BRITISH COUNCIL

Pratham
Every Child in School and Learning Well...

असर ASER
Evidence for Action

English Impact Report:
Investigating English Language Learning Outcomes at the Primary School Level in Rural India

Edited by Vivien Berry

English Impact Report: Investigating English Language Learning Outcomes at the Primary School Level in Rural India

Edited by Vivien Berry

ISBN 978-0-86355-721-7

10 Spring Gardens
London SW1A 2BN, UK

www.britishcouncil.org

Edited by

Vivien Berry
Senior Researcher, English Language Assessment, British Council

Photographs by

Christopher Tribble
www.ctribble.co.uk

Contributors

Professor Raghavachari Amritavalli
Professor, The English and Foreign Languages University, Hyderabad

Dr Rukmini Banerji
Director, Pratham-ASER Centre

Ranajit Bhattacharyya
General Manager, Pratham-ASER Centre

Savitri Bobde
Head, Assessment Unit, Pratham-ASER Centre

Dr Debanjan Chakrabarti
Head, English Language Policy Research and Publications, British Council, India

Jamie Dunlea
Researcher, English Language Assessment, British Council

Karen Dunn
Consultant, Lancaster University

Professor Barry O'Sullivan
Senior Adviser, English Language Assessment, British Council

Professor Jason Rothman
Professor of Literacy and Multilingualism, University of Reading

Professor Jeanine Treffers-Daller
Professor of Second Language Education and Director of the Centre for Literacy and Multilingualism, University of Reading

Contents

Foreword
The English Impact Report

Martin Davidson
Chief Executive, British Council

I am especially pleased to be in Delhi for the launch of the British Council and Pratham-ASER's first report on trends in English learning performance. As education work becomes increasingly focused on quality, this analysis of a large-scale survey of learning outcomes and the emphasis on actual performance is important. There is a growing interest in what the world's children are learning and how this learning can be measured and assessed. These are questions that will drive improvements in learning outcomes.

Issues such as these are becoming increasingly important as English as a *lingua franca* continues to grow across the world. The British Council remains committed to supporting access to English as a means of international communication and opportunity, but within the context of respect for local education traditions and languages. And this is certainly true of India where our partnerships in education and English language are a key part of the relationship between our two countries – a relationship which continues to grow.

This year I was privileged to attend the UK–India Education Forum, where I discussed future education collaborations with the UK Minister for Universities and Science, David Willetts, and the Indian Minister of Education, Pallam Raju. I met the Minister of Education from Maharashtra, Rajendra Darda, at the UK–India English Partnerships Forum, with whom the British Council is working to deliver the Maharashtra English Language Initiative for Primary Schools, an ambitious teacher education project which aims to train and support 67,200 primary school teachers across the state. Over the coming years, through developing such relationships and projects, we aim to work with up to 1.5 million teachers of English in India.

This report is the first fruit of our partnership with the Pratham-ASER Centre. By sharing our experiences and expertise, by building partnerships between our two countries and contributing to the evidence that will drive educational excellence, I believe we can make a real difference to learning outcomes in both our countries. And this is not only true of India. I hope that our collaboration will be seen as an example for similar partnerships elsewhere in the world.

Messages

Dr Madhav Chavan

Co-founder and CEO, Pratham Education Foundation

Teaching English to primary school children has been a matter of some controversy in India – due both to our colonial past and also for pedagogical reasons. But people are increasingly voting with their feet. There is a strong and clear demand for English in urban and in rural areas. This is true for households across the economic spectrum. Delivering English on the ground, however, is a challenge. Contrary to what some people think, English is not widely spoken or heard, especially in Indian villages. Children have hardly anywhere to go to hear English or practise speaking English. Where they live, most adults do not have experience of English either. Yet, English vocabulary has been creeping into most Indian languages. Perhaps that is the best way for India to get ready for English learning.

Working with children, Pratham has identified another challenge for learning English – the fact that many Indian children have difficulty reading their own language. Weakness and deficits in the first language are bound to get compounded by the time the child learns another language. As a country, we need to find new teaching/learning practices and delivery mechanisms for English for all ages. People of all ages want to learn English, and to do that we need to understand how much English people at different ages and in different parts of the country know.

We at the Pratham-ASER Centre are happy to collaborate with the British Council to explore how English can be easily and effectively brought into the lives of Indian children. There is no doubt that it is a key to individual progress and a requirement for India's economic growth.

Rob Lynes
Director, British Council, India

On a cold winter morning our Project Monitoring and Observation team in Muzaffarpur, Bihar, spoke to a very bright girl from a rural farming community. Sandhya attended the local high school where Himanshu Shekhar, a teacher educator who has benefitted from the British Council's DFID-funded Bihar Language Initiative in Secondary Schools (BLISS), taught. Asked why she thought English was important for her, Sandhya's unhesitant reply in Hindi was: 'People think you are as good as an illiterate if you don't know English.' Her mother, who sent all her four children to school despite the family's very modest means, added approvingly: 'She wants to be a teacher when she grows up.'

As I watched this conversation on the film our team made, I was reminded of the observations made in the Position Paper on English in India's National Curriculum Framework of 2005: 'English is in India today a symbol of people's aspirations for quality in education and a fuller participation in national and international life.'

Sandhya's aspirations could be true of almost every one of the estimated 230 million pupils who enrol in primary school every year in the world's largest school education system. I have come across very similar stories from almost all of the 11 states the British Council has worked in to improve the standards of English teaching, in partnership with the state governments.

Pupils, parents and policy makers pin enormous hopes on English to help lift individuals, families and communities out of poverty and indignity by providing access to education, employment, opportunity and social mobility. However, research shows that all's not well with English learning across India, especially at the primary level where the foundations are supposed to be laid.

Pratham, which is perhaps the world's largest NGO, has every January since 2005 published the Annual Survey of Education Report, or ASER (which means 'impact' in Hindi), which has become a not-to-be-missed event in education and development circles in India. In 2012 ASER was conducted in 567 of the 640 districts in India, reached more than 16,000 villages, nearly 330,000 households and surveyed almost 600,000 children aged three to 16.

It is therefore apt that the British Council and the Pratham-ASER Centre have come together to provide an in-depth analysis of the English learning outcomes data that has been collected by ASER since 2007.

I am also glad that this research project is a truly cross-cultural and cross-continental collaboration, with our colleagues in the English and Exams research team in the UK working with colleagues in India, and also involved the recently established Centre for Literacy and Multilingualism at the University of Reading. The involvement of the University of Reading in this project is also an indicator that opportunities abound in India for the English language education sector in the UK to actively take part in India's story of development.

By putting language learning outcomes in primary schools under an analytical lens, we hope to come up with a list of recommendations that will help policy makers and implementers take evidence-based decisions about the teaching and learning of English, teacher education and, most importantly, continuing professional development opportunities for teachers.

This research will also help the British Council and our partners to identify further research questions and areas to focus on, both in geographic and academic terms, in the future.


Narration
(1) Direct Narr
Ram said to his

Introduction

Ranajit Bhattacharyya
Pratham-ASER Centre

Debanjan Chakrabarti
British Council, India

About this report

This report is the result of a truly cross-cultural and multidisciplinary collaboration between Indian and UK institutions, and stems from a strategic partnership that the British Council has with Pratham. Colleagues from the Pratham-ASER Centre worked in tandem with the English and Exams research teams in India and the UK.

Following introductory messages from executives of the British Council and the Pratham Education Foundation, we begin with an article by Rothman and Treffers-Daller that explores the links between multilingualism and cognition in young learners. They illustrate how being able to communicate using several languages benefits society through fostering intercultural understanding; they also outline the cognitive advantages gained by multilingual individuals who switch between languages on a daily basis.

Amritavalli then lays out the policy context of English in the Indian school curriculum and the role of language within India's multilingual context in the early years of schooling. She makes a powerful plea for a bespoke approach to English language teaching for practically every school child in India.

In the next article, Banerji and Bobde discuss the development and evolution of the tools used to conduct the ASER studies, effectively the largest household survey in India. Their focus is on the instrument used to collect data on English language learning.

Dunlea and Dunn then present what is essentially a pilot study to investigate innovative ways of analysing ASER data on English. They attempt to trace the trend in English performance for the years the data exist (2007, 2009, 2012) and the relationship between home or mother tongue (L1) literacy performance and English reading performance (for 2012 only).

The report concludes with comments from O'Sullivan, who puts forward suggestions arising from results of the analysis and discusses the possible relevance and implications of the analyses on English education in India.

Who is this report for?

While this report is primarily for those involved in the framing and implementation of English language policy in education systems in India, it has wider implications for countries with a similarly wide cache of multicultural and heteroglossic capital.

The report is also meant to be a useful tool for the wider community of scholarship on English language teaching, assessment and evaluation, and for institutions and individuals involved in measuring or seeking measurable outcomes from educational interventions.

What does this report hope to achieve?

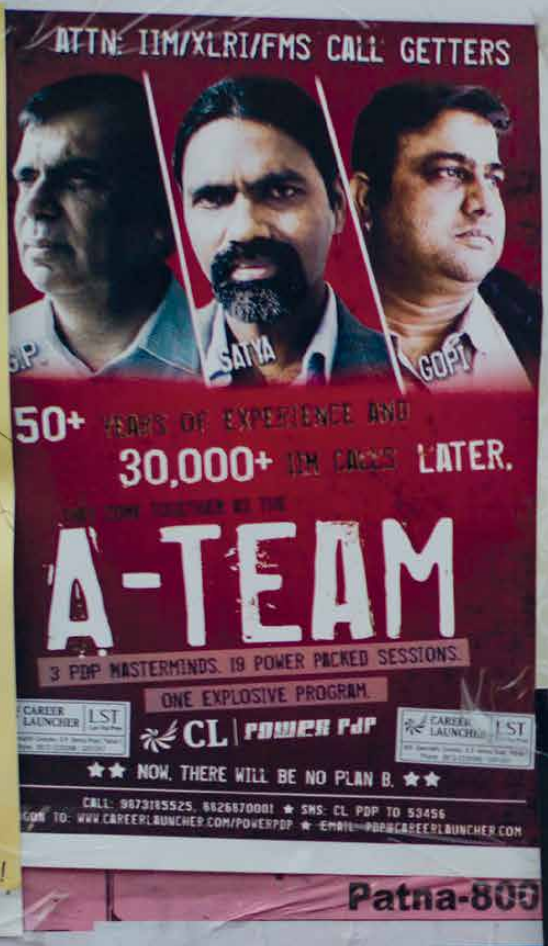
We place this report against the enormously complex backdrop of English language education in India, trying to match aspiration to reality and looking at evidence of how much English India's school-going children are actually learning, certain only of the fact that this is but the beginning of an enquiry and not the end of the journey.

The report is categorically not about whether India needs English or not, or how much English should be taught, by whom or of what kind. Instead, the report is meant to stimulate a wider debate about education and the role of languages in early years of schooling and whether English language learning outcomes is indeed one way of assessing the health of education systems.

The body of evidence presented in this report does not seek to draw hasty conclusions but instead attempts to point towards broader inferences that could have a bearing on language policies in public-funded education systems and the improvement of learning outcomes.

English Impact Report

#9162197779 #9934722375



दिल्ली अब आ
AN EXCLUSIVE UPSC/
श्रेय-
U.P.S.C./
के
G.S
By:- दुर्गेश कुमार
(एवं टीम दिल्ली
नामांकन जारी
Note:-हिन्दी एवं English माध्यम

1

Multilingualism in an international context

Jason Rothman and **Jeanine Treffers-Daller**
University of Reading

Abstract

In this paper we outline how using several languages benefits individuals as well as the society in which they live. On the societal level, the ability to communicate with a wide range of speakers of different languages with different cultural backgrounds promotes intercultural understanding. It also helps companies that employ multilinguals to obtain contracts abroad. For multilingual individuals, there are cognitive advantages, as they are mentally more flexible than monolinguals because they practise switching between languages on a daily basis.

Most people in the world speak more than one language. In fact, when we look at the map of the world and compare the number of countries (around 150) with the number of languages (in the many thousands) it is obvious that there must be lots of speakers of different languages in each country, and that many people must be bilingual (Grosjean, 2010). It is much less well known that there are very many multilinguals too, that is people who speak three, four or even five languages on a daily basis. This is particularly common in Asia, Africa and South America, but also occurs among immigrants in Europe and North America. In Malaysia, for example, it is very common to grow up with Chinese and Malay and to learn English at school (Azman, 1999). Multilingualism is strongly valued in Malaysia, as everyone realises that being able to speak three languages gives Malaysia a global competitive edge. Given the proliferation of native languages in the context of India, multilingualism thrives there in a similar way as well. It is the economic value of multilingualism that is so important in the booming informal market in sub-Saharan Africa: being able to communicate in the language of one's customer is, of course, a distinct advantage. In Europe and North America it is often the immigrants who are the most multilingual. It is not difficult to find a taxi driver in New York who emigrated from Haiti, and speaks Haitian French Creole, Standard French and, of course, English (Myers-Scotton, 2006).

The advantages of being able to speak more than two languages are not just of a practical nature and are not to be understated. Knowing about the cultures associated with different languages can help to overcome intercultural differences in communication. In our globalised world we come across people from different cultures wherever we go: at work, at school, when we go shopping or take part

in sports. Because of differences in our cultural background, we sometimes formulate sentences in ways our listeners do not expect. An overseas student in the UK might find it difficult, for example, to choose the right politeness level for asking a question. Writing 'Please check' in an email to a tutor at university could be seen as impolite, as the tutor might expect to see a longer, less direct request, such as 'Could you please comment on my draft?'. Listeners who are aware of intercultural differences through their knowledge of other languages and cultures will be more likely to accept formulations that are different from the expected norms, which facilitates mutual understanding and respect.

It is very common for bilinguals and multilinguals to switch between their languages. Roma children in Bulgaria have been found to switch effortlessly between Romani, Turkish and Bulgarian, for example, although they can also choose to use Bulgarian only if they meet a monolingual speaker of that language (Kyuchukov, 2002). It is normal for multilinguals to switch between their languages whenever they meet others who know the same languages. For researchers, a key question is how multilinguals can juggle different languages in their minds, making sure the right language is chosen depending on the situation and the listeners, and suppress languages that are irrelevant at the moment of speaking.

It is this multi-competence, the existence of two or more languages in the mind of the speaker, which makes multilinguals special (Cook, 2008). If we compare the French sentences produced by French-Dutch bilinguals in Brussels to speakers from France who speak only French, we often note there are differences in pronunciation, word choice or word order (Baetens Beardsmore, 1971). As listeners, we often try to pick up on those subtle differences in expressions which show where speakers are from, and whether they are native speakers or not. For psychologists and linguists it is interesting to find out what this means for the information we have stored about our languages in our minds. It probably means that French speakers from France or from Brussels have (slightly) different mental grammars, and multilinguals who know French, Dutch and English will not be the same as Dutch-French bilinguals.

As cognitive scientists, linguists and psychologists, we put aside the negative associations many people have with mixing two languages, speaking with a foreign accent or even false beliefs about the relative utility or superiority of some languages compared to others. It is true that language produced by bilinguals and multilinguals is different from that of monolinguals, but it is important to try and understand more deeply why this is so and what the positive effects of knowing more than one language are. From a cognitive and linguistic perspective, despite fine-grained beliefs to the contrary, there are no 'harder' or more 'useful' languages. All languages fulfil the same remit of linking meaning-to-sound/gesture correlations for communication and are effortlessly acquired by children sufficiently exposed to them. By-products of knowing more than one language can be labelled as advantages or disadvantages, depending on the outcome result applied to specific tasks or situations.

What research has revealed in the past decades is that knowing more than one language results in real changes to the mind/brain, some of which entail advantages for certain tasks while others present some challenges (see Bialystok, 2009 for review). We now irrefutably know that exposing children to multiple languages is not confusing for the child in the least. This might seem counter-intuitive and is even in sharp contrast to what many well-intentioned people believe and share with concerned parents, educators and policy makers. However, the research findings on the topic do not lend any support to this once-prevailing view. This does not mean that multilingual children follow the exact same developmental paths as monolingual children. How could they? By definition, they are not monolingual. Children exposed to multiple languages will often display what appear to be delays in the acquisition process. For example, a child exposed to Hindi, Bengali and English at an early age might display smaller vocabularies in each of these languages compared to appropriately matched monolingual children. Not only do they eventually catch up and are able to communicate in three languages later, we now know that they are not really delayed at all. When you add the sum of their three vocabularies together they far exceed the lexical knowledge of monolinguals and when you test their linguistic and cognitive development independently they show no signs of true delays.
It should not be alarming, but rather expected, that multilingual children mix the languages they are acquiring at various levels. Research has shown that such behaviour is not at all random, but governed by universal principles that conform to expected paths of language learning (e.g. Muysken, 2000). Bilingual children also acquire much earlier concepts about language and its social milieu. Even young multilingual children know that these codes are different entities and are used socially in different ways and with different people.

Beyond being able to communicate with many more people and all that this entails, what are the cognitive benefits, if any, associated with early acquisition of more than one language? Studies show that early multilingualism sharpens certain cognitive functions, such as the ability to suppress irrelevant information and working memory, to name just two. These benefits are useful for later language learning and information processing more generally, as well as other everyday tasks (Bialystok, 2009). Benefits also extend beyond childhood, and recent research has shown that early multilingual acquisition correlates with later onset of symptoms for neurological degenerative disorders such as Alzheimer's and dementia. Essentially, we could think of language acquisition in general as one form of mental exercising, and just as we might expect more and earlier cardio-exercise to correlate with increased health benefits throughout our lifetime, so too do the benefits of this early mental linguistic exercise for the mind.

On all planes, there is so much more to be understood about multilingualism. Future research endeavours to reveal the links between how knowing multiple languages shapes the way we perceive the world and how studying multilingualism will open new doors to understanding and create unique glimpses into the mind. Although we have just scratched the surface and have not even conceived of all the relevant questions there are to be asked about multilingualism, what we definitively know is that multilingualism is a good thing, not only for our global world but also cognitively for individuals.

References

Azman, H (1999) *Multilingual Literacies in Rural Malaysia: National Goals and Local Practices*. Nedlands, WA: University of Western Australia.

Baetens Beardsmore, H (1971) *Le francais regional de Bruxelles* (Institut de Phonetique, Conférences et Travaux, vol. III).

Bialystok, E (2009) Bilingualism: the good, the bad and the indifferent. *Bilingualism: Language and Cognition* 12 (1), 3–11.

Cook, V (2008) *Second language learning and language teaching.* 4th edition. London: Arnold.

Grosjean, F (2010) *Bilingual. Life and reality.* Cambridge, MA: Harvard University Press.

Kyuchukov, H (2002) Code-switching among trilingual Turkish-speaking Roma children in Bulgaria. *Psychology of Language and Communication* 6 (1), 7, 75–84.

Muysken, P (2000) Bilingual speech. *A typology of code-switching*. Cambridge: Cambridge University Press.

Myers-Scotton, C (2006) *Multiple voices. An introduction to bilingualism*. Oxford: Blackwell.

2

An English for every schoolchild in India

Raghavachari Amritavalli
The English and Foreign Languages University, Hyderabad

Abstract

English in India spans the first-second-foreign language spectrum because learning opportunities for the language are mediated by variation in environmental exposure and teacher competence. Systems delivering comprehensible input in multilingual, language-across-the-curriculum contexts could allow teachers to be learners, and learners to outpace teachers. Evaluation geared to language proficiency rather than curricular achievement would allow curricular freedom, certification of diverse attainment levels and provide alternate routes to success in English for those leaving school without a 'pass in English'.

A policy for a range of English-teaching contexts

India is a country of continental diversity, especially in its linguistic landscape which, in the seventh decade of our independence, retains and continues to assimilate English into itself. Initially thought of as the language of the intelligentsia, English today is the language of opportunity. A policy for its place at school necessarily needs to temper academic wisdom and ground reality with this imperative of parental aspiration.

English is for some Indians a first language of public (academic, societal, creative …) discourse. For many it is a second language, and for many others a foreign language. The school contexts in which English is taught mirror this contextual diversity; their adequacy is affected by the twin variables of teacher competence and environmental exposure to English. Consequently, the policy for teaching English articulated in the National Curriculum Framework (NCF) avoids methodological prescriptivism. It aims, rather, at a curricular cohesion grounded in guiding principles for language teaching and acquisition, which could accommodate a variety of implementations suitable to local needs and resources, and inform and rejuvenate them. The policy attempts to identify delivery systems for comprehensible input to the child, whether in the classroom or outside of it.

Language and learning in and out of school

The learning of English at school is best approached as a focused harnessing of the natural human ability to learn languages, given an environment of meaningful exposure to them. The curricular challenge is to find ways of approximating the language learning opportunities of the classroom to those of naturally supportive language learning environments. Happily, this integrative thrust in the policy for English coheres with the larger vision in the NCF of integrating the life of the child within and outside school. For early language teaching, this thrust could translate into initiatives for interaction of the English class with the English speakers available in the community, whether face to face or through media (audio, video, print); a better understanding of children's language use and language learning through stories and story reading with an adult, or through oral activities and play; and an emphasis on using language to understand simple spoken and written texts, and to similarly express oneself (even if imperfectly), rather than on language learnt as an object to display to evaluators as evidence of time and effort expended on its teaching.

In other words, language learning is a process of knowledge building in the learner, not a product for knowledge transfer from the teacher. Incidentally, this understanding of what constitutes language teaching at school needs to inform the teaching of languages other than English (report in *The Times of India*, 2013;[1] Nag and Snowling, 2011).

An input-rich communicational environment

A first methodological consequence is that to teach English is to maximise the learning opportunities for it. Our experience of multilingualism shows that children (and often, adults) naturally pick up the languages they are meaningfully exposed to. Learning opportunities for English may include, but not be limited to, methods of second or foreign language instruction such as communicative activities or activity-based learning. Such methods must find their place within a broader concept of an input-rich environment; otherwise they may degenerate, where not ignored altogether, into activity for its own sake with no accompanying language (Internal Report, *Teacher Education through School-based Support in India*: a UKAID-GOI project, 2013) or be limited to the formulaic learning of set phrases.

Language is a 'dynamic' text (Amritavalli, 1999). What counts as exposure is the encounter with new occurrences of comparable language samples, rather than mastery learning by repeated teaching of a single prescribed textbook. In the Bangalore Project (Prabhu, 1987), the 'texts' for language learning were the classroom discourse created by the teacher and the students around each task, which resulted in the 'recurrence' of language. Second or foreign language learners spontaneously acquire 'teacher talk', the language of classroom management. Whole Language approaches (Mangubhai, 2011; Jangid and

[1] The Telugu quarterly examination papers for classes VII to X, set 'in line with the recommendation of the official language commission' … 'had tough questions in reading, comprehension and writing, a pattern very alien to most students.' The report quotes a teacher: 'Only students with good understanding of Telugu can answer the questions.'

Amritavalli, 2011) validate the genre of 'predictable' stories in which events, and the language that narrates them, recur.[2]

Such approaches allow as input both spoken and written discourse. In India, reading has historically been the route into English for autonomous learners. Reading aloud to and with pre-readers is known to promote literacy (Adams, 1990), and could counter the fear of 'unseen' passages for comprehension in tests or examinations. Teachers and learners must evolve for their own classrooms a balance between 'prescribed' texts and learner-chosen texts from class libraries with print or audio-visual materials (Big Books, multilingual books, Reading Cards, learner magazines/newspaper columns, edutainment programmes). The language environment of disadvantaged learners can be enriched by developing schools into community learning centres.

English does not stand alone

English is taught by Indians to Indians so that we may interact with one another and with the world. The acknowledgement that English is a global language in a multilingual country has the (second) methodological consequence that we need not insulate it from our other languages in the classroom (as the audio-lingual era did), any more than in our everyday lives. It has been an abiding national vision that the teaching of English creates multilinguals to enrich all our languages. We have seen emerge in this century a rich mix of English and other languages in television and film, reflecting the current educated urban experience.

Within the school system, the Kendriya Vidyalayas, or Central Schools, have emerged as successful models of bilingual (English-Hindi) school education. However, 'English medium' schools, old or new, may continue to subscribe to an isolationist perspective on English. The result is not only a loss of one's own languages from the arena of modernity, but an injection of the burden of sheer incomprehension into the language classroom. It is common sense to use our existing knowledge, including the knowledge of other languages, to help us make sense of what is said or written in the new language. One's other languages can also help to scaffold expression in the new language.

A judicious multilingualism

We need, however, to distinguish the displacing of English by known languages in the classroom (as when prescribed English texts are explained in them) from the use of known languages to provoke and sustain an effort to engage with English texts in the classroom. The classroom discourse need not be monolingual, whether in English or in an Indian language. At the early levels of schooling, there are no teachers specialised for English teaching; a single teacher may handle more than one language, as well as the rudiments of the sciences. If that teacher could use English when they teach these other subjects, and other languages in the English classroom, the barriers between languages and 'languages' and 'subjects' could

[2] Amritavalli (1999) analogises language learning to raga recognition in Indian classical music, which develops through exposure to recurrent but not repeated samples of free renditions of the raga.

be removed. English could occur in tandem with known language(s) for learning activities. The tasks in the Bangalore Project (Prabhu, 1987) were input in English and required responses in English, but known languages were used to make the task language comprehensible.

Some of our classrooms are multilingual enough that the teacher may not share a language other than English with students; but groups of students may be permitted to create multilingual discourse among them. The legitimate use of other languages to promote the learning of English is a matter of context-sensitive understanding emerging from a tolerance of some language mixing, and the infusion of thought-content into the language class. Currently, the 'mother tongue' enters the class as a surreptitious intruder, keeping out its use in, for example, pre- and post-reading discussion activities, or in bilingual dictionaries or multilingual texts. Children's publishing in India has traditionally been parallel across languages; recently, multilingual story books have emerged for beginning readers. Such materials, popular with thoughtful young parents, do not find a place in our classrooms.

The goals of language teaching and the question of standards

The early years of schooling aim to develop proficiency in ordinary English as a base for later academic or professional language use. Given an input-rich environment, children's language learning can, but may not always, outpace the teacher's competence in English. The diversity of learning environments, therefore, must allow for a diversity of attained standards. These could be appropriately certified by evaluation (including continuous evaluation), not of achievement within particular syllabi, but of language proficiency with respect to national benchmarks. This would balance curricular freedom with standards of attainment, and open up alternative routes for English certification (and therefore instruction) outside schooling, addressing the problem that English (along with mathematics) is a principal reason for failure at the Class X school-leaving examination. A student may, where appropriate, be certified to 'pass without English' after ten years of schooling.

The policy envisages an input-rich environment that promotes the teacher's English alongside that of students. The teacher may, as a learner, develop a feel for 'the occurrence of learning' (Prabhu, 1987), which happens in a 'zone of proximal development' (Vygotsky, 1978) or an 'i+1' zone (Krashen, 1985). Adams (1990: 35) found that successful early reading instruction depends (irrespective of method, materials, objectives, class size or organisation) 'on the atmosphere – the momentum, support and expectations – created by the classroom teacher … the teacher's ability to stay tuned to that delicate interval between ease and difficulty for the students and to keep the instruction within it.' This feel for the pace of instruction can develop only when teachers are allowed to assume responsibility for their own and their students' learning.

References

Adams, MA (1990) *Beginning to read: thinking and learning about print*. Cambridge, MA: MIT Press.

Amritavalli, R (1999) 'Language as a Dynamic Text,' in R Amritavalli (Ed.) *Language as a Dynamic Text: essays on language, cognition and communication* (pp.1–17). Hyderabad: Allied Publishers (CIEFL Akshara series).

Jangid, G and Amritavalli, R (2011) Whole language, story reading, and children's writing. *The EFL Journal* 2/2: 15–37.

Krashen, SD (1985) *The Input Hypothesis: issues and implications*. London: Longman.

Mangubhai, F (2011) Reading, whole language, and second language learning: a personal journey. *The EFL Journal* 2/2: 1–13.

Nag, S and Snowling, M (2011) Reading comprehension, decoding skills and oral language. *The EFL Journal* 2/2: 85–105.

National Council of Educational Research and Training (2005). *National Curriculum Framework*.

Prabhu, NS (1987) *Second Language Pedagogy*. Oxford: Oxford University Press.

Teacher Education through School-based Support in India: a UKAID-GOI project. (TESS India) (2013) *Development Testing Report: Elementary English*.

The Times of India (2013) Telugu paper proves to be an acid test. October 01, 2013.

Vygotsky, LS (1978) *Mind in Society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

3

Evolution of the ASER English tool

Rukmini Banerji and **Savitri Bobde**
Pratham-ASER Centre

Abstract

This paper describes the evolution of the English tool framework and the Annual Status of Education Report (ASER) tool first used in 2007 and then used in ASER in subsequent years. The paper also discusses ASER's approach to the development of, and importance of developing, an English tool that is simple yet rigorous.

Background

Over the years there has been much debate in both policy and political circles in India over when children should begin to study English in school.[1] Today, in most states in the country, English as a subject is introduced early in primary school to meet the high and growing popular demand for it. The massive practical challenge across the country is how to deliver English effectively to millions of children. The task is even more arduous considering that close to 50 per cent of all children in the primary school age group are not fluent readers even in their own language (Pratham, 2007; 2009; 2012).

As part of the development of the National Curriculum Framework 2005, a great deal of attention was paid to the teaching of English. The position paper of the national focus group outlines key issues for the teaching of English.[2] The paper stresses 'communicative competence' and argues that teaching should aim to encourage the use of English in meaningful contexts and develop children's ability to use the language in a variety of contexts spontaneously and appropriately (page 3). It also includes a section on evaluation, which mentions that 'measurement should be about language proficiency' and that national benchmarks for language proficiency need to be evolved that teachers and children can use.[3]

[1] For example, see the debates and discussions around this issue in the National Knowledge Commission in 2006 and 2007.

[2] The National Curriculum Framework devotes an entire section to the teaching of English. See www.ncert.nic.in/new_ncert/ncert/rightside/links/pdf/focus_group/english.pdf

[3] The NCF position paper states: 'It is how one evaluates that will decide whether a child will want to be evaluated. While even the most child-centred methods of evaluation will be anxiety provoking for some, there is no question that a system of evaluation must be in place. The question is how and how much.' (page 16)

To facilitate a discussion of how English teaching should be designed in India it would be useful understand how much English children in India currently know. In 2007, when we tried to find answers to this basic question, we found that there were no empirical answers.[4] It is with this background of policy debate, practical challenge and lack of evidence in mind that the assessment of English was included as a key domain in ASER in 2007 and subsequent years.

Developing a framework for assessing children's basic English

In any discussion about ASER tools, it is always important to keep the basic objective of the exercise in mind. ASER is primarily an attempt by citizens to understand the status of schooling and basic learning of the children in their district. The tools are aligned to achieving this objective. Each year, ASER tools are used by thousands of volunteers who assess hundreds of thousands of children in their homes or their communities, meaning the tool needs to be simple to understand, use and interpret. While more sophisticated measurements may be possible with more qualified field investigators and more time, the biggest challenge in ASER is to make the tool as simple as possible without sacrificing rigour.

There was another major reason to keep the English tools for ASER 2007 simple. Across the country it was hard to predict what the English ability of the field investigators would be.[5] At the district level orientation, all field investigators were also given the same test. Based on the results of this assessment, a session was conducted to train the team further to ensure that all field investigators were at least able to use all four samples of the given test.

The third reason to keep the tool simple was to ensure that different stakeholders could use the evidence to bring about change in children's learning outcomes. Simplicity would ensure that not only academics, but also teachers in schools, officials at state, district or block level, NGOs and parents, could understand the basic level of the child and work towards the appropriate action. We also wanted the tool to be useful for teachers to quickly get a sense of the basic levels in their classes and enable them to track the children's progress.

As a starting point, we analysed state textbooks for English across the country, especially the textbook for the first year in which English is introduced in primary school. In keeping with the simplicity and ease of the overall ASER approach, the English tool had to be easy to use and the tasks that children were being asked to do had to be simple. The entire set of tasks could not take more than five to ten minutes per child, and could also not be so difficult that children would be too nervous to even attempt them. As in the case of other ASER tools (other languages and arithmetic), we attempted to see if the tasks could be progressive so that the level that a child reached comfortably could be recorded.

---

[4] Currently in India, there is neither national benchmark nor any country-wide measurement of basic English competency for children in elementary school. Each state textbook indicates the expectations of the state with respect to children's learning. So, as a country we do not know how much English our children know (or, for that matter, how much English our teachers know).

[5] ASER field investigators are volunteers from local district institutions and organisations.

The challenge was to be able to do all of this in a way that was not complicated. The framework was developed with these key elements in mind.

With English, as with the other ASER tools, the same set of tasks is given to all children between the age of five and 16. Given this and the fact that we did not have any benchmarks to refer to, an additional challenge was to build enough variation into the tasks so that the test could be given to children across the five to 16 age group.

What do states in India want their children to learn in English?

In the first year in which English is introduced as a subject, the learning goals for children are centred on the basic abilities of listening, speaking, reading and writing. Implicit in these competencies is basic comprehension of what is being read. Although the format, design and layout of the textbooks vary, most textbooks follow a similar pattern as far as content is concerned.

Table 1: Learning goals in English in the first year of English teaching in primary schools in India

LETTERS				WORDS				SENTENCES	
Listening	Speaking	Reading	Writing	Listening	Speaking	Reading	Writing	Listening	Speaking
Small and capital letters introduced with words and pictures. Many textbooks are similar to workbooks where practice writing can be undertaken in the book itself.				Building up of vocabulary is done with sets of words from familiar contexts such as home, animals, family, school, etc. Sets of similar sounding words are often used. In most cases, the introduction of alphabets is done with words and pictures so that pictures facilitate letter and word recognition.			Practise writing words. Usually copy and write.	A number of different activities including recitation of poems so that children can listen and repeat. Listen and repeat sentences after the teacher. Greetings. Instructions and a few basic sentence structures are practised.	

After reviewing and analysing state textbooks, we decided that a very basic framework would suffice for the ASER English tool. Whether English is introduced early (grade 1) or late (grade 3 or later), the learning goals expected of children do not seem to vary too much. In common with reading in regional languages, English too would be kept at a foundational level – of decoding and reading and very simple comprehension. The tasks to be included in ASER were basic letter recognition, ability to read simple everyday words and the reading of simple sentences. Of the words that were read, we wanted children to tell us the meaning in their own language; we had a similar goal with the sentences.

Developing the ASER English tool

The journey from developing the framework to finalising the actual tool was time consuming and intense. In three months, from June to September 2007, we conducted six pilots in different parts of India involving substantial numbers of children in each round. We had to have a tool that could be used from Nagaland to Ladakh to Kerala. Apart from content, the pilots were also very useful for figuring out many other elements of the tool such as design and layout and testing procedures. Examples included how much text should be on each page? How large should the font be? Should the instructions for the field investigator be on the actual tool or should they be on a separate sheet? Even the choice of font turned out to be an important issue – depending on which was used, some made letters such as 'a' or 'g' either easier or harder to read.[6]

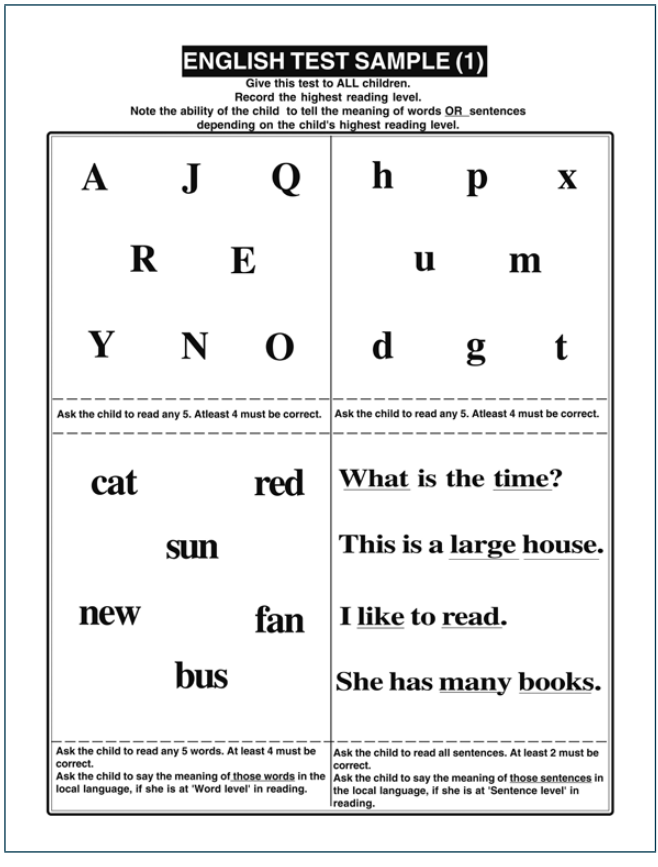
Let us look at an example of how the list of words evolved over the six pilots. First of all we showed children pictures of objects and asked them to name the object. There was significant variation in children's responses to the pictures. For example, the responses to a picture of a cat ranged from 'tiger' to 'dog' to 'animal', while many children did not respond at all to the picture of a comb, a candle or a tap. It seemed that much more experimentation would be needed with pictures before questions like these could be used on scale, so the decision was taken to drop this item. We eventually decided to construct a list of English words where each word on the list had to be (a) a familiar object that children were likely to know, (b) easy to read in English (grade 1), and (c) the equivalent word in their own language had to be easy as well. All three criteria had to be met in all states for them to be included as a word on this list. After each round of piloting, the responses of children to particular words were documented and the word list was modified based on the experience of the pilot. Words that had two different consonant sounds one after the other such as 'crow', 'ship', 'cold', 'star' and 'gold' were dropped to keep the word list simple. Eventually we had a list of three-letter words that met these conditions and could be used across India.

The ASER English tool is like the other ASER reading tools in that the tasks are of an increasing level of difficulty. In this way, at least as far as reading ability is concerned, a child can be marked at the level that he/she reached comfortably. Unlike the other language reading tasks, where children are reading in a language that they speak, in English we introduced very basic comprehension. We ask children to tell us the meaning of the words and sentences that they have read.[7] To maintain equivalency, the same set of tools has been used in ASER in 2007, 2009 and in 2012. For ASER 2007, 2009 and 2012, the English tool has been administered to assess the basic reading and comprehension abilities of all surveyed children (in the five to 16 age group) in English. The tool is a 'floor-level' test that has five levels for reading and two levels for comprehension. The tool has four test forms.

[6] While Times Roman is the font used in many state government textbooks, there are other fonts used as well.

[7] Detailed test administration instructions can be found in the ASER Report 2012 (pages 20–21).

Image 1: A sample of the English tool



ENGLISH TEST SAMPLE (1)

Give this test to ALL children.
Record the highest reading level.
Note the ability of the child to tell the meaning of words OR sentences depending on the child's highest reading level.

A J Q R E Y N O	h p x u m d g t
Ask the child to read any 5. Atleast 4 must be correct.	Ask the child to read any 5. Atleast 4 must be correct.
cat red sun new fan bus	What is the time? This is a large house. I like to read. She has many books.
Ask the child to read any 5 words. At least 4 must be correct. Ask the child to say the meaning of those words in the local language, if she is at 'Word level' in reading.	Ask the child to read all sentences. At least 2 must be correct. Ask the child to say the meaning of those sentences in the local language, if she is at 'Sentence level' in reading.

Concluding thoughts

The ASER English tool has been used with over 1.5 million children from every rural district in India during 2007, 2009 and 2012 (on average, over 522,000 children a year) making it perhaps the most widely used basic English assessment for children in the non-English-speaking world.

The evidence generated in all three years points to the fact that language reading skills, both in regional language and even more so in English, need urgent attention throughout India. Less than half the children in grade 5 can read simple words in English (Pratham, 2012) and, of the children who can read words, approximately 40 per cent cannot tell the meanings of the words they have read. This evidence needs to be widely used in thinking and planning English instruction for our children. Our children have told us what they can do. Now it is up to us to effectively enable them to do more.

References

National Curriculum Framework (2005) Available from
www.ncert.nic.in/new_ncert/ncert/rightside/links/pdf/focus_group/english.pdf

Pratham (2007) Annual Status of Education Report (ASER).
Available online at www.asercentre.org/education/India/status/p/143.html

Pratham (2009) Annual Status of Education Report (ASER).
Available online at www.asercentre.org/education/India/status/p/143.html

Pratham (2012) Annual Status of Education Report (ASER).
Available online at www.asercentre.org/education/India/status/p/143.html

4

English language learning outcomes at the primary school level in rural India: taking a fresh look at the data from the Annual Status of Education Report

Jamie Dunlea
British Council

Karen Dunn
Lancaster University

Abstract

This paper investigates innovative ways of exploring data from the Pratham-ASER Annual Status of Education Report. The paper describes the application of various statistical analysis techniques to investigate trends in English as a second or foreign language (L2) reading performance over time, as well as the relationship between first language (L1) literacy and L2 reading ability. The paper does not attempt to provide definitive answers to these questions, but instead focuses on suggesting useful techniques for exploring them.

Introduction

The Pratham-ASER Centre's annual survey of educational outcomes in rural India is an innovative, and certainly ambitious, large-scale educational survey. The background and methodology of the survey, including the design and validation of the testing tools, are described in detail elsewhere and will not be the focus of this paper (e.g. Pratham, 2012; Ramaswami and Wadhwa, 2010; Banerji and Bobde, this volume; Vagh, 2009). It is worth noting here, however, several important features of the survey that impact on the kind of data it produces and, consequently, on the analysis undertaken for this study: its scale, annually returning results from approximately 700,000 children; its focus, attempting to build a comprehensive snapshot of areas which often remain out of the research and policy limelight; and the methodology itself. The procedures, survey instruments and testing tools are, on the one hand, a product of the severe operational constraints that go hand in hand with the scale and focus of the survey. Overcoming these constraints is made possible by the co-operation of a large number of volunteers in local areas.

At the same time, this participation is not just the answer to a logistical problem. Banerji (2012: 8) notes that the survey 'is fundamentally based on participation and involvement of ordinary people.' Understanding these constraints was an important part of deciding on appropriate analysis techniques, but understanding the aims, impact and underlying philosophy of the survey was equally important for interpreting the results of the study and making the limited set of recommendations that conclude this paper.

Aims and scope of this study

The study was undertaken within the context of a wider collaboration between the British Council and the Pratham-ASER Centre to contribute innovative ideas to the research agenda for English language education in India. The authors were given access to data from the Annual Status of Education Report for the years 2007, 2009 and 2012 – those years in which children were tested for English in addition to L1 literacy and mathematics. The authors were invited to bring a fresh perspective to the data on English language learning outcomes: to consider what new relationships or trends might be investigated and what kind of statistical analysis techniques might be useful for doing so. It is important to stress that the authors were invited to undertake this study as outsiders, allowing them to approach the data without preconceptions and with no predetermined position to promote in relation to the debates surrounding English language educational policy in India. From the outset the study recognised the limitations imposed by this position – a lack of in-depth knowledge of the context. It was thus designed as a pilot study intended to provide information and suggestions and does not attempt to provide answers to policy questions. The suggestions in the report deal as much with evaluating the usefulness of the analysis procedures employed as they do with interpreting the possible trends that those procedures have begun to identify. As such, the statistical modelling approach employed, the rationale for the selection of the approach, and evaluation of the effectiveness of the approach will be described in some detail. It is thus hoped that the paper will be of interest, and accessible, to readers from both perspectives: those without a statistical background who wish to learn more about the trends and recommendations the study highlights, and those interested in the technical aspects of the analysis.

Research questions

The authors initially reviewed the raw data available for the study in conjunction with the reports published online for those years (Pratham, 2007; Pratham, 2009; Pratham, 2012). Three research questions were formulated, which provided the opportunity to explore the data in ways not already presented in the published reports and demonstrated the usefulness of various analysis techniques:

1. Are there any trends in L2 English reading performance in primary school across the different years of the survey?
2. Is there a relationship between L1 literacy and L2 English reading performance in primary school?

3. Given the wide age range of students within each school year, does age have a significant impact on L2 English reading performance within grades in primary school?

The reasons for exploring research questions 1 and 3 are reasonably self-evident and will be dealt with first. The background to research question 2 will be dealt with in more depth below. Research question 1 was approached more from the perspective of identifying some exploratory techniques that could be usefully employed for investigating trends in performance. Obviously, eyeballing the data will reveal differences in descriptive statistics across years. However, such differences can be superficial and should not be the basis of substantive policy decisions without examining whether those differences are significant both in a statistical sense and also in terms of real-world impact. Regarding research question 3, the published reports (e.g. Pratham, 2012) have already made reference to the wide range of ages present in each grade level. The study thus took the opportunity to model age as a variable in predicting L2 reading ability within each grade.

The combination of L1 literacy and L2 English reading data produced by the survey provides an interesting chance to investigate research question 2. The issue has received considerable attention in the literature, often revolving around the original question posed by Alderson (1984, cited in Alderson, 2000: 23): 'Is second-language reading a reading problem or a language problem?' The first alternative is the basis of what is also known as the Linguistic Interdependence Hypothesis (LIH), and the second the Language Threshold Hypothesis (LTH) (Alderson, 2000; Bernhardt and Kamil, 1995; Cummins, 1979). Bernhardt and Kamil (1995) reviewed a number of studies, concluding that L1 reading consistently accounts for approximately 20 per cent of the variance in L2 reading performance. They emphasise, however, that 'in those studies that are able to account for language proficiency, this construct seems to be a substantially more powerful predictor of L2 reading ability' (Bernhardt and Kamil, 1995: 30). Alderson (2000: 23-39) concludes that 'the importance of both factors ... is clearly acknowledged,' but also stresses that 'a linguistic threshold exists which must be crossed before first-language reading ability can transfer to the second-language reading context.' Far from being resolved then, the available evidence has indicated that the original question was overly simplistic.

Research has continued, attempting to clarify more complex interactions, in line with Bernhadt and Kamil's (1995: 32) call to 'consolidate LIH and LTH,' rather than promote one or the other. Indeed, recent research continues to identify the importance of L1 reading as a predictor of L2 reading ability (Alderson and Huhta, 2010; Sparks et al., 2012) while extending the focus of research. Sparks et al. (2012) investigated L1 print exposure as an additional predictor and examined the impact on L2 proficiency across skills, not just reading. In Japan, Yanase (2012) examined the relationship between a broad range of L1 activities, including reading habits, to identify criterial differences between high school students at the A1, A2 and B1 levels of L2 proficiency. Yamashita (2004: 15) investigated whether attitudes to L1 reading were also transferrable to L2 reading, concluding that 'EFL learners'

positive feeling towards L2 reading is likely to originate, at least to some extent, from their positive attitude towards L1.' Clearly the relationship between L1 literacy and L2 reading ability remains on the research agenda. Interestingly, with the exception of one study reviewed by Bernhadt and Kamil (1995), the above studies have focused on secondary school students or adult learners.[1] Indeed, Bernhardt and Kamil (1995) recommended against focusing on primary school learners because of the risk of confounding cognitive development with the effect of L1 reading, as both L1 reading and L2 reading would still be in a formative, unstable state. However, we also need to recognise the worldwide trend to introduce L2 English education into formal education systems at ever earlier stages (Graddol, 2006; Graddol, 2010). Within India this trend is clearly evident (Meganathan, 2011; NCERT, 2005), and all states now offer English in either grades 1 or 3 of primary school (NCERT, 2012), much earlier than the recommendations of the Position Paper of the National Focus Group on the Teaching of English (NCERT, 2005). The data thus presented the opportunity to extend what remains an important area of research to an increasingly important sector, primary school, and to examine the impact of L1 literacy in a multilingual context virtually unique to India.

It is important to note an important caveat related to research question 2. Although we are referring to L1 literacy, the interpretation of this variable is not entirely clear-cut. To take account of its linguistic diversity, India has long promoted a 'three languages formula', which can be generalised as the L1 or regional language, Hindi or one other Indian language, and English or one other European language (Graddol, 2010; Meganathan, 2011; NCERT, 2005). Another central pillar of educational policy has been to offer children, particularly in primary school, the opportunity to receive schooling in their home language (Graddol, 2010; Meganathan, 2011; NCERT, 2005). The challenges associated with making these linguistic choices available across one of the world's largest school systems are considerable, and will likely impact differently on different states and on different schools within states. Meganathan (2011: 19) reports that in 2002, 92.39 per cent of primary schools taught through the mother tongue. The reality in individual schools, however, is likely to be more complex. Graddol (2010: 54) notes that many states struggle to meet the challenge of providing mother-tongue instruction in such a linguistically diverse context, forcing some children to 'start their educational career in a language which may not be used at home.' In relation to the current study, while the instructions for administering the reading test call for children to choose the language they are tested in (Pratham, 2012), we need to be cautious in interpreting that choice as conclusively the home language. In some cases it may instead be the language of schooling, which is not in fact the L1 of the child, while in other cases the child may have chosen their home language, which may not be the language they are exposed to as a medium of instruction.

[1] While Cummins (1979) does cover studies involving learners in primary school, these studies focused on the impact of different variations of bilingual and immersion education on general academic performance, and not on identifying the specific contribution of L1 reading ability to predicting L2 reading ability.

The study

We will begin by providing an overview of the cleaning principles employed to derive the final data sets for analysis. The statistical approach taken to investigate research question 1 is somewhat different to the modelling approach taken for questions 2 and 3. The methodology and results will be presented separately for each research question in turn.

Identifying important characteristics of the data used for the study

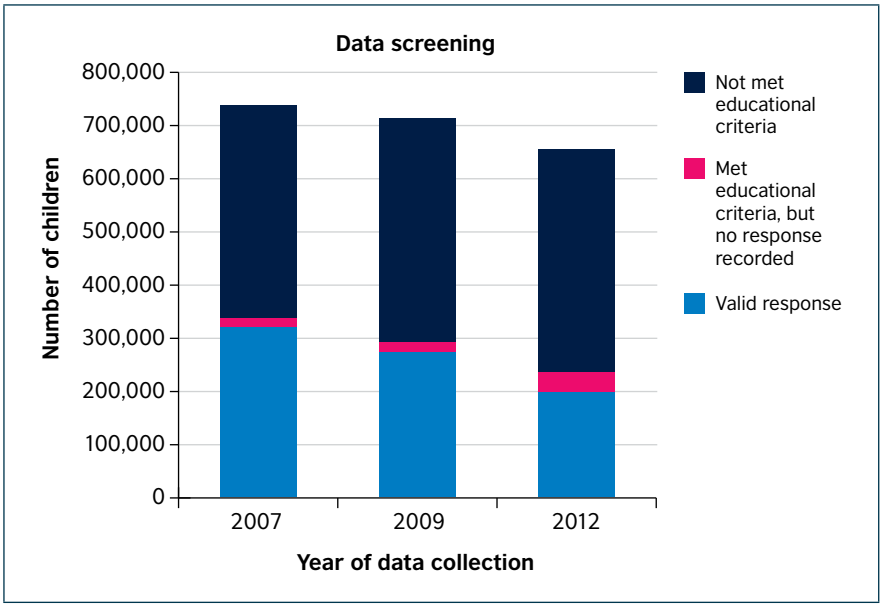
Because of the exploratory nature of this study, and to aid interpretation of the results, it was decided to narrow the focus to facilitate the application of various statistical analysis procedures. Firstly, due to the large differences in sample sizes across states, it was decided to only analyse the data at a national level. The Annual Status of Education Report takes account of these differences when collating state-level results. However, the probability proportional to size sampling method used for doing so would have added a degree of complexity to the modelling procedures used that were not necessary for the exploratory nature of this pilot study.

The Indian education system generally follows a 10-2-3 pattern, with the first ten years further divided into primary (1–5) and upper primary (6–10). The survey covers children in grades 1–10; however, only data from children in primary school grades 1–5, and for whom testing data for both L1 literacy and English were available, were used in this study. Further, children who elected to be tested for L1 literacy in English were dropped from the study. Given the research questions, it made little sense to compare L1 English reading to L2 English reading. It was also possible that children choosing this option were doing so for very different reasons, including that they were studying in a school with English as the medium of instruction. The context of learning and everyday language use for these children would potentially differ in important ways from children who were 'typically' learning English as a second or foreign language in rural schools. There is also considerable variation in the implementation of English medium instruction (NCERT, 2005; NCERT, 2012). The results of data screening on the final samples used for analysis across the three years are summarised in Table 1 and Figure 1, below.

Table 1: selection of cases from data sets for 2007, 2009 and 2012

	2007	2009	2012
N-size prior to screening process	735,662	712,214	654,116
Met educational criteria, but no response recorded	16,129	17,730	36,283
(% of total N-size)	(2.2%)	(2.5%)	(5.5%)
Not met educational criteria	400,030	422,536	420,581
(% of total N-size)	(54.4%)	(59.3%)	(64.3%)
Final N-size	319,503	271,948	197,252
(% of total N-size)	(43.4%)	(38.2%)	(30.2%)

Figure 1: Overview of case selection



The most significant aspect of the data, in terms of the impact on statistical analysis procedures, relates to the nature of the test results. The testing tools employed in the survey are designed around principles of mastery (Pratham, 2012; Vagh, 2011). Survey volunteers use the tests to probe the level of the child, and the highest level of performance the child is able to achieve is recorded in the survey data. For L1 literacy, for example, there are five different potential levels of performance, ranging from unable to recognise individual letters to able to read a story. For the purposes of this study, these levels were treated as an ordinal scale and the data were recoded to provide a score for each child on an ordinal L1 Reading scale[2] and a separate score on an English Recognition scale.[3] Initially, an English Comprehension scale was also considered for those questions that ask children to go on to provide the meaning of English words or sentences. The English Comprehension scale, however, was not included, since only a subset of students were asked the comprehension questions. Whether children were asked the comprehension questions depended on their answers to the English reading recognition questions (see Pratham, 2012, for an explanation of the testing tools and procedures). While this approach is appropriate for the mastery focus of the original survey, the data were not suitable for treatment as an ordinal scale.

While ordinal scale data provide more possibilities in terms of statistical analysis than the kind of categorical data represented by yes/no mastery classifications, ordinal data itself do not meet the assumptions of many statistical analysis

[2] 0=cannot read letters in L1, 1=read letters, 2=read words, 3=read paragraph (Std. 1 level text), and 4=read story (Std. 2 level text)

[3] 0=no L2 English recognition, 1=read upper case letters, 2=read lower case letters, 3=read words, and 4=read at sentence level

procedures that require continuous data on an interval scale, and this restricted the modelling options available for the study (for a discussion of the properties of different scales, see, for example, Bachman, 2004, in relation to language assessment, or Field, 2000, for a discussion in relation to research in the social sciences). The points on the English Recognition scale are treated in the analysis as representing observed manifestations of an underlying ability of English language recognition. It is important that the stepped nature of the data is acknowledged; to treat this as a continuous scale would be to increase the possibility of Type 1 errors, i.e. enhancing the possibility of finding relationships where none exist. For the purposes of the current analysis, the scale is assumed to approximate a single-dimensioned scale of L2 recognition ability. This is clearly a broad assumption; however, it is based on the order in which these steps are treated in the curriculum for L1 and L2 reading in Indian primary schools.

Research question 1: methodology and results

A comparison between the levels of L2 reading was only carried out for the L2 Recognition variable, given the issues with the construction of the Comprehension variable discussed above. Owing to processing limitations on the computers available to the authors at the time of the study, and the limited time available, it was not possible to test for the comparison across a data set containing case-level information from all three years (N=788,703). It was considered that a stratified random sample of ten per cent of this data set would suffice for the purposes of calculating the significance of differences across the three data sets. The proportion of valid cases from each year was maintained as per the original data set. A comparison of the frequencies for the re-sampled data set and the full data set are shown in tables 2 and 3, below.

Table 2: Number of cases across all three years of interest

		Frequency	Per cent	Valid Per cent	Cumulative Per cent
Valid	2007	319,503	40.5	40.5	40.5
	2009	271,948	34.5	34.5	75.0
	2012	197,252	25.0	25.0	100.0
	Total	**788,703**	**100.0**	**100.0**	

Table 3: Number of cases included in stratified random sample

		Frequency	Per cent	Valid Per cent	Cumulative Per cent
Valid	2007	31,950	40.4	40.4	40.4
	2009	27,262	34.4	34.4	74.8
	2012	19,927	25.2	25.2	100.0
	Total	**79,139**	**100.0**	**100.0**	

As discussed above, the data for the English Recognition scale were treated as lying on an ordinal scale. Many tests of statistical significance require continuous data on an interval scale that meets parametric assumptions. However, non-parametric tests are available for categorical and ordinal data (Field, 2000). For this study, the Kruskal-Wallis Test, a non-parametric form of ANOVA, was used to assess whether there is a significant difference between the outcomes on the Recognition scale for the three years. The Kruskal-Wallis Test can compare independent variables with more than two levels (i.e. data from three or more independent groups). It is a generalised form of the Mann-Whitney Test, which is a non-parametric version of the independent t-test. The results of the comparison performed here indicate a significant difference between the three years, with the highest L2 recognition scores occurring in 2009, followed by 2007 and then 2012. This test was repeated five times using different re-samples of the data. Although the exact results differed slightly each time, the substantive outcomes were comparable.

Table 4: Ranks estimated from the Kruskal-Wallis Test across three years

	Year	N	Mean Rank
L2 Recognition	2007	31,950	39,955.39
	2009	27,262	39,998.55
	2012	19,927	38,365.78
	Total	**79,139**	

Table 5: Test Statistics[a,b] from the Kruskal-Wallis Test

	L2 Recognition
Chi-Square	77.841
df	2
Asymp. Sig.	.000

a. Kruskal-Wallis Test
b. Grouping Variable: Year

Mann-Whitney tests were then carried out for each pair of years separately. These tests show that there is no significant difference between 2007 and 2009 Recognition scores, but that there is a significant difference between 2007 and 2012, and 2009 and 2012 Recognition scores, with the lower outcomes observed in 2012. As with the Kruskal-Wallis Test reported above, these individual comparisons were also replicated with several re-samples of the data. Output from the Mann-Whitney tests is shown in Tables 6–11, below.

Table 6: Ranks estimated from the Mann-Whitney Test comparing L2 Recognition between 2007 and 2009

	Year	N	Mean Rank	Sum of Ranks
Recog_Ord	2007	31,950	29,593.10	945,499,508.00
	2009	27,262	29,622.21	807,560,570.00
	Total	**59,212**		

Table 7: Test Statistics[a] from the Mann-Whitney Test comparing L2 Recognition between 2007 and 2009

	L2 Recognition
Mann-Whitney U	435,082,283.000
Z	-.212
Asymp. Sig. (2-tailed)	.832

a. Grouping Variable: Year

Table 8: Ranks estimated from the Mann-Whitney test comparing L2 Recognition between 2009 and 2012

	Year	N	Mean Rank	Sum of Ranks
Recog_Ord	2009	27,262	24,007.85	654,501,891.00
	2012	19,927	23,030.19	458,922,564.00
	Total	**47,189**		

Table 9: Test Statistics[a] from the Mann-Whitney Test comparing L2 Recognition between 2009 and 2012

	L2 Recognition
Mann-Whitney U	260,369,936.000
Z	-7.896
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Year

Table 10: Ranks estimated from the Mann-Whitney test comparing L2 Recognition between 2007 and 20012

	Year	N	Mean Rank	Sum of Ranks
Recog_Ord	2007	31,950	26,337.79	841,492,521.00
	2012	19,927	25,299.59	504,144,982.00
	Total	**51,877**		

Table 11: Test Statistics from the Mann-Whitney Test comparing L2 Recognition between 2007 and 2012

	L2 Recognition
Mann-Whitney U	305,592,354.000
Z	-7.882
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Year

Research questions 2 and 3: methodology and results

The statistical modelling approach was employed to answer research questions 2 and 3. Before describing the models employed and the results of the analysis, it will be useful to give an overview of the data, presenting this in terms of the outcome variable (L2 English reading, here operationalised as performance on the English Recognition scale) and the explanatory variables (Age of Child and L1 Reading). Only the data for 2012 was used in the investigation of research questions 2 and 3.

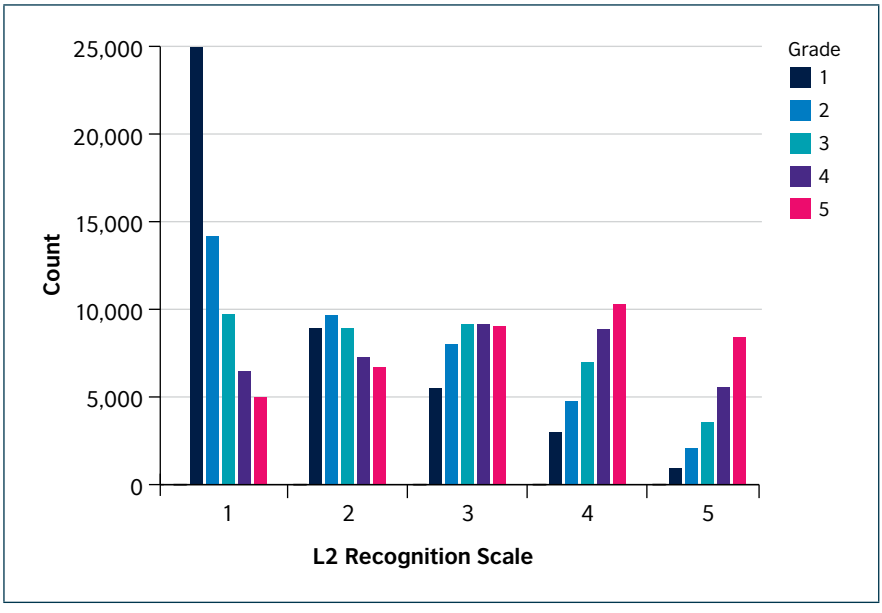
Outcome variable: L2 Recognition

Children in all grade levels displayed L2 Recognition skills at all levels on the scale, although there is a trend towards higher performances among children in the higher school grades. Frequencies are given in Table 12, while the clustered bar chart shown in Figure 2 illustrates the pattern of performances on the L2 Recognition Scale across grades.

Table 12: L2 Recognition Scale outcomes for different grade levels

		Grade					Total
		1	2	3	4	5	
L2 Recognition Scale	**0**	24,952	14,187	9,716	6,457	4,991	**60,303**
	1	8,906	9,675	8,935	7,254	6,691	**41,461**
	2	5,532	8,036	9,140	9,168	9,046	**40,922**
	3	3,013	4,769	6,965	8,870	10,320	**33,937**
	4	915	2,110	3,588	5,587	8,429	**20,629**
Total		**43,318**	**38,777**	**38,344**	**37,336**	**39,477**	**197,252**

Figure 2: Outcomes on L2 Recognition Scale for each grade



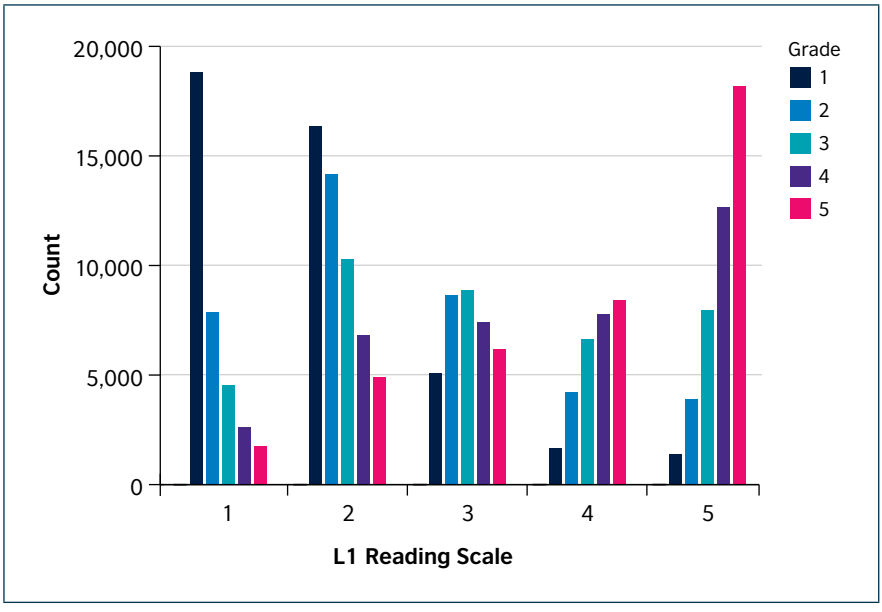
L1 Reading by grade level

L1 Reading shows a very similar pattern across the grade levels to that of L2 Recognition; however, there is a higher peak in the achievement among the children in grade 5 for L1 Literacy than for L2 Recognition.

Table 13: L1 Reading Scale outcomes for different grade levels

		Grade					Total
		1	2	3	4	5	
L1 Reading Scale	0	18,821	7,859	4,543	2,623	1,763	**35,609**
	1	16,378	14,158	10,293	6,844	4,920	**52,593**
	2	5,071	8,661	8,872	7,402	6,177	**36,183**
	3	1,676	4,199	6,650	7,785	8,425	**28,735**
	4	1,372	3,900	7,986	12,682	18,192	**44,132**
Total		**43,318**	**38,777**	**38,344**	**37,336**	**39,477**	**197,252**

Figure 3: Outcomes on the L1 Reading Scale for each school grade



The rank correlations between these two variables show that there is a strong positive relationship between them (r_s= .622 - .693); please see Table 14 for details.

Table 14: Correlations (Spearman's Rank) between L1 Reading and L2 Recognition at each grade level

Grade		Statistic
1	Correlation Coefficient	.693**
	Sig. (2-tailed)	.000
	N	43,318
2	Correlation Coefficient	.679**
	Sig. (2-tailed)	.000
	N	38,777
3	Correlation Coefficient	.684**
	Sig. (2-tailed)	.000
	N	38,344
4	Correlation Coefficient	.657**
	Sig. (2-tailed)	.000
	N	37,336
5	Correlation Coefficient	.622**
	Sig. (2-tailed)	.000
	N	39,477

**significant at 5% level

A cross-tabulation of the variables, meanwhile, shows that despite the association between these two variables there are children who perform at all combinations of the two levels across all grades (the lowest cell count is 18 for children in grade 1, scoring 4 on the L2 Recognition scale and 0 on the L1 Reading scale).

Table 15: L2 Recognition Scale cross-tabulation with L1 Reading

school_class			L1 Reading					Total
			0	1	2	3	4	
1	L2 Recognition Scale	0	17,537	6,487	784	100	44	**24,952**
		1	889	6,326	1,330	249	112	**8,906**
		2	294	2,846	1,636	502	254	**5,532**
		3	81	635	1,199	641	457	**3,013**
		4	20	84	122	184	505	**915**
	Total		**18,821**	**16,378**	**5,071**	**1,676**	**1,372**	**43,318**
2	L2 Recognition Scale	0	7,016	5,522	1221	293	135	**14,187**
		1	568	5,473	2,565	741	328	**9,675**
		2	199	2,573	3,041	1,386	837	**8,036**
		3	58	503	1,535	1,375	1,298	**4,769**
		4	18	87	299	404	1,302	**2,110**
	Total		**7,859**	**14,158**	**8,661**	**4,199**	**3,900**	**38,777**
3	L2 Recognition Scale	0	4,011	4,003	1,155	359	188	**9,716**
		1	311	4,072	2,736	1,113	703	**8,935**
		2	133	1,775	3,134	2,238	1,860	**9,140**
		3	51	357	1,566	2,304	2,687	**6,965**
		4	37	86	281	636	2,548	**3,588**
	Total		**4,543**	**10,293**	**8,872**	**6,650**	**7,986**	**38,344**
4	L2 Recognition Scale	0	2,288	2,693	891	336	249	**6,457**
		1	183	2,594	2,210	1,274	993	**7,254**
		2	92	1,216	2,629	2,559	2,672	**9,168**
		3	35	271	1,391	2,847	4,326	**8,870**
		4	25	70	281	769	4,442	**5,587**
	Total		**2,623**	**6,844**	**7,402**	**7,785**	**12,682**	**37,336**
5	L2 Recognition Scale	0	1,520	1,925	777	432	337	**4,991**
		1	131	1,881	1,896	1,386	1,397	**6,691**
		2	67	855	2,123	2,710	3,291	**9,046**
		3	26	212	1,132	3,047	5,903	**10,320**
		4	19	47	249	850	7,264	**8,429**
	Total		**1,763**	**4,920**	**6,177**	**8,425**	**18,192**	**39,477**

Age

There is considerable variation in the age of the children studying in each grade level. Although each grade level includes children across the entire range of ages contained in the data set (five to 16 years), the mean age increases for each successive grade level; please see Table 16, below. As expected, however, there is a concentration of numbers in each grade level around the age group for whom this year of schooling is intended. The situation is further clarified upon examining the pie charts presented in Figure 4 and the numbers in Table 17, in which the three ages that see the highest representation in each grade are highlighted.

Table 16: Summary statistics age at each grade level

Grade	N	Minimum	Maximum	Mean	Std. Deviation
1	43,305	5	16	6.33	1.348
2	38,755	5	16	7.54	1.447
3	38,325	5	16	8.52	1.418
4	37,315	5	16	9.50	1.447
5	39,451	5	16	10.50	1.370

Figure 4: Distribution of ages within each school grade

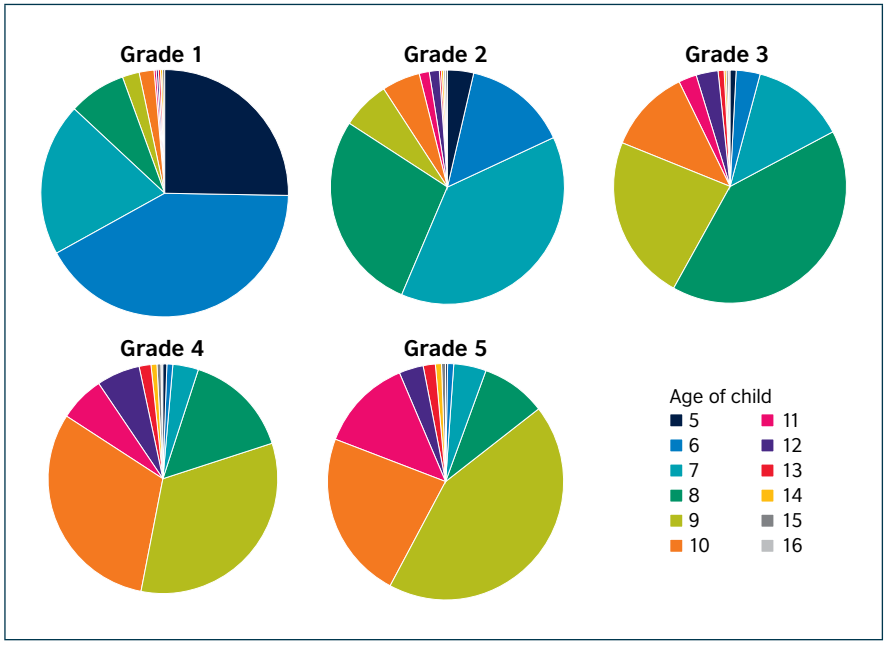


Table 17: Frequencies of children at each age studying in the different grade levels

		Grade					Total
		1	2	3	4	5	
Age of Child	5	**10,998**	1,496	368	223	70	**13,155**
	6	**18,086**	**5,586**	1,276	381	88	**25,417**
	7	**8,581**	**14,799**	**5,031**	1,348	353	**30,112**
	8	3,350	**10,841**	**15,651**	**5,571**	1,749	**37,162**
	9	924	2,565	**8,861**	**12,302**	3,547	**28,199**
	10	789	2,058	4,374	**11,582**	**17,081**	**35,884**
	11	157	518	1,047	2,478	**9,056**	**13,256**
	12	165	523	1,108	2,237	**4,988**	**9,021**
	13	61	145	318	627	1,317	**2,468**
	14	44	75	171	325	660	**1,275**
	15	66	63	87	164	357	**737**
	16	84	86	33	77	185	**465**
Total		**43,305**	**38,755**	**38,325**	**37,315**	**39,451**	**197,151**

The cross-tabulation in Appendix 1 shows that children of all ages studying in all grades demonstrated L2 Recognition levels across the range. The data do, however, become much sparser among the higher age groups. Again, this is to be expected since the grades covered are all primary school level.

Observed cumulative percentages

Since it is proposed that the ordinal outcomes may be modelled using a cumulative logit model (see section on modelling information, below, for details), it is useful to examine the observed cumulative percentages. It is a function of these that are modelled when taking this approach (Norušis, 2011).

Figure 5 illustrates that the older children tend to perform better on the L1 Recognition scale. The curves of the slopes alter as the age of the child increases, but they do not cross. By grade 2 (Figure 6) the relationship is not as clear cut, and some of the slopes cross.

Looking to each of the cumulative percentage plots in figures 7–9, the overall impression is that age does not have such a distinct relationship with each of the L2 Recognition levels as observed for grade 1.

On the other hand, the relationship between L1 Reading and L2 Recognition is surprisingly similar across the grade levels (see Figures 10–14).

Figure 5: Grade 1 L2 Recognition outcomes by Age

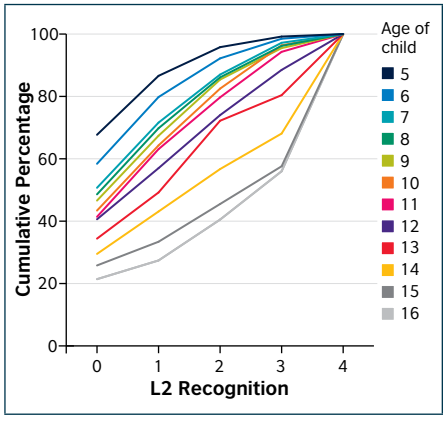


Figure 6: Grade 2 L2 Recognition outcomes by Age

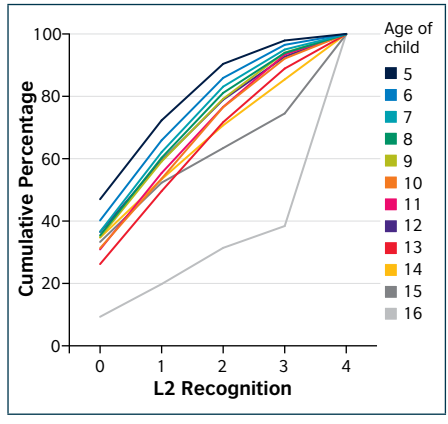


Figure 7: Grade 3 L2 Recognition outcomes by Age

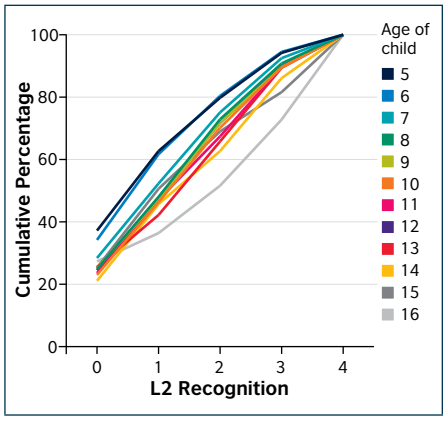


Figure 8: Grade 4 L2 Recognition outcomes by Age

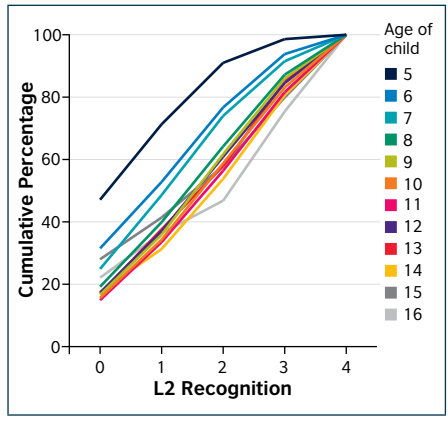


Figure 9: Grade 5 L2 Recognition outcomes by Age

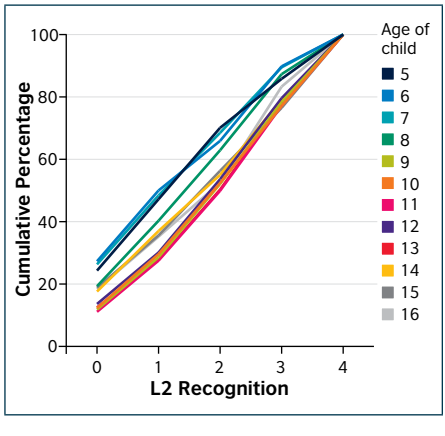


Figure 10: Grade 1 L2 Recognition outcomes by L1 Reading

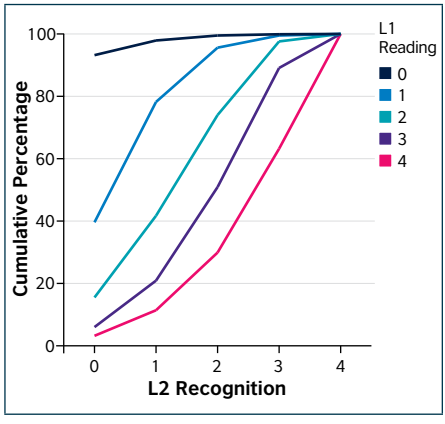


Figure 11: Grade 2 L2 Recognition outcomes by L1 Reading

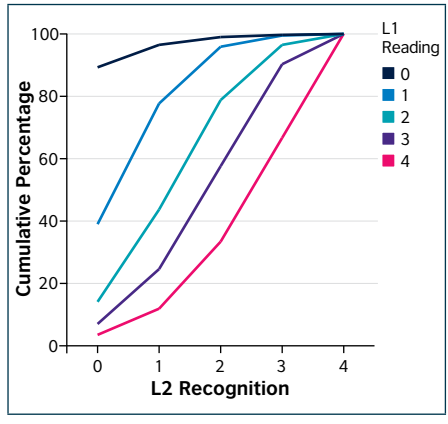


Figure 12: Grade 3 L2 Recognition outcomes by L1 Reading

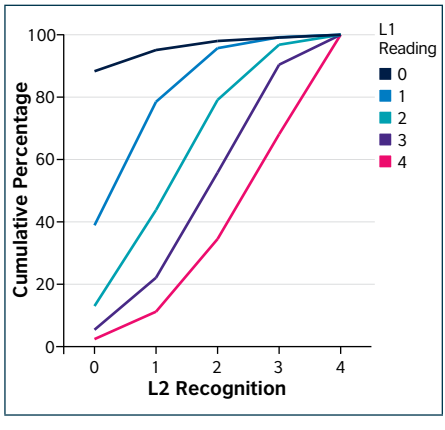


Figure 13: Grade 4 L2 Recognition outcomes by L1 Reading

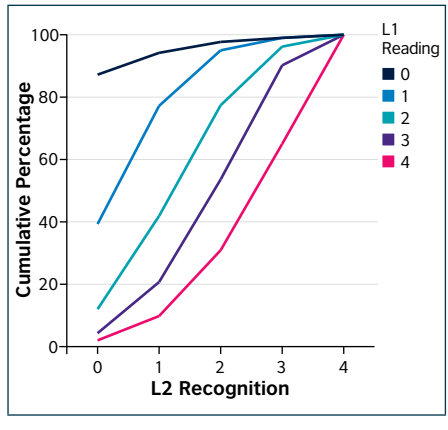
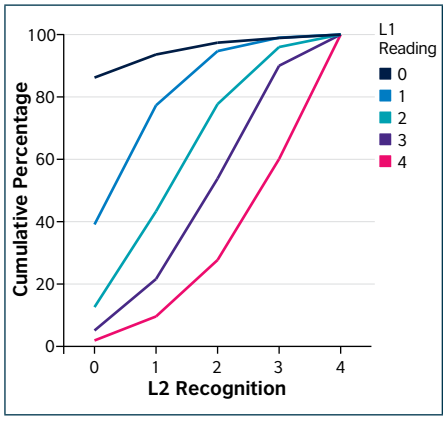


Figure 14: Grade 5 L2 Recognition outcomes by L1 Reading



Modelling information: description of models to be employed

The models employed in this analysis are from a family of generalised linear regression models (McCullagh and Nelder, 1989). These models are a statistical generalisation of the linear modelling process by which the outcomes of a given situation measured continuously (Y) are predicted by way of a series of covariates (X) that are accorded certain weights (β):

$$Y = X\beta + \varepsilon, \qquad \varepsilon \sim N(0,\sigma^2)$$

In this case Y is assumed to be normally distributed conditional on X.
The combination of covariates included in the model explains a certain portion of the variation in the outcome with a degree of error. In the case of the current data, however, the outcomes cannot be considered continuous. As noted in the section on data preparation, the information about L2 Recognition is recorded on an ordinal scale in which children are given the opportunity to demonstrate their L2 recognition knowledge in terms of a predefined scale.

In order to model these outcomes it is important to relax the assumption of normality. In a generalised linear model, outcomes are assumed to be realisations from an exponential family distribution with parameters driven by Xβ. Of interest here are models that predict dichotomous outcomes and – using an extension of this method – ordinal outcomes. Dichotomous (0;1) outcomes can be considered as following the Bernoulli distribution, for which the parameter π describes the probability of an outcome being equal to 1.

Employing an element in the model known as the 'link function' (η) to transform this parameter to a format that can be expressed as a value between - ∞ and + ∞ it is possible to model the expected outcomes using the same principles as described above for linear models. For Bernoulli data, a commonly used link function is the *log-odds ratio*, or *logit* function. This is the log of the odds that one outcome will be observed over the alternative outcome:

$$\eta = \ln \frac{\pi}{1-\pi}$$

Given its function in transforming the expected outcome to a manner that can be expressed in a linear format, η is often referred to as the 'linear predictor'. Once the relevant link function is established, the model can – as with a simple linear model – be expressed as a series of covariates with given weights:

$$\eta = X\beta$$

Weighted covariates (Xβ) are estimated to have a systematic relationship with the outcomes, as expressed via the linear predictor (η). The option exists to extend this concept to modelling ordinal data using cumulative logit regression, also known as the proportional odds model (McCullagh, 1980). Essentially, this approach models a series of ordered categorical outcomes by estimating the odds of observing a response in each successive category in terms of a series of thresholds. So, for

example, the first threshold would compare responses in the first category (Recognition=0) versus being in any of the higher categories (Recognition=1,2,3 and 4). This would model the threshold between having no L2 Recognition knowledge and having some knowledge. The second prediction will be for the odds of being in the lowest two categories (Recogntion=0 and Recognition=1) versus the others (Recognition=2,3 and 4), and so on. In other words, the effects are estimated to be cumulative. The odds modelled are:

prob (score of 0)/prob (score greater than 0)

prob (score of 0 or 1)/prob (score greater than 1)

prob (score of 0, 1, or 2)/prob (score greater than 2)

prob (score of 0, 1, 2, or 3)/prob (score greater than 3)

There will always be one fewer odds ratio modelled than there are categories in the outcome. The statistical principle upon which each of these odds is modelled is the same as that described above for dichotomous outcomes.

A key assumption in ordinal regression is that the effects of any explanatory variables are consistent or *proportional* across the different thresholds; hence this is usually termed the *assumption of proportional odds* (also known as the *assumption of parallel lines*). This assumes that the explanatory variables have the same effect on the odds regardless of the threshold. For example, if a set of separate binary logistic regressions were fitted to the data, a common odds ratio for an explanatory variable would be observed across all the regressions. In ordinal regression there will be separate intercept terms at each threshold, but a *single* odds ratio for the effect of each explanatory variable. This means that if the odds ratio varies considerably between the steps in the ordinal outcome then the model is not suited to the data.

Modelling assumptions

L2 Recognition is modelled as ordinal outcomes for all responses included in the 2012 data set and also for each individual grade level. Hypothesised explanatory variables are L1 Reading and Age of Child. In the initial instance these models are used to assess the appropriateness of modelling L2 Recognition under the assumption of proportional odds, as described above.

Looking to the observed cumulative frequencies displayed above suggests that this is not an untenable possibility; however, having run the statistical models it is possible to formally test this using the test of parallel lines. The lines referred to here are the regression slopes associated with each estimated threshold. The test compares the -2 log likelihood for the null model where the lines are constrained to the parallel to a 'general' model in which each of the slopes are estimated separately. If this general model represents an improvement on the null a significant result is recorded, indicating a departure from the assumption of proportional odds.

Table 18 shows the results of this test for the ordinal regression model run for the full 2012 data set, while Table 19 shows the results for models run for each of the grade levels individually.

Table 18: Test of Parallel Lines[a] for ordinal regression of L2 Recognition on Age of Child and L1 Reading – full 2012 data set

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	4,764.931			
General	2,332.107	2,432.824	15	.000

The null hypothesis states that the location parameters (slope coefficients) are the same across response categories

a. Link function: Logit

Table 19: Test of Parallel Lines[a] run for each grade level

School Grade	Model	-2 Log Likelihood	Chi-Square	df	Sig.
1	**Null Hypothesis**	1,552.699			
	General	1,021.521	531.178	45	.000
2	**Null Hypothesis**	1,617.942			
	General	1,173.779	444.163	45	.000
3	**Null Hypothesis**	1,831.139			
	General	.000[b]	1,831.139	45	.000
4	**Null Hypothesis**	1,837.187			
	General	.000[b]	1,837.187	45	.000
5	**Null Hypothesis**	1,837.468			
	General	.000[b]	1,837.468	45	.000

The null hypothesis states that the location parameters (slope coefficients) are the same across response categories

a. Link function: Logit

b. The log-likelihood value is practically zero. There may be a complete separation in the data. The maximum likelihood estimates do not exist.

The test for parallel lines gives a significant reading across all of the models and a warning indicates that problems have occurred with the convergence of the models for grades 4 and 5 (see footnote 'b' below the table). The first thing to note with regard to the significant findings is that this test is sensitive to the sample size and, as such, is more likely to find small differences significant in a larger data set – something which would certainly affect our model, since even the individual grade levels contain approximately 40,000 cases in each group. Additionally, models including continuous covariates are more susceptible to violating assumptions. Re-running the test of parallel lines for models incorporating L1 Reading and Age of Child as individual explanatory variables shows, however, that both variables contribute to the significance of these outcomes.

While it has been observed that in general this test is found to reject the null hypothesis more often than it accepts it (see, for example, Harrell, 2001: 335), the significant findings described above should not be ignored, especially in light of the warnings for grade 4 and 5 models. To overlook this issue may be to miss important substantive insights from the data. In this situation it is useful to explore the relationship between L2 Recognition and the explanatory variables further. By modelling the odds of achieving a score greater than each of the cut-points individually using a series of binary regression models, it is possible to gain an insight into where the violations of the proportional odds assumption exist; see Brant (1990) for more details. As described above, there are four points on the L2 Recognition ordinal scale at which the odds ratio is applied in an ordinal regression model. These are listed in Table 20.

Table 20: Successive odds ratios modelled in ordinal regression model

Notation	Response outcome modelled	Reference category
y>0	L2 Recognition levels (1,2,3,4)	No L2 Recognition
y>1	L2 Recognition levels (2,3,4)	L2 Recognition levels (0,1)
y>2	L2 Recognition levels (3,4)	L2 Recognition levels (0,1,2)
y>3	L2 Recognition levels (4)	L2 Recognition levels (0,1,2,3)

The series of binary regression models revealed that there are indeed distinctions in the estimated odds ratios at the different cut-points in the response data for both the individual grade levels and the 2012 data modelled together. The full sets of estimates are given below. An example from the modelling of the grade 1 data shows that the log-odds estimate for the age of child variable is .075 (odds-ratio = 1.078) when the first cut-point (y>0) is modelled, while it is twice this at .150 (odds-ratio = 1.162) when the last cut-point (y>3) is modelled. The two interim estimates are between this. This pattern is repeated for the age estimates at the other grades, with a number of the estimates at the first cut-point showing as not significantly different to zero. A similar progression in the magnitude of the odds estimates is seen for the L1 Reading variable. While something of the detail and the implications of these estimates are explored in more detail below, the key point to note is that both variables display non-proportionality in the odds ratios. And, in fact, examination of the estimates themselves also raises some interesting points regarding the relationship between L2 Recognition and the explanatory variables – L1 Reading and Age – which would be overlooked had the analysis proceeded under the assumption of proportionality.

The decision was thus taken to model outcomes as a series of binary regressions. Although this approach involves creating a series of models for each grade level – an exercise that is less parsimonious than modelling the outcomes simultaneously using an ordinal regression model – it allows an in-depth exploration of the data in a manner that has greater statistical tractability. As an exploratory exercise, the focus is on understanding as much as possible about the nuances in relationships between the variables. The models created during this exercise are presented below since they will form the basis of the major substantive findings.

Binary regression models for each individual grade level (2012 data)

As described above, binary regression models were run for each individual grade level. These are considered to be the best means of investigating the relationships of interest in this exploratory analysis, namely between the L2 Recognition of the participants, their L1 Reading and their age.

Before looking to the nature of the estimates themselves, the first point to note is the pattern of significance in the hypothesised explanatory variables (to recap, Age was included in the models as a continuous covariate and L1 Reading as an ordered categorical variable). Significance information from the Wald (Type III) test, which tests the null hypothesis that a parameter is 0, is presented in Table 21. Each block of three significance readings (for intercept, Age and L1 Reading respectively) for each grade level (1–5) represents information from a separate binary regression model. It can be seen from this overview that age does not play a significant role at the five per cent level in estimating L2 Recognition above zero (y>0) or above lower case letter recognition (y>1) among the higher grades.

Table 21: Tests of Model Effects for all grade level binary regression models

Grade	Source	Type III			
		y>0 p-value	y>1 p-value	y>2 p-value	y>3 p-value
1	(Intercept)	.000	.000	.000	.000
	Age	.000	.000	.000	.000
	L1 Reading	.000	.000	.000	.000
2	(Intercept)	.000	.000	.000	.000
	Age	**.875**	.026	.000	.000
	L1 Reading	.000	.000	.000	.000
3	(Intercept)	.000	.000	.000	.000
	Age	**.314**	.008	.000	.000
	L1 Reading	.000	.000	.000	.000
4	(Intercept)	.000	.081	.000	.000
	Age	**.403**	**.128**	.000	.000
	L1 Reading	.000	.000	.000	.000
5	(Intercept)	.000	.045	.000	.000
	Age	**.487**	**.065**	.002	.010
	L1 Reading	.000	.000	.000	.000

Parameter estimates for the series of binary regression models are presented in Tables 22–26, below. Each column of information represents the estimates derived from a distinct binary regression model for the grade in question. Both log-odds (B) and odds (Exp[B]) are shown for each model. The outcomes modelled in each successive model are distinguished by the cut-point on the L2 Recognition ordinal scale as per the information in Table 20, above. The estimates indicate the relationship between the relevant explanatory variable (Age or L1 Reading) and the odds of observing the higher outcome at the given cut-point. So, for example, the first column on each table (y>0) provides an estimate of the odds of a child performing at any point on the L2 Recognition scale above 0 (no L2 recognition knowledge). The second column (y>1) provides an estimate of the odds of a child performing above 1 (recognition of L2 upper case letters), and so on. All estimates shown are significant at the five per cent level (in most cases significant at the one per cent level). Where an estimate is non-significant this is marked 'NS'.

A clear pattern emerges in the estimates. Firstly, for each grade level, the age of the child has an increasing effect on the odds of demonstrating L2 Recognition knowledge at each higher threshold, i.e. the estimate for the first threshold (y>0) is in each case lowest (in the case of the higher grades not significantly different to zero), and highest for the final threshold (y>3). This finding would not have been allowed for using the cumulative logit model, for which all these odds ratios would have been equal. Secondly, the effects for L1 Reading show a stepped increase in the impact of each higher level on the odds of demonstrating L2 Recognition. This pattern holds for each threshold model, confirming a strong relationship between L1 Reading scores and the L2 Recognition scale.

Table 22: Grade 1 binary regression model estimates

Parameter	y>0		y>1		y>2		y>3	
	B	Exp(B)	B	Exp(B)	B	Exp(B)	B	Exp(B)
(Intercept)	-3.077		-4.644		-5.991		-7.781	
Age of Child	.075	1.078	.129	1.138	.124	1.132	.150	1.162
[L1 Reading=4]	5.921	372.784	5.740	311.064	5.909	368.338	6.063	429.663
[L1 Reading=3]	5.330	206.438	5.113	166.168	5.119	167.168	4.650	104.585
[L1 Reading=2]	4.288	72.821	4.138	62.677	4.132	62.302	3.077	21.693
[L1 Reading=1]	3.025	20.594	2.541	12.692	2.113	8.273	1.544	4.683
[L1 Reading=0][a]	0	1	0	1	0	1	0	1

[a] This is the reference category for L1 Reading

Table 23: Grade 2 binary regression model estimates

Parameter	y>0		y>1		y>2		y>3	
	B	Exp(B)	B	Exp(B)	B	Exp(B)	B	Exp(B)
(Intercept)	-2.108		-3.467		-5.096		-6.918	
Age of Child	NS	NS	.020	1.020	.063	1.065	.112	1.119
[L1 Reading=4]	5.448	232.293	5.308	201.946	5.288	197.947	5.322	204.793
[L1 Reading=3]	4.710	111.052	4.429	83.848	4.306	74.143	3.807	45.015
[L1 Reading=2]	3.928	50.805	3.566	35.375	3.304	27.221	2.728	15.302
[L1 Reading=1]	2.567	13.027	2.068	7.909	1.487	4.424	.980	2.664
[L1 Reading=0][a]	0	1	0	1	0	1	0	1

[a] This is the reference category for L1 Reading

Table 24: Grade 3 binary regression model estimates

Parameter	y>0		y>1		y>2		y>3	
	B	Exp(B)	B	Exp(B)	B	Exp(B)	B	Exp(B)
(Intercept)	-2.109		-3.178		-4.259		-5.324	
Age of Child	NS	NS	.024	1.024	.040	1.040	.062	1.064
[L1 Reading=4]	5.747	313.250	5.045	155.244	4.561	95.679	4.031	56.317
[L1 Reading=3]	4.884	132.158	4.228	68.580	3.686	39.885	2.546	12.756
[L1 Reading=2]	3.917	50.249	3.217	24.953	2.583	13.237	1.374	3.951
[L1 Reading=1]	2.471	11.834	1.680	5.366	.820	2.271	NS	NS
[L1 Reading=0][a]	0	1	0	1	0	1	0	1

[a] This is the reference category for L1 Reading

Table 25: Grade 4 binary regression model estimates

Parameter	y>0		y>1		y>2		y>3	
	B	Exp(B)	B	Exp(B)	B	Exp(B)	B	Exp(B)
(Intercept)	-1.833		-2.920		-4.212		-5.128	
Age of Child	NS	NS	NS	NS	.049	1.050	.052	1.053
[L1 Reading=4]	5.835	342.065	5.004	149.008	4.548	94.443	4.011	55.202
[L1 Reading=3]	5.025	152.170	4.129	62.116	3.599	36.562	2.417	11.212
[L1 Reading=2]	3.912	49.999	3.112	22.466	2.512	12.330	1.400	4.055
[L1 Reading=1]	2.356	10.549	1.564	4.778	.800	2.226	NS	NS
[L1 Reading=0][a]	0	1	0	1	0	1	0	1

[a] This is the reference category for L1 Reading

Table 26: Grade 5 binary regression model estimates

Parameter	y>0		y>1		y>2		y>3	
	B	Exp(B)	B	Exp(B)	B	Exp(B)	B	Exp(B)
(Intercept)	-1.741		-2.878		-3.938		-4.801	
Age of Child	NS	NS	NS	NS	.028	1.028	.027	1.027
[L1 Reading=4]	5.807	332.620	4.940	139.770	4.603	99.783	4.108	60.825
[L1 Reading=3]	4.754	116.048	3.979	53.464	3.490	32.786	2.331	10.288
[L1 Reading=2]	3.772	43.467	2.960	19.298	2.395	10.968	1.349	3.854
[L1 Reading=1]	2.276	9.738	1.460	4.306	.747	2.111	NS	NS
[L1 Reading=0][a]	0	1	0	1	0	1	0	1

[a] This is the reference category for L1 Reading

Since the Age of Child is a continuous variable, the estimates associated with this covariate indicate the impact on the odds for each increase of one year in a child's age. Where significant, the impact on the odds ratio is positive, indicating that an increase in age increases the odds of achieving in the higher category in each model. The largest estimated effect of age is on the odds of a child in grade 1 gaining the highest level – sentence recognition – in the L2 Recognition task. Looking to the grade 1 exp(B) estimates in Table 22 for the final threshold (y>3), the odds ratio is 1.16. This means that for each year older a child is (L1 Reading held constant) the odds that he or she will be able to recognise L2 sentences are increased by a factor of 1.16. So, for example, the odds of a child aged eight are slightly over one and a half times that of a child aged five of achieving this level of L2 Recognition.

The L1 Reading variable, on the other hand, as an ordered categorical variable, shows five estimates. These estimates indicate the impact on the odds of children at each successive level on the L1 Reading scale performing at the higher level of L2 Recognition. The baseline group is L1 Reading=0. So, for example, looking to the grade 1 exp(B) estimates in Table 22 for the first model, a child whose L1 Reading score is 2 will have more than 72 times the odds of demonstrating some degree of L2 Recognition than a child whose L1 Reading score is 0. Meanwhile, a child with an L1 Reading score of 4 has over 372 times the odds of scoring in the higher L2 Recognition category than a child with an L1 Reading score of 0. On the face of it these estimates seem surprisingly large; however, if the difference between not being able to read even a word in your L1 and being able to read a word (L1 Reading=2) or a paragraph (L1 Reading=4) is considered, these large increases in odds are less surprising.

Model fit

A number of statistics are employed to consider the fit of the binary regression models described above:

Likelihood ratio test

A likelihood ratio test assesses whether the predictions made in the model are better than those that could be made based purely on the marginal probabilities. A significant value indicates that the model represents a significant improvement over the baseline intercept-only model (or thresholds model). A significant value ($p<.000$) is found for each of the models across each of the five grades. This indicates that the information contained in the explanatory variables is adding to our understanding of the patterning of L2 Recognition in the model.

Pearson's chi-square statistic

This statistic provides an indication of whether the data are consistent with the fitted model; in other words, whether the predicted cell counts are significantly different to the observed cell counts. The null hypothesis is that there is no significant difference between the two. The majority of models fitted here had a significant p-value; the exceptions were the grade 5 model for $y>2$ ($x^2= 53.077$ on 54 d.f.; $p=.510$) and the grade 1 model for $y>3$ ($x^2= 71.942$ on 54 d.f.; $p=.052$). With the exception of these two models, the estimated outcomes are therefore considered to vary significantly from the observed data. It is unsurprising to find a significant value for this statistic in the current modelling exercise; with large sample sizes this test statistic is very sensitive to any variations in the observed and predicted counts. Additionally, these test statistics can be highly sensitive to empty cells in the observed data, which is often the case in models that employ variables with continuous covariates – as per the age covariate in the current models. Empty cells are shown to account for up to just under 20 per cent of the observed data in the grade-level groupings, with the highest percentage of empty cells being seen in the $y>3$ models.

Table 27: Cells with zero frequencies in each model

	Number of cells[a] with zero frequencies (%)			
Grade	**y>0 models**	**y>1 models**	**y>2 models**	**y>3 models**
1	14 (11.7%)	8 (6.7%)	11 (9.2%)	22 (18.3%)
2	8 (6.7%)	4 (3.3%)	10 (8.3%)	20 (16.7%)
3	6 (5.0%)	3 (2.5%)	4 (3.3%)	13 (10.8%)
4	5 (4.2%)	0 (0.0%)	5 (4.2%)	16 (13.3%)
5	3 (2.5%)	3 (2.5%)	7 (5.8%)	12 (10.0%)

[a] i.e. dependent variable levels by combinations of predictor variable values

In fact, it is perhaps surprising to even find that two of the models yield a non-significant statistic here, since the expectation in modelling the data using the limited number of explanatory variables was not to account perfectly for the observed outcomes but to establish a broad idea of some of the main trends and relationships between the variables. This is an important point to note in considering the role of fit statistics in this context.

Pseudo-R^2 statistics

In modelling observed data, the aim is to express observed responses in terms of the underlying pattern of relationships. In ordinary least squares (OLS, or linear) regression modelling, the R^2 statistic indicates the amount of variability in observed responses explained by the model. In the case of logistic regression (of which binary regression is a type), such a statistic is not derived in quite such a straightforward fashion. In order to replicate the usefulness of the R^2 statistic, a number of 'pseudo' R^2 statistics have been derived. It is to be cautioned that these pseudo R^2 statistics are, however, not to be interpreted as direct equivalents to R^2 statistics in OLS regression (Cohen et al., 2003:503). Unfortunately there is no generally accepted consensus as to the best statistic to employ. The current analysis reports Nagelkerke's statistic, since this has the virtue of being reported on a scale between 0–1, with values closer to 1 representing a closer fit with the data. As with the information contained in the Pearson's statistic reported above, this information indicates where the model makes improvement on the null model. The calculation of this statistic is analogous to the OLS R^2 statistic in that it is based on the ratio of likelihoods to reflect the improvement of the full model over the intercept-only model.

The Nagelkerke statistics for the models reported here range from .539 to .289. It can be seen from the information in Table 28 that using this criterion, the better performing models are those that model data for the lowest cut-point on the L2 Recognition scale. The reduction in the statistic for each progressive grade level indicates that the relationship between dependent and explanatory variables lessens (slightly) for the higher grade levels. Overall, however, these figures indicate that the combination of explanatory variables make a reasonable contribution towards explaining L2 Recognition at each of the cut-points.

Table 28: Nagelkerke's Pseudo R^2 statistics

Grade	y>0 models	y>1 models	y>2 models	y>3 models
1	.539	.451	.313	.415
2	.469	.432	.403	.344
3	.476	.452	.392	.300
4	.474	.439	.380	.286
5	.440	.404	.364	.289

Discussion

Descriptive statistics for English reading performance show a slight drop in the percentage of students scoring at the top of the scale in 2012. Given the ordinal nature of the data, non-parametric tests were used to test if the differences were statistically significant. Results indicated that there was no statistically significant difference in performance on the English reading recognition scale between 2009 and 2007, but that the differences between 2012 and 2007 and 2012 and 2009 were significant. These findings should be considered with extreme caution. As noted above, tests of significance are sensitive to sample size, and the sample sizes used in this study are likely to derive significant results from even minor differences. The results at this stage should be considered more as an example of what procedures could be employed to investigate apparent differences in the descriptive statistics across years. Other potential alternatives for investigating changes in ability over time are considered in the recommendations below.

The relationship between L1 literacy and L2 English reading

A significant positive relationship indicates that as levels of L1 Reading proficiency increase, so do levels of L2 Recognition. This is shown by the incremental increase in odds ratios with each higher level of L1 Reading. For example, looking to the exp(B) estimates in Table 26, it can be seen that – assuming Age is held constant – the odds of children performing at L1 Reading level 1 having a non-zero L2 Recognition score are nine times higher than children scoring at L1 Reading level 0. For children scoring at L1 Reading level 2, the increase in odds is 43 times; for those scoring at level 3 it is 116 times; and for those at level 4, the odds of their scoring more than zero on the L2 Recognition scale is over 332 times greater than those children at level zero.

This stepped increase in the odds is in line with expectation, given the established relationship in the literature between L1 literacy and L2 reading. This is by no means presented in this context to indicate a causal relationship. These findings can be viewed as the first step in building a more sophisticated insight into this relationship, which may well involve bringing in other background or explanatory variables, perhaps within a multi-level modelling context. Please see the Recommendations section for further discussion of this point.

One factor that must be taken into account when moving forward with investigations into the relationship between L1 and L2 literacy is the need to differentiate the odds ratios across the model. Achievements at the lower levels of the L1 Reading scale do not tell us so much about the variations in higher levels of the L2 Recognition (especially at y>2 and y>3). This pattern is clear across all grade-level models, but particularly those for the children in grades 4 and 5. In hindsight, it is perhaps to have been expected that this non-proportionality in the odds ratios would exist. In fact, it was this phenomenon that contributed towards the disruption of the proportional odds requirement made by the ordinal regression models, discussed at length above. Despite the promise of a more statistically parsimonious explanation of the outcomes from cumulative logit models, it was found not to be useful to assume proportional odds in estimating the children's tendencies to achieve different levels of L2 Recognition. To have done so would have been to mask this finding. The approach chosen here to accommodate for this non-proportionality in the odds estimates at the different cut-points in the data represents a useful means of extrapolating the differences in the estimates. However, there is potential in future research to employ a more streamlined model that does not assume proportionality (see, for example, Williams, 2006).

The relationship between age and reading within grades

This research question concerns the relationship between the Age covariate and the L2 Recognition outcomes in the models.

Where age plays a significant role in the binary regression models described above, this shows that, having already taken into account the strong relationship between L1 Reading levels and L2 Recognition, there is an additional effect which shows that the older the child the higher the odds of achieving the higher level at each of the successive cut-points. The exceptions to this are the cut-points between zero L2 Recognition and upper case letter recognition (y>0), and between upper case and lower case letter recognition (y>1). The age of the child did not play a significant role in describing the odds of scoring in these categories for the children in higher grades (please refer back to the information shown in Table 21).

This finding is of particular interest because it shows that children's L2 performance can be differentiated by their age even when they have the same level of L1 literacy and are studying at the same grade level. The effect sizes are not huge, but the consistency of these findings suggests that it is certainly an observation that would merit further investigation in future analyses.

A final point to note here is that these estimates must be understood in light of the fact that the majority of information within each grade level is from children either at the typical age for that grade, or one or two years above/below that (see the descriptive statistics). While there has been enough information from children at different ages within each grade to produce stable estimates from the data, it is acknowledged that the children who fall outside of the general age range for a

grade will potentially represent situations in which special educational circumstances apply. While these models have demonstrated a significant effect of the age of a child within a given grade level, a more detailed account would address some of these additional factors. Once again, these findings suggest no more than that the effect for age is worth pursuing. Certainly, if there is an effect for age, then this may have pedagogical implications for the materials and procedures used for children of different ages. But, given the enormous difficulties already faced by many of the rural schools that are the focus of the survey, any such suggestions would need to be weighed carefully within the wider context of setting priorities for improving materials and facilities along with teacher training.

Recommendations

This exploratory study has indicated that L1 reading ability and age do have an impact on L2 reading scores. The regression analyses could be replicated for 2007 and 2009 to check if the results hold across data sets. The analyses could also be replicated for the upper primary grades (grades 6–10) in the data. It would also be possible to consider introducing extra variables, such as L1 language, government or private school, etc. to examine the general trends in finer detail. Language distance (that is differences between the L1 and English) could indeed impact on the effect of L1 literacy, and the type of school in which children learn could be associated with differential access to facilities and resources or to different approaches to language education. It is, however, likely that these variables also interact in complex ways with other background and context variables. One way of teasing out such interactions may be to examine some of the background information collected in the survey to derive socio-economic indicators for inclusion in the regression analysis.

As noted in the rationale for research question 2, Sparks et al. (2011) investigated the impact of print exposure in addition to L1 reading performance and found that print exposure in the L1 is a significant predictor of L2 proficiency. As the current data contain some questions on the presence of newspapers and reading materials in the home, it may be useful to incorporate these in the regression models as indicators of print exposure. Data on the children already includes an item on whether they attend the government school which was visited separately for the school survey (see Pratham, 2012, for a description of how information on schools is also collected as a part of the survey). Linking the data could provide a way of investigating how many learners were choosing an L1 for the literacy test that differed from the language they were being schooled in (language of schooling is recorded at the school level in the school survey).

As observed in the main body of the report, age is a complicated variable here. The models reported above employ age as a continuous variable. During the course of analysis some explorations into different means of coding this variable (for example, as a binary variable) were considered. None of these were ultimately beneficial to the model. However, there is an argument to suggest that students who fall outside of the usual range of ages for a grade level should be accounted for differentially, since – more likely than not – they represent distinct educational circumstances.

Future analysis would also benefit from using a weighted resample from the data based on the number of districts in each state and the population numbers in each district, in order to provide a representative picture across rural India at a national level. This would reduce the N-size, but would provide opportunities to apply some more complex statistical approaches. Multi-level models, such as those that can be constructed using Generalised Linear Mixed Modelling approaches, can account for observations that are hypothesised to be clustered in groups (McCulloch and Searle, 2001). In the educational field, this has been commonly used to refer to homogeneity in the performances of students in the same class, or school, because of their shared educational experiences (see, for example, Goldstein, 1995). In the current circumstances it may be useful to model shared variation at the household level; in this case measures of, for example, parental literacy or access to media in the home could be modelled for their relationship with the L1 and L2 literacy levels of children in the study. It would be equally possible to incorporate school-level factors in much the same vein.

Given the processing and software limitations of the current project it was not possible to fully explore the options for relaxing the assumption of proportional odds while modelling L2 Recognition as an ordinal outcome. A number of models exist to achieve this; it would be a matter of assessing which is the most suitable for the data (see, for example, Williams, 2006). Such an approach would represent a more parsimonious approach from a statistical perspective; however, the series of binary regression models presented here provide adequate information about the patterns in the data required by the exploratory approach taken for this study.

Potential modifications to the data collection tools

Consideration could be given to ways of combining the English reading recognition and comprehension scales. This could be achieved with minimal changes to the tools used or the testing procedures in order to derive a single ordinal scale. Making such modifications would of course impact on comparability with previous years, but would facilitate more powerful analyses of the data, and particularly the relationship between L1 literacy and L2 proficiency.

More wide-ranging modifications

If a longer-term view of the survey is taken, consideration could be given now to planning new testing tools and scoring procedures, which would derive continuous data on an interval scale to maximise statistical information and facilitate more powerful statistical analyses. Any revision of the testing materials should also take into consideration the potential to extend the testing of L2 English beyond reading. Simple listening or speaking tasks could be devised and administered according to the same basic principles and procedures currently employed, which prioritise efficiency and enable the reach and access of the survey by utilising volunteers in rural districts to administer the testing tools. If there is concern regarding the ability of all survey volunteers to adequately administer and assess such tests, then it may be possible to administer the extra English tests to only a smaller sub-sample. For this sub-sample, survey volunteers could be recruited, for example, from university students with sufficient English ability, and extra training for the

purpose could be provided. A more balanced focus on all four skills is encouraged in the National Position Paper on the Teaching of English (NCERT, 2005). While recognising that the study in its present form takes account of the reality on the ground, in which reading is still the dominant form of classroom exposure to the target language for many learners, given the high profile of the Annual Status of Education Report, there is real potential for positive washback if the survey promoted a broader view of English proficiency with its testing instruments and reports.

This study has made some exploratory attempts to demonstrate techniques that could be used with the data as it is currently collected. Some of these techniques were used to investigate whether apparent differences in performance over the years are in fact substantive. Other methods of investigating this question are worth consideration. From the perspective of testing and assessment research, much could be gained by undertaking an anchoring study to link the assessment tools used across different years of the survey. Such a study could investigate the possibility of using methodology such as in Saida and Hattori (2008) to estimate difficulty parameters using Rasch or other Item Response Theory models for items on test forms administered in previous years. Equating test forms across years using Rasch, and anchoring the item parameter estimates on a common scale, would allow the re-analysis of item response data across years, thus allowing ability estimates for learners to be placed on the same common scale. This would facilitate investigation of the question of substantive changes in children's L2 reading ability across years.

Conclusion

The Annual Status of Education Report is a dynamic and important initiative, not just for the insights that the survey results can provide on educational issues, but also because of the process of the survey itself. The survey embraces citizen involvement and the promotion of inquiry and transparency among ordinary citizens about an issue central to their lives: education. The practical constraints that accompany the goals and scale of the survey are very real. In this study, we have taken the opportunity to investigate the impact of some of those constraints on the data derived from the survey, in terms of what kind of analyses can be applied to the data. Our suggestions, however, must not be viewed in isolation and need to be weighed against the practical implications of attempting to implement change to the current survey and the potential impact of any changes on the very real positive features of the survey as it currently exists. While this study has contributed important insights into the statistical procedures that can be employed with the data and has suggested some useful lines of enquiry for future research, we take this opportunity to reiterate our position as outsiders bringing a fresh perspective to the data. This position has enabled us to identify and investigate facets of the data not currently reported, such as the relationship between L1 and L2 reading performance. At the same time, it is important to reiterate that our suggestions and recommendations are just that: ideas offered for the consideration of those working within the incredibly rich context from which these data were collected.

Appendix 1: L2 Recognition Scale cross-tabulation with Age of Child

Grade			Age of Child												Total
			5	6	7	8	9	10	11	12	13	14	15	16	
1	L2 Recognition Scale	0	7,445	10,559	4,347	1,627	401	368	65	67	21	13	17	18	**24,948**
		1	2,075	3,877	1,797	709	193	164	34	27	9	6	5	5	**8,901**
		2	1,009	2,251	1,323	545	168	142	26	28	14	6	8	11	**5,531**
		3	378	1,133	874	340	128	80	23	24	5	5	8	13	**3,011**
		4	91	266	240	129	34	35	9	19	12	14	28	37	**914**
	Total		**10,998**	**18,086**	**8,581**	**3,350**	**924**	**789**	**157**	**165**	**61**	**44**	**66**	**84**	**43,305**
2	L2 Recognition Scale	0	703	2247	5,422	3,841	803	716	160	191	38	26	21	8	**14,176**
		1	379	1,433	3,769	2,701	571	501	127	120	34	14	12	9	**9,670**
		2	271	1,117	3,113	2,259	587	413	110	102	32	13	7	10	**8,034**
		3	112	594	1,764	1,385	400	304	87	72	25	11	7	6	**4,767**
		4	31	195	731	655	204	124	34	38	16	11	16	53	**2,108**
	Total		**1,496**	**5,586**	**14,799**	**10,841**	**2,565**	**2,058**	**518**	**523**	**145**	**75**	**63**	**86**	**38,755**
3	L2 Recognition Scale	0	137	436	1,431	3,902	2,039	1,097	247	272	81	36	22	9	**9,709**
		1	94	351	1,203	3,614	2,069	974	246	258	53	42	22	3	**8,929**
		2	63	239	1,149	3,896	2,119	1,067	213	266	75	29	16	5	**9,137**
		3	53	180	875	2,801	1,710	778	233	200	75	40	11	7	**6,963**
		4	21	70	373	1,438	924	458	108	112	34	24	16	9	**3,587**
	Total		**368**	**1,276**	**5,031**	**15,651**	**8,861**	**4,374**	**1,047**	**1,108**	**318**	**171**	**87**	**33**	**38,325**

Appendix 1: L2 Recognition Scale cross-tabulation with Age of Child (continued)

Grade			Age of Child												Total
			5	6	7	8	9	10	11	12	13	14	15	16	
4	L2 Recognition Scale	0	105	120	336	1,070	1,913	1,923	370	388	106	57	46	17	**6,451**
		1	54	81	319	1,151	2,305	2,232	460	440	130	45	22	11	**7,250**
		2	44	91	342	1,352	3,015	2,983	568	534	127	73	27	8	**9,164**
		3	17	65	236	1,284	2,985	2,811	624	529	170	87	36	22	**8,866**
		4	3	24	115	714	2,084	1,633	456	346	94	63	33	19	**5,584**
	Total		**223**	**381**	**1,348**	**5,571**	**12,302**	**11,582**	**2,478**	**2,237**	**627**	**325**	**164**	**77**	**37,315**
5	L2 Recognition Scale	0	17	24	93	337	436	2,010	1,009	680	165	116	67	34	**4,988**
		1	16	20	78	368	613	2,868	1,484	822	202	128	60	31	**6,690**
		2	16	14	71	396	809	4,011	2,029	1,173	296	119	74	32	**9,040**
		3	11	21	74	426	875	4,479	2,476	1,295	364	160	72	57	**10,310**
		4	10	9	37	222	814	3,713	2,058	1,018	290	137	84	31	**8,423**
	Total		**70**	**88**	**353**	**1,749**	**3,547**	**17,081**	**9,056**	**4,988**	**1,317**	**660**	**357**	**185**	**39,451**

References

Alderson, JC (2000) *Assessing Reading*. Cambridge: Cambridge University Press.

Bachman, LF (2004) *Statistical Analyses for Language Assessment*. Cambridge: Cambridge University Press.

Banerji, R (2013) 'The birth of ASER', in *Annual Status of Education Report* (ASER) 2012. Pratham.

Bernhardt, EB and Kamil, ML (1995) Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics* 16/1: 15–34.

Bharat, R and Wadhwa, W (n.d.) Survey design and precision of ASER estimates. Pratham Technical Report. Available online at www.asercentre.org/sampling/precision/reliability/validity/p/180.html

Brant, R (1990) Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics* 46/4: 1171–1178.

Cohen, JP, West, SG and Aiken, LS (2003) *Applied Multiple Regression/Correlation Analysis for the Behavioural Sciences*. Lawrence Erlbaum Associates: New Jersey.

Cummins, J (1979) Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research* 49/2: 222–251.

Field, A (2009) *Discovering statistics using SPSS*. London: Sage.

Goldstein, H (1995) *Multilevel statistical models*. London: Edward Arnold.

Graddol, D (2010) *English Next*. New Delhi: British Council.

Graddol, D (2010) *English Next India*. New Delhi: British Council.

Harrell, FE (2001) *Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis*. NY: Springer-Verlag.

Huhta, A, Alderson, JC, Nieminen, L and Ullakonoja, R (2011) *Diagnosing reading or writing in a second or foreign language*. Paper presented at American Association for Applied Linguistics Conference, March 2011, Chicago.

Meganathan, R (2011) 'Language policy in education and the role of English in India: From library language to language of empowerment', in Coleman, H (ed) (2011). *Dreams and Realities: Developing Countries and the English Language*. London: British Council.

McCullagh, P (1980) Regression Models for Ordinal Data. *Journal of the Royal Statistical Society*. Series B (Methodological), Vol. 42, No. 2, pp. 109–142

McCullagh, P, and Nelder, J (1989) *Generalized linear models*. London: Chapman and Hall.

McCulloch, C, and Searle, S (2001) *Generalized, linear, and mixed models*. New York: John Wiley and Sons.

National Council of Educational Research and Training (NCERT) (2006) *National Focus Group Position Paper on Teaching of English*. New Delhi: NCERT

National Council of Educational Research and Training (NCERT) (2012) A Study on the Teaching of English in Government Schools at the primary level in India. (Unpublished). New Delhi: NCERT.

Norušis, MJ (2011) IBM SPSS *Statistics 19 Advanced Statistical Procedures Companion*. Upper Saddle River, NJ: Prentice Hall

Pratham (2007) Annual Status of Education Report (ASER). Available online at www.asercentre.org/education/India/status/p/143.html

Pratham (2009) Annual Status of Education Report (ASER). Available online at www.asercentre.org/education/India/status/p/143.html

Pratham (2012) Annual Status of Education Report (ASER). Available online at www.asercentre.org/education/India/status/p/143.html

Saida, C and Hattori, T (2008) Post-hoc IRT equating of previously administered English tests for comparison of test scores. *Language Testing* 25/2: 187–210.

Sparks, RL, Patton, J, Ganschow, L and Humbach, N (2012) Do L1 Reading Achievement and L1 Print Exposure Contribute to the Prediction of L2 Proficiency? *Language Learning* 62/2: 473–505

Vagh, SB (n.d.) Survey Design and Precision of ASER Estimates. Pratham Technical Report available online at www.asercentre.org/sampling/precision/reliability/validity/p/180.html

Williams, R (2006) Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *The Stata Journal*, 6, Number 1, pp. 58–82

Yanase, K (2012) *Nihongo de no gengo seikatsu ga eigo gakushu ni oyobosu eikyou [The influence of language use in Japanese on English learning]*. Paper presented at the Japan Association of College English Teachers 51st International Convention, September 2012, Aichi, Japan.

Yamashita, J (2004) Reading attitudes in L1 and L2, and their influence on L2 extensive reading. *Reading in a Foreign Language* 16/1: 1–19.

5

Looking back and looking forward

Barry O'Sullivan
British Council

As I write this concluding chapter I will take the opportunity to reflect on the tremendous amount of work and co-operation between the newly formed research team here at the British Council in London and our colleagues at the British Council in India and at the Pratham-ASER Centre. The fruition of this collaboration offers the unique insight into the reality that is English language ability and use that is to be found in the less-well-researched side of modern India that can be found in this report.

Background

Rothman and Treffers-Daller (Chapter 1) present a clear overview of current thinking in the area of multilingualism. In their chapter, they highlight the fact that there is no empirical evidence to support the traditional view that asking children to deal with a number of languages can lead to confusion. On the other hand, being multilingual has been shown to have considerable social, educational and long-term cognitive/medical benefits.

In Chapter 2, Amritavalli presents a view of English as a first, second and foreign language in India. This complex mix of roles for English is reflected in her observations on the advantages of the English language (for internal and external communication) and its likely disadvantages (in effectively limiting the opportunities for the creation of a true multilingual society). Amritavalli suggests that the language of learning should not be limited to a single source (English or Hindi, for example) but should switch to meet the needs of the learners – a position supported by Rothman and Treffers-Daller. My interpretation of the central theme in Amritavalli's chapter is that, in the Indian context, English should be seen as a key language in the multilingual armoury of learners, but that the importance of other languages should not be overlooked. The different languages contribute in different ways to the learners' perception of themselves as national and international citizens.

Of course, the link between the national and international is of particular significance, given the role of the English language in the world today. That this role is often portrayed as transcending international borders is certainly justified in the *de facto* position of English as the language of international communication (whether this be for business, political or academic purposes). The possible negative implications of English 'taking over' our communication needs is hinted at by Amritavalli, who argues for a multilingual approach to learning, in which languages are seen to equally contribute to social communication, a position also supported by Rothman and Treffers-Daller.

In the third chapter of the volume, Banerji and Bobde (Chapter 3) present a fascinating insight into the rationale and approach taken by the Pratham Group in the ASER project. This overview neatly contextualises the report and allows the reader to more fully understand why the instruments used have been developed in the way they have been.

The view, therefore, that emerges from the three chapters that set the background to this report, is that while English is a hugely important element of the educational process in India, its true value should be seen in terms of its role in the multilingual society that is India. This role is important in defining India's position in the world, but also in defining India's people, since English is, in this context, a language for both national and international communication.

The research undertaken

All this background sets the scene perfectly for the research questions presented by Dunlea and Dunn in Chapter 4. In this chapter, Dunlea and Dunn set out three research questions:

1. Are there any trends in L2 reading performance in primary school across the different years of the survey?
2. Is there a relationship between L1 literacy and L2 English reading performance in primary school?
3. Given the wide age range of students within each school year, does age have a significant impact on L2 English reading performance within grades in primary school?

In summary, these have been answered as follows:

1. Yes, there is a trend. While the indication from the analysis indicates that the level of performance has been declining over the years, the difference between the initial two years (2007 and 2009) is not statistically significant. This means that any observed difference may be due to chance and should not be interpreted as reflecting a real decline. However, the decline in performance between both the 2007 and 2009 data collections and 2012 is statistically significant and can be interpreted to reflect a meaningful (if small) decline.

2. The answer again is yes, there appears to be a strong positive relationship between L1 reading performance and L2 performance. However, the limitations of the instruments used mean that this finding, while clearly important, should be interpreted with great caution. This is because the instruments used were not actually designed to answer this question and the analysis was performed retrospectively – suggesting that the findings are likely to be an accurate reflection of the underlying situation but that a more focused approach is needed in order to allow for more certainty.

3. The response to this question was complicated by the fact that even within a single school grade there was a broad range of ages to be found. However, it was noted that the majority of learners within a grade seemed to represent a range of three years. With this in mind, the researchers investigated the impact of age at the grade level and found that the older the learner, the higher the likelihood was that they would demonstrate a higher level of performance. When a similar comparison was made across grades, the same trend was observed. So, the answer to the question was that yes, age does appear to have a significant impact on performance level.

The discussion and findings reported in Chapter 4 have a number of implications for policy and for the Pratham-ASER Centre, which can be seen as either data related (i.e. the structure of the data that emerges from the collection instrument), instrument related (the type and format of the individual questions within the instrument) or analysis related (the type of analyses that might be undertaken both of existing data and of the new data types that emerge from changes made to the instruments).

Since Dunlea and Dunn have commented in some detail on the recommendations that stem from their research I will not repeat those here. Instead, I would like to focus on how the co-operation between the British Council and the Pratham Group may be manifested in the future, based on their work.

The future

It is clear to me that the success of the current co-operation should point to continued collaboration between the two organisations. The expertise brought to the analysis and interpretation of the findings by the British Council researchers can add significantly to the impact of ASER in the future. By working together to rethink and re-interpret the surveys undertaken during the years to now (2007, 2009 and 2012) we can develop a more detailed understanding of this complex issue and begin to make important recommendations to policy-makers based on empirical data.

It would appear from the analysis reported in this volume that some short-term changes to the way in which questions are set in the current instrument can add significantly to the value of the work, while more considered longer-term consideration of the assessments used is likely to result in even greater gains. The recommendations stemming from this report tend to reflect current thinking

in relation to statistical data analysis. The limitations reported are related to the way in which the instruments are operationalised, rather than in the constructs that underpin the approach. In fact, the way in which the instruments have been conceived reflects the most recent cognitive model of reading ability and progression, namely that of Khalia and Weir (2009) in the way that progression is modelled (i.e. from the phoneme level to the textual and beyond). It may be valuable, for instance, to look again at the measures to incorporate the construct measured into a single scale (i.e. combining the elements of recognition with comprehension).

Another issue to emerge from the work reported in this volume is that of test comparability. While the non-parametric tests used in the statistical analysis here represent one approach to dealing with the issue, they should not be seen as a long-term solution. More rigorous approaches to test equation could be explored in future co-operative partnerships. These suggestions are, of course, not intended to undermine the considerable work that has already been undertaken in developing the current instrument; see, for example, Ramaswami and Wadhwa's (2010) analysis of issues around sampling and Vagh's (2009) interesting work on the validation of the instruments.

The complexity of the process of gathering data from such huge numbers of respondents across such a huge country cannot be ignored. Any changes to the way in which the data are collected (be these large or small) must be considered with tremendous care, as practicality must be a key consideration in any future collection approach. Fortunately, it is clear from the success we have had in working together to create this report that we can build instruments that reflect the reality of data collection on the ground as well as the well-understood constructs that currently underpin the approach taken and the requirements and expectations of complex statistical modelling.

References

Khalifa, H and Weir, CJ (2009) *Examining Reading*. Cambridge: Cambridge University Press.

Ramaswami, B and Wadhwa, W (2010) Survey Design and Precision of ASER Estimates. Available online at: http://img.asercentre.org/docs/Aser%20survey/Technical%20Papers/precisionofaserestimates_ramaswami_wadhwa.pdf

Vagh, BS (2009) Validating the ASER Testing Tools: Comparisons with Reading Fluency Measures and the Read India Measures. Available online at: www.asercentre.org/sampling/precision/reliability/validity/p/180.html

ISBN 978-0-86355-721-7

www.britishcouncil.org
www.pratham.org
www.asercentre.org