

Progress in International Reading Literacy Study: Measuring and making international comparisons of student achievement in reading

Progress in International Reading Literacy Study (PIRLS) is an ongoing, cyclical international study of student achievement in reading. It is directed by the International Association for the Evaluation of Educational Achievement (IEA).

Origins and context

The IEA was founded in 1958. It has evolved from a collective of research bodies into a professional organisation with a secretariat based in Amsterdam (NLD) and a centre devoted to data processing and research based in Hamburg (DEU). Beyond this professional organisation, IEA has over 70 members that are governmental and non-governmental educational research institutions from countries in Africa, Asia, Australasia, Europe, the Middle East and the Americas. Most of the members represent national education systems. IEA also maintains funding and non-funding partnerships (IEA, n.d.-a).¹

According to IEA's founders, the different education systems across the world together form a kind of educational laboratory, and comparative research into these different systems can reveal important relationships between inputs and outcomes, relationships that would not necessarily be detected if any one system were studied in isolation (IEA, 2014a).

¹ Some of IEA's recent funding partnerships have been with the European Commission, the Ford Foundation, the Inter-American Development Bank, the United Nations Development Program and the World Bank. Its non-funding partnerships include those with the Educational Testing Service (ETS), UNESCO's International Institute for Educational Planning (IIEP UNESCO), Organization of Ibero-American States (OEI), Partnership for Educational Revitalization in the Americas (PREAL), Programme d'Analyse des Systèmes Educatifs CONFEMEN (PASEC), Statistics Canada, the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), and the West African Examinations Council (WAEC) (IEA, n.d.-a).

IEA studies seek to understand the processes and products of education by administering cognitive assessments and collecting background data to examine the *intended curriculum*, the *implemented curriculum* and the *attained curriculum* (IEA, n.d.-a). The intended curriculum is concerned with the national, social and educational contexts. It covers what is described in curriculum policies and publications, and how the education system is structured to facilitate the learning that is described in these policies and publications. The implemented curriculum is concerned with the school, teacher and classroom contexts. It covers what is actually taught in the classrooms and how it is taught, including the characteristics of the individuals who are teaching. The attained curriculum is concerned with the learning outcomes and characteristics of students. It covers what students learn, what they think about what they learn, and their backgrounds (I. V. S Mullis & Martin, 2013).

This three-aspect concept of the curriculum has been used in many of the 30 comparative research studies IEA has conducted since its inception. While PIRLS and Trends in International Mathematics and Science Study (TIMSS) measure performance in basic school subjects, studies have also been conducted in areas such as literature, advanced mathematics and physics, civics and citizenship, and computer and information technology (IEA, 2014b, 2014c).

At the international level, PIRLS is managed by the TIMSS & PIRLS International Study Center at the Lynch School of Education at Boston College. Each participating entity has a research coordinator team that is responsible for the local implementation of the study.

PIRLS has been conducted every five years since 2001. Although PIRLS was a follow-up to IEA's 1991 Reading Literacy Study. The number of participating countries

has grown from 35 in the first cycle to 50 in the fifth cycle, with a further 11 benchmarking entities (such as provinces).

PIRLS is funded by participants and through IEA's funding partnerships.

Purpose

PIRLS measures the reading comprehension performance of students and collects a wide array of contextual information about students, schools, curricula and educational policies and systems. PIRLS is designed to inform educational practice and policy by providing an international perspective of teaching and learning in reading literacy (IEA, n.d.-b).

Measurement objectives

Assessment domains

PIRLS is an international assessment that is both curriculum-based and standardised. The development of cognitive tests that represent the curricula of all participants involves extensive research, consultation and consensus-building. The initial PIRLS Assessment Framework built on the 1991 Reading Literacy study, which provided the basis for the definition of reading literacy and developing the research instrument. The first cycle of the PIRLS assessment Framework was developed through the collaboration of almost 40 participating countries (Campbell, Kelly, Mullis, Martin, & Sainsbury, 2001).

The PIRLS assessment framework is updated each cycle. Updating the frameworks is a collaborative process, involving the following participants: National Research Coordinators from participating countries/benchmarking entities, the TIMSS & PIRLS International Study Centre, chief subject consultants and international expert committees.

In updating the Framework, two competing interests need to be balanced: the frameworks must maintain continuity to enable trend measurement, whilst being adjusted to stay relevant for changing educational contexts (I. V. S Mullis & Martin, 2012). This is achieved by progressively replacing old items and texts with new ones. No item feature for more than three assessment cycles, but core trend texts are retained (I. V. S Mullis & Martin, 2012).

In 2016, PIRLS was further increased to 20 passages to include a second assessment option—PIRLS Literacy (which was earlier known as prePIRLS), which is a less difficult reading assessment that is equivalent in scope to PIRLS. Also, the ePIRLS assessment option was introduced, which is an assessment of online reading (I.V.S Mullis & Martin, 2019).

In 2021, PIRLS Literacy has been incorporated into the main PIRLS assessment. There are three levels of passage difficulty that are combined into two levels of booklet difficulty. The use of less or more difficult booklets varies with the reading achievement level of the students in the country. This is a group adaptive approach (I.V.S Mullis & Martin, 2019).

The understanding of reading has been shifting from merely demonstrating fluency and basic comprehension to demonstrating the ability to apply what is understood or comprehended to new situations or projects. The definition of reading presented in the 2016 Assessment Framework is:

Reading literacy is the ability to understand and use those written language forms required by society and/or valued by the individual. Readers can construct meaning from texts in a variety of forms. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment.

Reading is assessed at the fourth grade by two reading purposes—literary and informational. There are then four comprehension strategies that are assessed within each of the purposes: retrieval, inferencing, integrating, and evaluation (I.V.S Mullis & Martin, 2019). Table 1 presents the PIRLS 2016 assessment matrix of purposes and comprehension processes.

Table 1: Reading Purposes and Comprehension Processes in the PIRLS 2016 assessment framework

Purposes for Reading	Processes of Comprehension
Literary Experience	Focus on and Retrieve Explicitly Stated Information
Acquire and Use Information	Make Straightforward Inferences Interpret and Integrate Ideas and Information Evaluate and Critique Content and Textual Elements

The PIRLS passages are classified by their primary purposes. Passages classified as literary have questions addressing theme, plot events, characters, and setting; passages classified as informational have questions about the content of the passages. The comprehension processes are evaluated across all passages, as the comprehension processes that readers use are more similar than different for both purposes.

Contextual information

The PIRLS context questionnaires aim to facilitate a better understanding of the contextual factors that affect how students learn to read and their educational opportunities. Linking this data with achievement results in the PIRLS assessment can help interpret those results and inform policy to improve reading skills (I.V.S Mullis & Martin, 2019). The context questionnaires collect data about five influences on student reading development: home, school, classroom, national and student attributes.

The home context is divided into two aspects: the environment for learning and the emphasis on children's literacy skills. The environment for learning includes: resources for learning in the home, whether parents like reading, and the language spoken in the home. The emphasis on learning includes: early literacy activities, early literacy tasks when beginning primary school, and parental expectations of children's education and attitudes towards reading (I.V.S Mullis & Martin, 2019).

The school context includes: school resources generally, and specifically for reading instruction, the school climate for learning, the degree of discipline and safety in the school, and the school's emphasis on reading instruction (I.V.S Mullis & Martin, 2019).

The classroom context focuses on factors related to the teaching of reading. This includes: student engagement, strategies, types of texts assigned, organisation for teaching, library resources and classroom teaching. Additionally, data about characteristics of teachers are collected, including: teacher preparation, professional development and teaching experiences. In 2021, data will also be collected about information technology in the classroom (I.V.S Mullis & Martin, 2019).

At the national level, contextual factors are divided into the organisation of the education system and the reading curriculum. The former includes: languages of instruction, system for pre-primary education, age of entry and retention and number of years of schooling. The latter includes: reading curriculum in the primary grades and strategies for students with reading difficulties (I.V.S Mullis & Martin, 2019). This information is provided in the PIRLS encyclopaedia, which each participating country/benchmarking entity is required to provide. They are usually written by staff from ministries of education or research institutions (I.V.S Mullis & Martin, 2019).

Information about student attributes that are sought relate to students' basic demographic characteristics and their attitudes towards reading. The attitudinal information includes: whether students like reading, are confident readers, their familiarity with digital devices and whether they like the assessment passages. The demographic characteristics sought include gender and age (I.V.S Mullis & Martin, 2019).

The contextual questionnaires are updated, ensuring that they reflect changes in education practices and contexts. Existing scales are updated and new scales are added.

Target population and sampling methodology

The target populations for PIRLS are defined with reference to UNESCO's International Standard Classification of Education (ISCED) scheme. The PIRLS target population is the grade that represents four years of schooling, counting from the first year of ISCED Level 1; this corresponds to the fourth grade in most countries. To better match the assessment to the achievement level of students, countries have the option of administering PIRLS at the fifth or sixth grade (IEA, n.d.-b). The IEA explains that this population has been targeted because 'Typically, at this point in their schooling, students have learned how to read and are now reading to learn' (I. V. S. Mullis & Martin, 2019).

The above target population includes all students within the definition. However, participants are permitted to make school-level and student-level exclusions for political, organisational and operational reasons, providing these exclusions are based on the clearly defined grounds and rules defined in the methods and procedures manual (LaRoche, Joncas, & Foy, 2017)

There are technical standards for the sampling precision of estimates. These standards are usually met with a single intact class from 150 schools that yields approximately 4000 students for each target population. For countries participating in both PIRLS and PIRLS Literacy, the required student sample size is doubled, resulting in around 8000 students (LaRoche et al., 2017).

Some participating countries/benchmarking entities sample more than one class per sampled school, this enables a larger student sample, the better estimation of school-level effects and internal level comparisons, such as between national regions. They may also be required to sample more than 150 schools if the standard class size is particularly small, if schools stream students by ability, if high levels of non-response are expected, or if the PIRLS standards for sampling precision have not been met in previous cycles (LaRoche et al., 2017).

The National Research Coordinator from the participating country/benchmarking entity are responsible for developing and implementing the national sampling plan, with support from IEA and Statistics Canada. The research coordinator constructs a complete and accurate sampling frame, based

on the international two stage sampling design (LaRoche et al., 2017).²

In the second stage of sampling, one or more intact classes are sampled with equal probability of selection using systematic random sampling.³ Class sampling is undertaken by the research coordinator using software developed by IEA for use in its surveys. PIRLS samples intact classes rather than directly sampling students for two reasons. Firstly, the study examines students' curricular and instructional experiences, and these are often organised at the classroom level. Secondly, it minimises disruption at the schools if the assessment includes all students in some classes rather than some students from all classes (LaRoche et al., 2017).

Assessment administration

Within a participating country/benchmarking entity, after schools have been sampled the national research centre is responsible for identifying and training school coordinators. The school coordinators are tasked with providing the national research centre with information for within-school sampling of classes; identifying and training test administrators; updating tracking forms; organising the time and place for test administration; distributing questionnaires; maintaining the security of test booklets; and managing the receipt and return of all assessment materials (LaRoche et al., 2017).

Managing the activities on the day of test administration is the responsibility of the test administrators. At the start of test sessions, the test administrators must read instructions that are standardised across all participating countries/benchmarking entities. A test session is divided into two parts. The duration of each part and the duration of the break between the two parts are also standardised across all participating countries/benchmarking entities.⁴

Numerous steps are taken to ensure the quality of the assessment. Both the school coordinators and the test administrators are supported in their work by manuals that are developed by the TIMSS & PIRLS International Study Centre and translated and adapted by national research centre staff as required. Test administration, scoring, and data entry and processing, are standardised as much as possible. To achieve this, PIRLS has developed and documented procedures, protocols, software and training, and also initiated an

2 If explicit stratification is used then one sampling frame must be constructed for each explicit stratum.

3 Since small classes increase the risk of obtaining unreliable estimates, if a sampled school is identified as having small classes, these classes are grouped together into pseudo-classes that have adequate numbers of student before the second stage of sampling.

4 Each of the two parts of a test session is 40 minutes; the break between the two parts cannot exceed 30 minutes (Johansone, 2017).

independent quality assurance program. Furthermore, International Quality Control Monitors visit a sample of schools in each country to observe test administration and send national quality monitors to 10% of schools (Johansone, 2017).

Since PIRLS is a comparative international survey, assessment booklets must be standardised across countries. This includes the translation and adaption of test items from the source language to the target language(s), based on standard, international agreed procedures. The IEA then manages the processes of quality assurance, engaging external reviewers. The IEA provides feedback on the translations and adaptations, which the National Research Centres are expected to review and act on where necessary (Ebbs & Wry, 2017)

In 2021, PIRLS will transition to a digital format, with half of the countries delivering PIRLS via a digital platform. The use of digital technology assists in the group adaption design, where participating entities can choose to use varying levels of difficult booklets (I.V.S Mullis & Martin, 2019).

Reporting and dissemination

After each assessment cycle, PIRLS results are reported in international reports prepared by the TIMSS & PIRLS International Study Centre. Each report begins with some introductory information about the history and context of PIRLS, the nature of the current assessment, and the range of participating countries/benchmarking entities. Student achievement results are presented next, followed by the background questionnaire data.⁵

Results are reported for each participating country in terms of means and distributions of student achievement. Trends in achievement over multiple cycles, cohort comparisons, achievement differences by gender and trends in achievement differences by gender are also reported.

Student achievement results are reported with reference to four points on the PIRLS international benchmark scale: advanced (625), high (550), intermediate (475), low (470) (I. V. S Mullis & Prendergast, 2017).

The PIRLS International Benchmarks are given not only as numerical proficiency scores but also as detailed proficiency descriptions. These descriptions of what benchmark scores mean in terms of knowledge and skills are developed by the TIMSS & PIRLS International Study Center and the item review experts through data analysis and conceptual analysis of the assessment items. Examples of anchor items (i.e. items that

5 All PIRLS reports can be downloaded from <https://www.iea.nl/publications/study-reports/international-reports-iea-studies>

function best for students with achievement at or near a benchmark) are also provided.

Student achievement is also reported using the scales for each reading purpose and process sub-scales. Average achievement on each sub-scale is compared to average achievement on the relevant overall scale. Trends in average achievement and average achievement disaggregated by gender are also reported.

A variety of background data for students, teachers and schools is reported and linked with average achievement scores. A number of policy-relevant questionnaire scales are presented, covering areas including: resources available at home for learning and education, resources available at school, teacher working conditions, school climate and students' attitudes towards learning. In the PIRLS 2016 international results report, chapter headings included: student achievement, performance at international benchmarks, achievement in reading purposes and comprehension processes, home environment support, school composition and resources, school climate, school safety, teachers' and principals' preparation, classroom instruction, and student engagement and attitudes.

Participating entities also produce their own reports. The analysis and format is similar to the international reports, but with less emphasis on international comparison and more focus on each participating entities own issues of concern and relevance. This includes providing intra-country comparisons (such as between different provinces), regional comparisons (such as metropolitan versus rural), comparing school types, (such as private and public), and a focus on specific demographic groups (such as particular ethnicities).

While the results reports present the data from the student, teacher and school questionnaires, the data from the curriculum questionnaire are presented in the PIRLS encyclopaedias. These data are not analysed, but simply presented in a way that enables easy comparison.

In addition to the results reports and encyclopaedias that are produced each cycle, PIRLS also produces technical reports (also called 'Methods and Procedures') that describe in detail all technical aspects of the assessment.

PIRLS results reports, encyclopaedias, technical reports, assessment frameworks and other documentation for all cycles can be downloaded from the website of the TIMSS and PIRLS International Study Centre.⁶ The international databases for all cycles, and accompanying user guides, can be downloaded from

the TIMSS and PIRLS website.⁷ IEA's Data Processing Centre has developed the IEA IDB Analyser and IEA Data Visualiser software applications to facilitate the analysis and visualisation of data from IEA studies. These applications can be downloaded from IEA's website.⁸

Influence

PIRLS appears to be influencing a degree of policy convergence amongst participating countries with regards to curriculum and teacher education. Many of the PIRLS countries have given more attention to teacher education and reading instruction, modifying university programs and providing professional development. Nearly all the PIRLS 2016 countries have institutionalised objectives to improve reading instruction. A number of programs to improve early learning and readings skills have been initiated. Many countries have updated their curriculum, with greater emphasis on comprehension strategies, analytical skills and informational reading, in line with the PIRLS assessment. Additionally, there has generally been greater emphasis on enhancing student motivations to read and reading for pleasure (I. V. S. Mullis, Martin, M. O., Goh, S., & Prendergast, C. (Eds.), 2017).

PIRLS has also highlighted where there is underachievement within a country, thereby enabling governments to provide more support for disadvantaged populations, if they choose to. Through the implementation of statewide monitoring of educational achievement, PIRLS has enabled the success of initiatives to be evaluated (I. V. S. Mullis, Martin, M. O., Goh, S., & Prendergast, C. (Eds.), 2017). In general, countries that perform poorly compared to similar countries (such as regional neighbours) or whose performance has declined, have tended to initiate policy changes (Cresswell, Schwantner, & Waters, 2015).

A study under the auspices of the IEA reviewed the impact of PIRLS 2006 in 12 countries. It found that PIRLS had a wide range of influences, including: structural change to education systems, the establishment of dedicated research and evaluation units, the implementation of policies focussed on boosting education quality, and curricula changes (Schwippert & Lenkeit, 2012).

For example, in Hong Kong the curriculum was modified, with greater emphasis on reading for pleasure, rather than for mere instrumental purposes. 'Reading for pleasure' was made a key learning area and teaching objective, with greater resources provided to primary school to support this. Specifically,

6 <https://www.iea.nl/publications/study-reports/international-reports-iea-studies>

7 <https://timss.bc.edu/databases-landing.html>

8 <https://www.iea.nl/data-tools/tools#section-308>

in some schools, the first lesson of the day was reserved for pleasurable reading. Furthermore, existing assessments were modified to absorb the PIRLS theoretical content, and formative assessments were introduced to monitor progress. Additionally, guides were produced for parents to support them in developing the reading ability of their preschool children. The guidebook includes: information about the role of parents in fostering language skills in babies and the importance of reading to one's children; advice to parents how to encourage good reading habits in children, and suggestions to parents on how they can create an environment that encourages reading. In addition to new initiatives, existing programs were given more prominence, such as 'Read to Learn' (I. V. S. Mullis, Martin, Goh, & Prendergast, 2017).

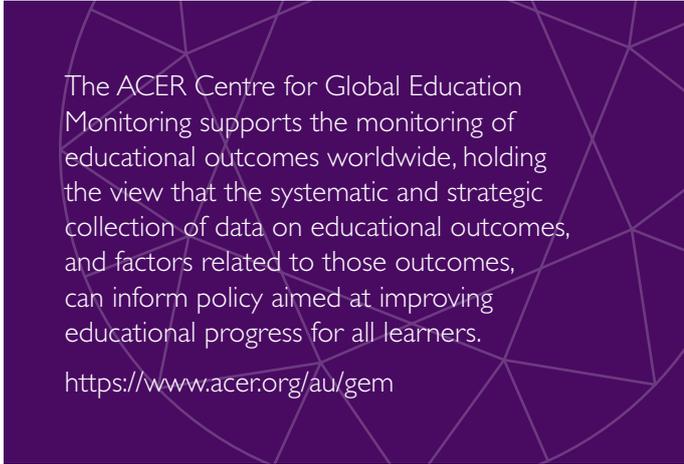
An evaluation of the impact of PIRLS and TIMSS in low and middle income countries for the World Bank found that generally, PIRLS greatly influenced understanding of education by policy makers (Gilmore, 2005). This impact was largely due to the use of international comparisons, especially relating to student achievement, curriculum, teaching methods and education resources (Gilmore, 2005). Furthermore, teachers' practices were likely influenced through the wide distribution of PIRLS reports to teachers and the public (Gilmore, 2005).

The ACER Global Education Monitoring Centre supports the monitoring of educational outcomes worldwide, holding the view that the systematic and strategic collection of data on educational outcomes, and factors related to those outcomes, can inform policy aimed at improving educational progress for all learners.

References

- Campbell, J., Kelly, D., Mullis, I. V. S., Martin, M. O., & Sainsbury, M. (2001). *Framework and Specifications for PIRLS Assessment 2001*. Chestnut Hill, MA: International Association for the Evaluation of Educational Achievement.
- Cresswell, J., Schwantner, U., & Waters, C. (2015). *A Review of International Large-Scale Assessments in Education: Assessing Component Skills and Collecting Contextual Data*. Melbourne: Australian Council for Educational Research.
- Ebbs, D., & Wry, E. (2017). Translation and Layout Verification for PIRLS 2016. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and Procedures in PIRLS 2016*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).
- Gilmore, A. (2005). *The Impact of PIRLS (2001) and TIMSS (2003) in Low- and Middle-Income Countries: an Evaluation of the Value of World Bank Support for International Surveys of Reading Literacy (PIRLS) and Mathematics and Science (TIMSS)*. Retrieved from New Zealand:
- IEA. (2014a). IEA: Brief history. Retrieved from http://www.iea.nl/brief_history.html
- IEA. (2014b). IEA: Completed studies. Retrieved from http://www.iea.nl/completed_studies.html
- IEA. (2014c). IEA: Studies. Retrieved from <http://www.iea.nl/studies.html>
- IEA. (n.d.-a). About Us [Pamphlet]. Retrieved from http://www.iea.nl/fileadmin/user_upload/IEA_Documents/IEA_Brochure.pdf
- IEA. (n.d.-b). PIRLS. Retrieved from <https://www.iea.nl/studies/iea/pirls>
- Johansone, I. (2017). Survey Operations Procedures in PIRLS 2016. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and Procedures in PIRLS 2016*. Chestnut Hill, MA: International Association for the Evaluation of Educational Achievement.
- LaRoche, S., Joncas, M., & Foy, P. (2017). Sample design in PIRLS 2016. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and Procedures in PIRLS 2016*. Chestnut Hill, MA: International Association for the Evaluation of Educational Achievement.
- Mullis, I. V. S., & Martin, M. O. (Producer). (2012, 19 February 2014). Developing Assessment Frameworks for Measuring trends in a Changing Context [PowerPoint slides] Retrieved from http://www.iea.nl/fileadmin/user_upload/General_Assembly/53rd_GA/GA53_Developing_Assessment_Frameworks.pdf
- Mullis, I. V. S., & Martin, M. O. (2019). *PIRLS 2021 Assessment Frameworks* (L. S. o. E. TIMSS & PIRLS International Study Center, Boston College and International Association for the Evaluation of Educational Achievement (IEA) Ed.). Chestnut Hill, MA: International Association for the Evaluation of Educational Achievement.
- Mullis, I. V. S., & Martin, M. O. (Eds.). (2013). *TIMSS 2015 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).
- Mullis, I. V. S., Martin, M. O., Goh, S., & Prendergast, C. (2017). *PIRLS 2016 Encyclopedia: Education Policy and Curriculum in Reading*. Chestnut Hill, MA: International Association for the Evaluation of Educational Achievement.

- Mullis, I. V. S., Martin, M. O., Goh, S., & Prendergast, C. (Eds.). (2017). *PIRLS 2016 Encyclopedia: Education Policy and Curriculum in Reading*. Retrieved from <http://timssandpirls.bc.edu/pirls2016/encyclopedia/>
- Mullis, I. V. S., & Prendergast, C. (2017). Using scale anchoring to interpret the PIRLS and ePIRLS 2016 achievement scales. In I. V. S. Mullis & M. O. Martin (Eds.), *PIRLS 2021 Assessment Frameworks*. Chestnut Hill, MA: International Association for the Evaluation of Educational Achievement.
- Schwippert, K., & Lenkeit, J. (2012). *Progress in Reading Literacy in National and International Context: the Impact of PIRLS 2006 in 12 Countries*. Munster: Waxman.

The background of the text box is a dark purple color with a faint, light purple geometric pattern of overlapping circles and lines, resembling a network or a globe's structure.

The ACER Centre for Global Education Monitoring supports the monitoring of educational outcomes worldwide, holding the view that the systematic and strategic collection of data on educational outcomes, and factors related to those outcomes, can inform policy aimed at improving educational progress for all learners.

<https://www.acer.org/au/gem>