.E	Est./S.E	Two-Tailed P-Value
INSTR WITH				
INTOR	0.819	0.023	36.395	0.000
PAR	0.852	0.025	34.173	0.000
SELF	0.429	0.031	13.868	0.000
OUGHT	0.799	0.033	24.320	0.000
EX	0.638	0.030	21.577	0.000
MB	0.774	0.034	22.891	0.000
IDEAL WITH				
INTOR	0.714	0.027	26.273	0.000
PAR	0.640	0.030	21.645	0.000
SELF	0.700	0.028	24.707	0.000
OUGHT	0.734	0.033	22.450	0.000
EX	0.763	0.029	26.197	0.000
MB	0.717	0.035	20.765	0.000
INSTR	0.576	0.031	18.323	0.000
MB3 WITH				
MB1	0.299	0.035	8.494	0.000
PAR1 WITH				
PAR2	0.221	0.035	6.296	0.000
Intercepts				
INTOR1	3.736	0.092	40.763	0.000
I1	3.368	0.080	42.290	0.000
PAR1	4.108	0.103	39.848	0.000
PAR2	4.401	0.115	38.227	0.000
SELF1	3.235	0.073	44.515	0.000
O1	3.841	0.090	42.736	0.000
INTOR2	4.266	0.115	37.104	0.000
I2	3.417	0.083	41.334	0.000
MB	14.097	0.098	41.668	0.000
O2	3.954	0.102	38.872	0.000
INSTR1	4.810	0.152	31.600	0.000
INSTR2	4.602	0.120	38.202	0.000

	Estimate	S.E	Est./S.E	Two-Tailed P-Value
MB	34.080	0.104	39.298	0.000
SELF2	3.045	0.066	45.930	0.000
EX2	3.737	0.088	42.269	0.000
INSTR3	4.500	0.122	36.772	0.000
SELF3	3.240	0.071	45.425	0.000
INTOR3	4.000	0.098	40.970	0.000
EX3	4.109	0.102	40.170	0.000
SELF4	3.304	0.073	45.568	0.000
PAR3	4.439	0.123	36.225	0.000
I3	3.849	0.094	41.017	0.000
O4	3.551	0.082	43.429	0.000
I4	3.294	0.074	44.791	0.000
MB	24.022	0.095	42.489	0.000
INTOR4	4.343	0.110	39.495	0.000
EX4	4.073	0.095	43.085	0.000
PAR4	4.672	0.131	35.652	0.000
Variances				
INTOR	1.000	0.000	999.000	999.000
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				
INTOR1	0.779	0.026	29.990	0.000
I1	0.748	0.027	28.239	0.000
PAR1	0.683	0.028	24.255	0.000
PAR2	0.700	0.029	24.240	0.000
SELF1	0.651	0.028	23.089	0.000
O1	0.631	0.028	22.213	0.000

A :: FNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU- < 2535%A AFNOAG//XCKPW?9BE//17B=6>?9BEB17B=ZV28TZQU- < 2535%
NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?NOAGWBN7=?9WNOAGWB::E=?9WB8FW?9BEB17B981%>?9BEB17?
?//GW9216BENGLISH?IMPACTB::17B=6PW8FW?4>JHY4HS83QJH?//GWBSRIBLANKAFW?9>?//GWB::17B=6PW8FW?4>JHY4H933QJH
G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8G%2AXGML1981^9#DRM<L%A9M5<ZV28TIQU<2535%AXG%L1UM8
FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?FWB::17B=6PW?9BEB17OAGWB::7=?9=W8FW8FW?9BEBW8FW8FW?

	Estimate	S.E	Est./S.E	Two-Tailed P-Value
INTOR2	0.493	0.027	17.951	0.000
I2	0.658	0.027	24.175	0.000
MB1	0.619	0.030	20.347	0.000
O2	0.626	0.033	19.185	0.000
INSTR1	0.441	0.030	14.816	0.000
INSTR2	0.491	0.028	17.316	0.000
MB3	0.613	0.030	20.719	0.000
SELF2	0.502	0.031	16.282	0.000
EX2	0.629	0.026	24.069	0.000
INSTR3	0.501	0.032	15.616	0.000
SELF3	0.410	0.026	16.062	0.000
INTOR3	0.531	0.031	17.319	0.000
EX3	0.486	0.030	16.373	0.000
SELF4	0.520	0.029	18.210	0.000
PAR3	0.436	0.031	14.195	0.000
I3	0.552	0.031	17.973	0.000
O4	0.763	0.028	27.726	0.000
I4	0.554	0.027	20.623	0.000
MB	20.596	0.031	19.309	0.000
INTOR4	0.556	0.029	19.007	0.000
EX4	0.509	0.027	18.838	0.000
PAR4	0.461	0.030	15.520	0.000

