"Navigating the Digital Divide: The Role of ICT in Enhancing Educational Outcomes in Thailand"

Leigh Pearson, Ph.D. Candidate
International College, Mahidol University
School of Global Studies, Thammasat University
leigh.pea@mahidol.edu

Last Updated: 17-11-2024

Note: The analyses and findings presented in this document are part of an ongoing PhD research project focused on ICT use in education within Thailand. These results represent preliminary findings and are intended to fulfill initial publication requirements. The complete research project, which includes further analyses and comprehensive conclusions, is described in full within the PhD-ICT-Education-Analysis repository. As the project progresses, additional insights and updates will be incorporated to provide a complete view of ICT's role in educational outcomes.

Contents

14 Ethical Considerations

1   Introduction

In an era where Information and Communications Technology (ICT) has become a cornerstone of global development, the role of ICT in education is especially crucial. Countries around the world are leveraging digital technologies to enhance learning, improve equity, and drive innovation in education.

The Thai education system has been a topic of concern for policymakers, educators, and parents for many years, and yet Thailand has made significant, often under-appreciated progress towards realising its aspiration of transforming into a high-income, knowledge-based economy. The country's vision details ambitious goals to upgrade the quality of education and provide its young people with 21st-century skills, as laid out by the National Strategy Committee (NSC, 2019) in its twenty-year National Strategy (2017–2036) and by the Office of the National Economic and Social Development Council (ONESDC, 2022) in its 12th and 13th National Economic and Social Development Plan (2017-2021 & 2023-2027).

Despite these efforts, Thai students' performance on international benchmarks has consistently lagged behind regional and global peers. For instance, the average scores for the five O-NET subjects have consistently fallen below 50 percent (National Institute of Educational Testing Service, 2021), PISA scores have been in an increasingly steep decline since 2012 in all three skills, particularly reading (OECD, 2023), and Thailand ranks 101st out of 113 participating countries for English proficiency (English First, 2023).

This discrepancy points to a critical gap between policy aspirations and on-the-ground realities. While there has been substantial progress in terms of ICT infrastructure and resources, the impact of these advancements on actual learning outcomes remains unclear. Furthermore, disparities in ICT access between urban and rural students, known as the 'first' digital divide, along with the 'second' digital divide, which reflects inequalities in ICT usage, together intensify educational inequalities, particularly affecting marginalized and disadvantaged students (Ma & Cheng, 2022).

This research, situated at the intersection of national policy and educational outcomes, aims to explore how ICT integration within Thai schools influences learning outcomes and equity. Using a combination of the PISA ICT Framework (OECD, 2019) and the Multi-Level Framework of Technology Acceptance and Use (MLFTAU; Venkatesh et al., 2016), this study will provide a nuanced understanding of both structural and behavioral factors that shape ICT use in educational settings.

Through an analysis of PISA datasets spanning multiple years and complemented by data from key national agencies, such as the National Statistical Office of Thailand, the Equitable Education Fund (EEF), and the Office of Basic Education Commission (OBEC), this research will assess the effectiveness of ICT in enhancing educational outcomes. Special attention will be given to how ICT impacts underprivileged students, who are often left behind in national efforts to digitize education.

By examining Thailand's experience, this research will contribute valuable insights to the global discourse on the role of technology in education, particularly in developing countries facing similar digital and educational challenges.

2   Utilization of Conceptual and Theoretical Frameworks

This study utilizes two primary frameworks—the PISA ICT Framework and the Multi-Level Framework of Technology Acceptance and Use (MLFTAU). These frameworks were chosen due to their complementary strengths in addressing both systemic and behavioral aspects of ICT integration, aligning with the study's goals to evaluate how ICT use influences educational outcomes in Thailand.

2.1   PISA ICT Framework

The PISA ICT Framework was developed by the OECD to specifically assess the impact of Information and Communication Technology (ICT) within educational systems across countries. This framework is adapted from the older CIPO model, which is structured around four components:

- Context: The broader environment in which the educational system operates, including socio-economic, policy, and infrastructure factors that shape ICT access and use.

- Input: Resources and support systems provided to schools, such as funding, ICT infrastructure, and training for teachers, as well as the personal characteristics of students and teachers.

- Process: The methods, pedagogies, and practices through which ICT is implemented within educational settings, both in-class and outside-the-classroom.

- Output: The measurable outcomes of ICT use, including academic performance, student engagement, and digital skills.

In the PISA ICT Framework, these four elements are adapted to reflect the specific dynamics of ICT usage in education. It provides a structure for analyzing both in-school and outside-the-classroom factors, such as the availability of ICT resources, policies supporting digital education, and the impact on students' performance and engagement.

The PISA ICT Framework was chosen for this study for the following reasons:

Systemic Analysis of ICT Integration

The PISA ICT Framework retains the CIPO model's structured approach to analyzing educational systems by dividing ICT impact into in-school (in-class) and outside-the-classroom components. This layered approach enables a comprehensive assessment of both the resources available and the contextual factors that influence ICT use, providing a clear structure for understanding ICT's role across various educational settings.

Alignment with International Benchmarks

The PISA ICT Framework is tailored to evaluate ICT's impact within the context of PISA's global educational assessments, directly supporting the study's focus on assessing ICT-related educational outcomes in Thailand. By applying a framework already validated within the PISA system, this research can leverage internationally recognized metrics and benchmarks to evaluate ICT's effectiveness in enhancing learning.

Outcome-Focused Structure

The PISA ICT Framework is outcome-driven, making it suitable for measuring tangible educational results like student performance on PISA benchmarks. This outcome orientation aligns directly with the study's aim to evaluate how ICT usage impacts academic performance, engagement, and equity in Thailand's educational system.

2.2 Multi-Level Framework of Technology Acceptance and Use (MLFTAU)

The Multi-Level Framework of Technology Acceptance and Use (MLFTAU), developed by Venkatesh and colleagues, provides a behavioral and psychological perspective on the adoption and effective use of technology. This framework considers factors at multiple levels—individual, organizational, and social—that influence how technology is perceived and utilized. The key components of MLFTAU include:

- Individual Level: Focuses on personal factors that influence technology use, such as perceived usefulness, perceived ease of use, and behavioral intention. These factors impact whether individuals (students and teachers) are willing to use ICT tools.

- Organizational Level: Examines how institutional support, resources, and training affect technology adoption within organizations (e.g., schools). This level is important for understanding how school resources and policies influence ICT integration.

- Social Influence: Considers the impact of peer influence, societal expectations, and cultural norms on technology use. For example, students may be influenced by how their peers or teachers view and use ICT.

- Facilitating Conditions: Reflects the infrastructure, support, and resources that enable or constrain effective ICT use. This includes both material resources (such as access to computers) and institutional support.

The MLFTAU framework was selected for this study for the following reasons:

Focus on Behavioral Drivers

MLFTAU captures individual attitudes, beliefs, and behavioral intentions related to technology use among teachers and students. By examining constructs like performance expectancy, effort expectancy, social influence, and facilitating conditions, this framework allows the study to investigate the psychological factors that drive or hinder ICT adoption in educational settings.

Applicability to Educational Stakeholders

MLFTAU's emphasis on understanding user acceptance of technology is especially relevant to this study's focus on teachers' and students' ICT usage. Given that effective ICT integration relies on both access and willingness to use technology, MLFTAU adds depth to the analysis by examining how individual motivations and institutional support affect ICT engagement.

Complementary to Systemic Analysis

While the PISA ICT Framework provides a structured view of ICT's systemic impact on educational outcomes, MLFTAU offers a granular understanding of the behavioral factors influencing ICT adoption. This dual approach captures both the structural and personal dimensions of ICT use, offering a holistic view of its role in Thailand's educational landscape.

In summary, the combined use of the PISA ICT Framework and MLFTAU enables a comprehensive analysis that spans both external, systemic factors and individual behavioral drivers. This integration aligns with the study's objectives to evaluate ICT's impact on educational outcomes while addressing the practical and psychological barriers to effective ICT use in Thai schools.

2.3 Application and Integration of Frameworks in Data Analysis

This study applies the CIPO and MLFTAU frameworks synergistically to examine the systemic and behavioral factors influencing ICT use in education. The CIPO model structures the investigation into how contextual, input, and process variables affect student outcomes, while the MLFTAU framework adds depth by analyzing individual attitudes and behaviors toward ICT acceptance within schools. This dual-framework approach enables a nuanced analysis that captures both the external factors shaping ICT access and the behavioral factors that drive ICT use.

In this analysis, data from multiple sources—including the PISA student and ICT questionnaires, along with national datasets like the EMIS and NSO—will be mapped onto the PISA ICT framework, which aligns with the CIPO model's dimensions of input, process, and output. Contextual information from national datasets provides insights into socio-economic and regional disparities, while process variables, such as ICT integration in classrooms, are explored using teacher and student responses from PISA questionnaires.

The MLFTAU framework adds a behavioral perspective to the analysis of ICT use. For example, regression models will assess how factors such as performance expectancy and facilitating conditions predict ICT usage among teachers and students, thereby influencing educational outcomes. This combined approach captures the systemic and behavioral complexities of ICT use, aligning with the research objectives of assessing ICT integration, identifying barriers and enablers, and making informed policy recommendations.

3 Objectives

The objectives of this research are to:

1. Assess the current state of ICT integration in Thai education and its alignment with national strategy aspirations.

By understanding how closely current ICT practices align with Thailand's strategic goals, this objective provides insights into whether investments and policy directives are translating effectively into school environments. This analysis will reveal gaps in ICT resources or infrastructure that may hinder educational quality, particularly in digital literacy and access.

The findings will help policymakers identify where the strategy falls short in practical application, enabling targeted adjustments. For example, if schools lack sufficient devices or broadband access, the government can prioritize these resources in underperforming areas, thereby supporting equitable learning environments and improving digital competencies across regions.

2. Identify the enablers and barriers to effective ICT use within educational settings.

Identifying the factors that encourage or obstruct ICT use in classrooms (such as teacher training, ICT resources, or administrative support) directly impacts student learning experiences. Schools with adequate enablers, like well-trained teachers and high-quality ICT infrastructure, are likely to see enhanced student engagement and better digital skill acquisition.

By pinpointing specific barriers, policymakers can implement targeted initiatives to address them, such as expanding teacher training programs or developing supportive policies. Removing these barriers supports more effective technology integration, enabling students to benefit from ICT resources and enhancing their overall learning outcomes, especially in digital skills essential for future job markets.

3. Analyze the relationship between ICT utilization and student performance on established educational benchmarks, using Plausible Values (PVs) and appropriate statistical models.

This objective explores how ICT usage correlates with academic performance across different socio-economic and geographic backgrounds, highlighting the role of ICT in bridging or widening educational inequalities. By examining this relationship, the study can determine whether ICT access is genuinely beneficial to student achievement or if disparities in usage contribute to unequal outcomes.

The results can guide policies aimed at reducing educational inequalities by improving ICT access for underserved groups, such as students in rural areas or from lower socioeconomic backgrounds. For instance, if ICT access is found to improve PISA performance, targeted support for disadvantaged students could be a powerful equalizer, potentially enhancing overall national performance on educational benchmarks.

4. Provide evidence-based recommendations to policymakers for enhancing ICT acceptance and utilization in Thai education.

Evidence-based recommendations can drive sustainable improvements in ICT use across Thai schools, addressing both systemic (CIPO model) and behavioral (MLFTAU) factors. These recommendations help ensure that the integration of ICT not only aligns with national goals but also supports enhanced educational equity and outcomes.

This objective synthesizes findings into actionable guidance, such as recommending specific infrastructure improvements, training initiatives, or policies that promote equitable ICT access. It can also propose frameworks for ongoing assessment and adjustment, ensuring that ICT policies remain responsive to emerging challenges and changing educational needs.

To clarify how each research objective aligns with the study's analytical frameworks, the following table maps each objective to the relevant dimensions of the CIPO model—Context, Input, Process, and Output—and the Multi-Level Framework of Technology Acceptance and Use (MLFTAU). Additionally, it distinguishes between in-school and outside-the-classroom contexts, reflecting the dual environment in which ICT influences student learning. This multi-layered approach underscores the study's comprehensive analysis of ICT integration by examining both structural factors within educational systems and behavioral factors that affect individual technology acceptance. By situating each objective within this combined framework, the table provides a visual guide to the study's scope, illustrating how it addresses ICT's role in enhancing educational outcomes in Thailand across different settings.

Table 1: Mapping Research Objectives to CIPO and MLFTAU Framework Dimensions

Objective	CIPO Framework \ Dimension	MLFTAU \ Dimension	Explanation
1. Assess the current state of ICT integration in Thai education and its alignment with national strategy aspirations.	Context, Input	Facilitating Conditions	Context: Evaluates external factors, such as socio-economic status and infrastructure availability, affecting ICT integration in and out of school. Input: Assesses resources (e.g., funding, technology, teacher training) provided to implement ICT in schools and its availability at home. Facilitating Conditions: Examines if infrastructure and support are in place to enable ICT adoption in both settings.
2. Identify the enablers and barriers to effective ICT use within educational settings.	Input, Process	Effort Expectancy, Social Influence, Facilitating Conditions	Input: Identifies resources and training available for ICT adoption in school. Process: Analyzes classroom ICT usage practices and obstacles. Effort Expectancy: Considers ease of ICT use for educators and students within the school. Social Influence: Examines peer or policy pressures to adopt ICT in classrooms. Facilitating Conditions: Identifies existing school-based support for effective ICT use.
3. Analyze the relationship between ICT utilization and student performance on established educational benchmarks, using Plausible Values (PVs) and appropriate statistical models.	Process, Output	Performance Expectancy	Process: Investigates ICT integration into teaching practices in school and the use of ICT for homework or self-directed learning at home. Output: Assesses the impact of ICT on academic performance both in-school (e.g., classroom engagement) and outside-the-classroom (e.g., homework effectiveness). Performance Expectancy: Analyzes if ICT is perceived to improve educational outcomes in both settings.

4. Provide evidence-based recommendations to policymakers for enhancing ICT acceptance and utilization in Thai education.	All (Context, Input, Process, Output)	All MLFTAU Dimensions	Context: Considers socio-economic and regional factors affecting ICT access in both school and home environments. Input: Recommends necessary resources and training in schools and for home access. Process: Suggests best practices for ICT integration in classrooms and strategies for encouraging productive ICT use at home. Output: Aims to improve student performance and equity through enhanced ICT policies across both in-school and outside-the-classroom environments. MLFTAU Dimensions: Provides recommendations to address all behavioral and structural factors affecting ICT adoption and effective u

The potential of this research extends beyond national boundaries, echoing the global challenges of digital technology integration within educational systems. By delineating Thailand's experience, insights of broader relevance are revealed, contributing to the international discourse on educational technology and its implementation in diverse cultural and economic contexts.

4 Methodology

This research employs a quantitative approach, leveraging secondary data from sources including the PISA assessments and national statistics to analyze the role of Information and Communications Technology (ICT) in shaping educational outcomes in Thailand. Two guiding frameworks are central to this study: the PISA ICT Framework, which is based on the CIPO model (Context, Input, Process, Output), and the Multi-Level Framework of Technology Acceptance and Use (MLFTAU).

The PISA ICT Framework provides a structural perspective on ICT integration within educational practices, examining contextual, input, process, and output factors. In contrast, the MLFTAU offers insights into the behavioral and psychological factors influencing ICT acceptance and usage among students and educators. Together, these frameworks allow for a comprehensive analysis of both systemic and individual factors affecting ICT's impact on education.

To comprehensively assess the impact of ICT on student performance, this study focuses on core variables such as ICT access, usage patterns, and educational outcomes. Analytical techniques, including regression models and hierarchical linear modeling, will be applied to examine these relationships, with Plausible Values (PVs) from PISA data utilized to account for measurement uncertainty in student performance. By addressing both systemic and individual dimensions, this methodology aims to generate nuanced and actionable insights, informing policymakers on effective strategies to enhance ICT integration and reduce educational disparities.

4.1 Data Sources

To conduct this analysis, the study leverages several key datasets.

Programme for International Student Assessment (PISA): Data from the PISA assessments for 2000, 2003, 2006, 2009, 2012, 2015, 2018, and 2022 will be utilized. These datasets provide student performance data in reading, mathematics, science, and creative thinking, as well as contextual information from student and school questionnaires. The study will focus on ICT-related variables as captured in the PISA surveys.

Thai National Educational Datasets: Additional data from the National Statistical Office of Thailand, the Information System for Equitable Education (iSEE 2.0), the Education Management Information System (EMIS), and the Management Information System (MIS) of NIETS will provide local context and complementary insights regarding ICT infrastructure and usage.

These datasets collectively enable a robust analysis of ICT's impact on educational outcomes across different contexts in Thailand. The PISA dataset, in particular, includes Plausible Values (PVs) and weights, ensuring that analysis reflects the complex sampling design used in the assessment.

4.2 Data Import and Setup

The PISA data for analysis is imported using the intsvy package, which is specifically recommended by the OECD for handling PISA data. This package is tailored for large-scale international assessments, allowing for the use of plausible values (PVs) and survey weights to account for the complex sampling design employed in these assessments. The PISA datasets used in this research include Thai-specific data from multiple cycles (e.g., 2000, 2003, 2006, 2009, 2012, 2015, 2018, and 2022), along with supplementary data from the National Statistical Office of Thailand and other relevant sources. Proper data import and preparation are essential to ensure that the analysis is based on clean, structured data.

The original Thailand-specific PISA dataset from PISA Thailand is provided in CSV format, which is compatible with intsvy and allows for efficient processing without further conversion. The CSV files are organized within the data/raw/ folder by cycle year, maintaining a clear and organized version history for multi-year analysis.

The following libraries are essential for importing and handling the PISA data and performing subsequent data manipulation:

```r
library(intsvy)     # For handling PISA data analysis, plausible values, and survey weights
library(readr)      # For reading CSV files
library(dplyr)      # For data manipulation
library(tidyr)      # For tidying data
library(ggplot2)    # For data visualization
```

This code block loads all the required packages, ensuring that the necessary tools for data handling and analysis are available.

To enable the loading of PISA data dynamically for each specified year, a function is created to construct the file path based on the year argument. This setup enables handling of data from multiple years without the need for hardcoded file paths.

```r
# Function to dynamically load the PISA data for a specified year
load_pisa_data <- function(year) {
    # Define file path dynamically
    csv_path <- paste0("data/raw/", year, "/pisa", year, "_data.csv")
    # Read the CSV
    pisa_data <- read_csv(csv_path)
    return(pisa_data)
}
```

To facilitate multi-year analysis, lists are created to store plausible values (PVs) and weight variables specific to each year. The example pv_list below contains the PVs for mathematics performance across different years, while the weights list holds the corresponding survey weights. This approach allows the code to dynamically select the appropriate variables based on the year being analyzed, ensuring flexibility and scalability.

```
# Example lists for PVs and weights by year, to facilitate multi-year analysis
pv_list <- list("2022" = c("PV1MATH", "PV2MATH", "PV3MATH", "PV4MATH", "PV5MATH"),
                "2018" = c("PV1MATH18", "PV2MATH18", "PV3MATH18", "PV4MATH18", "PV5MATH18"))
weights <- list("2022" = "W_FSTUWT", "2018" = "W_FSTUWT18")
```

To validate the successful import and basic structure of each dataset, a structural and summary analysis was conducted. This analysis included generating a summary and structure file for each dataset year, allowing a comprehensive examination of the dataset's layout and missing values. Due to the extensive nature of this output, the complete results are stored as separate files in the data/metadata/data_exploration/ subfolder, organized by year. This ensures that the RMarkdown document remains focused, while all necessary metadata files are accessible for reference. The following code provides an example of this structure generation process:

```
# Example: Generate structure and summary for 2022 and save in metadata folder
year <- "2022"
pisa_data <- load_pisa_data(year)

# Save structure as a .txt file
structure_file <- paste0("data/metadata/data_exploration/", year, "_structure.txt")
capture.output(str(pisa_data), file = structure_file)

# Save summary as a .csv file
summary_file <- paste0("data/metadata/data_exploration/", year, "_summary.csv")
write.csv(summary(pisa_data), summary_file)
```

In the next step, data for a specific year (in this example, 2022) is loaded using the load_pisa_data function. The gender variable (ST004D01T) is identified and converted to a factor with labels ("Female" and "Male") to facilitate analysis by gender. To confirm the successful processing of this variable, a count of the number of males and females is displayed. This serves as an initial data validation step, ensuring that the gender data is correctly labeled and ready for analysis.

```
# Load data for a specific year (2022) and validate gender data
year <- "2022"
pisa_data <- load_pisa_data(year)

# Convert gender variable (ST004D01T) to a factor with labels
if ("ST004D01T" %in% colnames(pisa_data)) {
    pisa_data$gender <- factor(pisa_data$ST004D01T, levels = c(1, 2), labels = c("Female", "Male"))
} else {
    stop("The 'ST004D01T' gender variable is missing in the data.")
}

# Display gender counts and verify they match total students
gender_counts <- table(pisa_data$gender)
cat("Gender Counts:\n")
```

```
## Gender Counts:
```

```
print(gender_counts)
```

```
##
## Female    Male
##   4451    4044
```

```r
total_gender_count <- sum(gender_counts)
total_students <- nrow(pisa_data)
cat(sprintf("\nTotal students (gender counts): %d\nTotal students in dataset: %d\n", total_gender_count
```

```
##
## Total students (gender counts): 8495
## Total students in dataset: 8495
```

```r
if (total_gender_count == total_students) {
    cat("The gender data matches the total number of students.\n")
} else {
    cat("Warning: Gender data does not match total students.\n")
}
```

```
## The gender data matches the total number of students.
```

After converting and counting the gender variable, the dataset dimensions are displayed, followed by a preview of the first 10 rows. These steps confirm the successful data import and provide an overview of the dataset's structure, ensuring that all necessary variables and observations are still present.

```r
# Get dimensions of the loaded data
data_dim <- dim(pisa_data)
cat(sprintf("The %s dataset contains %d rows and %d columns.\n", year, data_dim[1], data_dim[2]))
```

```
## The 2022 dataset contains 8495 rows and 1455 columns.
```

```r
# Display a preview of the first 10 rows
head(pisa_data, 10)
```

```
## # A tibble: 10 x 1,455
##      CNT   CNTRYID CNTSCHID CNTSTUID CYC   NatCen STRATUM SUBNATIO  OECD ADMINMODE
##      <chr>   <dbl>    <dbl>    <dbl> <chr> <chr>  <chr>      <dbl> <dbl>     <dbl>
##  1 THA       764 76400001 76400396 08MS  076400 THA97    7640000     0         2
##  2 THA       764 76400001 76400632 08MS  076400 THA97    7640000     0         2
##  3 THA       764 76400001 76400865 08MS  076400 THA97    7640000     0         2
##  4 THA       764 76400001 76400936 08MS  076400 THA97    7640000     0         2
##  5 THA       764 76400001 76401306 08MS  076400 THA97    7640000     0         2
##  6 THA       764 76400001 76401652 08MS  076400 THA97    7640000     0         2
##  7 THA       764 76400001 76401681 08MS  076400 THA97    7640000     0         2
##  8 THA       764 76400001 76401944 08MS  076400 THA97    7640000     0         2
##  9 THA       764 76400001 76402440 08MS  076400 THA97    7640000     0         2
## 10 THA       764 76400001 76402738 08MS  076400 THA97    7640000     0         2
## # i 1,445 more variables: Option_CT <lgl>, Option_FL <lgl>, Option_ICTQ <dbl>,
## #   Option_WBQ <dbl>, Option_PQ <dbl>, Option_TQ <dbl>, Option_UH <dbl>,
## #   BOOKID <dbl>, ST001D01T <dbl>, ST003D02T <dbl>, ST003D03T <dbl>,
## #   ST004D01T <dbl>, ST250Q01JA <dbl>, ST250Q02JA <dbl>, ST250Q03JA <dbl>,
## #   ST250Q04JA <dbl>, ST250Q05JA <dbl>, ST250D06JA <dbl>, ST250D07JA <dbl>,
## #   ST251Q01JA <dbl>, ST251Q02JA <dbl>, ST251Q03JA <dbl>, ST251Q04JA <dbl>,
## #   ST251Q06JA <dbl>, ST251Q07JA <dbl>, ST251D08JA <dbl>, ST251D09JA <dbl>, ...
```

The PISA 2022 Thailand dataset provides a structured array of variables categorized by prefixes to represent different dimensions of educational data. Key identifiers include variables with the CNT prefix, such as CNTRYID (Country Identifier), CNTSCHID (School ID), and CNTSTUID (Student ID), which are essential for distinguishing observations by country, school, and student.

Student responses (ST-prefixed variables) capture individual data, while ICT-related variables (IC prefix) provide insights into technology access and use. School questionnaire variables (SCH prefix) include institutional characteristics and policies that may influence educational outcomes. This foundational structure, combined with supplementary data from additional PISA cycles and the National Statistical Office, will enable a comprehensive examination of factors influencing educational outcomes in Thailand.

4.3 Data Preparation and Cleaning

Since this study focuses exclusively on Thailand, only Thailand-specific data files provided by PISA Thailand were used, eliminating the need to filter out data from other countries. This targeted dataset selection simplifies storage requirements and streamlines file handling, allowing for efficient data management across multiple PISA cycles (i.e., 2000, 2003, 2006, 2009, 2012, 2015, 2018, and 2022).

With the data imported and initial structure checks completed, the next steps involve preparing the data for analysis. This includes handling plausible values (PVs) and weights, addressing missing data, and performing any necessary transformations to ensure a clean and structured dataset.

Handling of Plausible Values (PVs) and Weights

The PISA dataset includes Plausible Values (PVs) for student achievement scores in reading, mathematics, science, and creative thinking. PVs are multiple imputed estimates of students' latent abilities rather than exact scores, which means they are essential for accurate statistical inferences and valid comparisons across student groups. Unlike single point estimates, PVs account for measurement error by providing a range of estimates, enhancing the validity of inferences drawn from these scores. For this study, PVs will be applied in all analyses involving student performance, aligning with best practices in large-scale assessments to ensure robust statistical conclusions.

Additionally, survey weights are provided in the PISA dataset to adjust for the complex sampling designs used, ensuring the sample accurately represents the broader population. Since PISA employs stratified sampling to capture diverse student populations, applying these weights corrects for selection probability differences, allowing for generalizable and unbiased parameter estimates. Weights are particularly critical in maintaining statistical rigor in this study, as they ensure that results are representative of Thailand's student population. In all analyses involving PVs, the corresponding survey weights will be applied to control for sampling biases and uphold the validity of statistical inferences.

Missing Data Handling

Missing data is a common issue in large-scale assessments like PISA, where students may skip questions or fail to complete the survey. Proper handling of missing data is crucial to maintain the representativeness of the sample and ensure the validity of the results. In this section, various methods are employed to address missing data, following OECD recommendations and best practices in educational research.

To ensure a thorough understanding of the missing data across multiple years of PISA data, missing data summaries were generated for each specified year using the coding below. For each variable, the percentage of missing values was calculated, with results saved in separate CSV files organized by year within the metadata/missing_data_summaries folder. Each file is named according to the format missing_pisa_data_summary_[year].csv, providing a structured overview of missing data without overloading this document with detailed output. Readers interested in the complete summaries for each year can refer to the CSV files in the scripts folder.

```
# List of PISA years to process
years <- c(2000, 2003, 2006, 2009, 2012, 2015, 2018, 2022)
```

```r
# Function to dynamically load PISA data for a given year
load_pisa_data <- function(year) {
    csv_path <- paste0("data/raw/", year, "/pisa", year, "_data.csv")
    pisa_data <- read.csv(csv_path)
    return(pisa_data)
}

# Loop through each year to calculate missing data summaries and save to metadata folder
for (year in years) {
    # Load data for the specific year
    pisa_data <- load_pisa_data(year)

    # Calculate percentage of missing data for each variable
    missing_data_summary <- sapply(pisa_data, function(x) sum(is.na(x)) / length(x) * 100)
    missing_data_summary <- data.frame(Variable = names(missing_data_summary),
                                       Missing_Percentage = missing_data_summary)

    # Define year-specific folder path in metadata/missing_data_summaries
    year_folder <- paste0("data/metadata/missing_data_summaries/", year)
    if (!dir.exists(year_folder)) {
        dir.create(year_folder, recursive = TRUE)
    }

    # Save missing data summary as a CSV file in the appropriate year subfolder
    summary_file <- paste0(year_folder, "/missing_pisa_data_summary_", year, ".csv")
    write.csv(missing_data_summary, summary_file, row.names = FALSE)
}
```

The code below processes the PISA data for each specified year, identifies and removes variables with 100%
missing data, and saves the cleaned datasets in the pre-existing year-specific folders within `data/processed/`.
Each cleaned file is named `cleaned_pisa_data_[year].csv`. This approach maintains an organized struc-
ture for multi-year data analysis, storing each year's cleaned data in its respective folder for easy reference
and subsequent analysis.

```r
# Load necessary libraries
library(dplyr)
library(readr)

# Function to dynamically load the PISA data for a specified year
load_pisa_data <- function(year) {
    # Define file path dynamically
    csv_path <- paste0("data/raw/", year, "/pisa", year, "_data.csv")
    # Read the CSV
    pisa_data <- read_csv(csv_path)
    return(pisa_data)
}

# List of years you have PISA data for
years <- c(2000, 2003, 2006, 2009, 2012, 2015, 2018, 2022)

# Loop through each year, load the data, clean it, and save the result
for (year in years) {
    # Load data for the specified year
```

```r
    pisa_data <- load_pisa_data(year)

    # Calculate the percentage of missing data for each variable
    missing_data_summary <- sapply(pisa_data, function(x) sum(is.na(x)) / length(x) * 100)
    missing_data_summary <- data.frame(Variable = names(missing_data_summary),
                                       Missing_Percentage = missing_data_summary)

    # Identify variables with 100% missing data
    high_missing_summary <- missing_data_summary %>%
      filter(Missing_Percentage == 100)

    # Remove variables with 100% missing data
    cleaned_data <- pisa_data %>% select(-one_of(high_missing_summary$Variable))

    # Save the cleaned dataset in the appropriate processed year folder
    output_path <- paste0("data/processed/", year, "/pisa", year, "_cleaned_stage1.csv")
    write_csv(cleaned_data, output_path)

    # Print a message confirming the process for each year
    cat(sprintf("Year %d: Removed %d variables with 100%% missing data. Cleaned data saved to %s\n",
                year, nrow(high_missing_summary), output_path))
}
```

```
## Year 2000: Removed 0 variables with 100% missing data. Cleaned data saved to data/processed/2000/pis
## Year 2003: Removed 0 variables with 100% missing data. Cleaned data saved to data/processed/2003/pis
## Year 2006: Removed 0 variables with 100% missing data. Cleaned data saved to data/processed/2006/pis
## Year 2009: Removed 0 variables with 100% missing data. Cleaned data saved to data/processed/2009/pis
## Year 2012: Removed 0 variables with 100% missing data. Cleaned data saved to data/processed/2012/pis
## Year 2015: Removed 1 variables with 100% missing data. Cleaned data saved to data/processed/2015/pis
## Year 2018: Removed 20 variables with 100% missing data. Cleaned data saved to data/processed/2018/pi
## Year 2022: Removed 137 variables with 100% missing data. Cleaned data saved to data/processed/2022/p
```

For a complete list of variables with high missing percentages, please refer to the supplementary file Missing Data Summary.

Variable Transformation

Detail any modifications or transformations applied to the variables. Provide the R code for these transformations.

Descriptive Analysis

Present basic statistics or visualizations to describe the data. Include the R code generating these descriptive statistics.

4.4 Variable Selection

The selection of variables for analysis in this study was conducted with a systematic approach rooted in the theoretical foundations as stipulated by the PISA ICT Framework and the MLFTAU. Each chosen variable reflects a deliberate intersection of these models, aiming to dissect the intricate dynamics of ICT utilization within educational domains both inside and beyond the classroom setting.

Variables were initially screened for their relevance to the core constructs of the study's guiding frameworks. The PISA ICT Framework provides a dichotomy of in-school factors (such as ICT curriculum integration, availability of technology, and teacher ICT proficiency) and outside-the-classroom influences (encompassing

access to technology at home and students' ICT engagement outside school hours). Similarly, the Multi-Level Framework of Technology Acceptance and Use (MLFTAU) offered a layered perspective on ICT adoption, spanning individual, organizational, and environmental influences.

An extensive review of current literature was performed to identify variables consistently linked with meaningful ICT integration outcomes. These variables were cross-referenced with the factors outlined in the theoretical frameworks to ensure their pertinence to the study's focus on technology acceptance and educational effectiveness.

A meticulous examination of the PISA 2022 dataset was undertaken to determine the availability of potential variables. This step was crucial in refining the variable pool to include only those with robust data support and exclude variables with extensive missing data or inadequate measurement reliability.

The final assortment of variables was curated to ensure tangible linkages to policy implications and educational practice. This was underpinned by the consideration of variables that can inform evidence-based interventions and strategic educational planning.

The selected variables encompass a comprehensive array of factors relating to the PISA ICT framework and the MLFTAU:

Table 2: Key Constructs of the PISA ICT Framework - In-School Dimension

Dimension	Construct	Definition	Relevance
Context	Socioeconomic Status	The economic and social position of students and their families, influencing access to and use of ICT.	Provides background for understanding ICT access disparities.
	School Location	The geographical location of the school (urban or rural) and its impact on ICT resources.	Helps identify regional differences in ICT integration.
	Cultural Attitudes	Cultural beliefs and attitudes towards technology and education at school.	Offers insight into the social context of ICT use at school.
Input	ICT Resources	Availability and quality of technology and digital tools in educational settings at school.	Assesses the infrastructure for ICT integration at school.
	Teacher Training in ICT	Extent of teacher preparation for integrating ICT into teaching practices at school.	Evaluates the readiness of educators for ICT use at school.
	School Policies on ICT	Policies and guidelines regarding ICT use in education set by schools or authorities.	Indicates the regulatory environment for ICT at school.
	Student Engagement (Attitudes)	Students' perceptions, beliefs, and feelings towards using ICT in their learning environment at school.	Affects the willingness to engage with ICT and can significantly influence the acceptance and effective use of technology in education.
Process	ICT Use in Teaching	How teachers incorporate ICT into their teaching methods and curriculum at school.	Reflects the application of ICT in instructional practices at school.
	Pedagogical Approaches with ICT	The use of ICT to support various teaching strategies, such as student-centered learning or collaborative learning.	Assesses how ICT is integrated into pedagogical practices to enhance teaching effectiveness.

	Student Engagement (Time)	The amount of time that students spend using ICT for learning activities during school hours.	Indicates the level of integration of ICT in daily educational activities and can be a predictor of skill acquisition and comfort with technology.
	Student engagement (Attention)	The degree of students' focus or concentration when using ICT for learning purposes in school.	Reflects the quality of ICT engagement and may correlate with deeper learning and retention of information.
	Student engagement (Effort)	The amount of energy or diligence that students apply to ICT-related tasks during school-based learning.	Suggests how motivated students are to use ICT resources for learning and may affect the outcomes of technology-enhanced education.
Output	Educational Outcomes	Impact of ICT on students' academic performance in areas such as reading, mathematics, and science at school.	Evaluates the effectiveness of ICT in enhancing learning at school.
	ICT Competencies	Skills and abilities of students to effectively use ICT for educational purposes at school.	Assesses the development of digital literacy and competencies at school.

Table 3: Key Constructs of the PISA ICT Framework - Outside-the-Classroom Dimension

Dimension	Construct	Definition	Relevance
Context	Socioeconomic Status	The economic and social position of students and their families, influencing access to and use of ICT.	Provides background for understanding ICT access disparities.
	Cultural Attitudes	Cultural/family beliefs and attitudes towards technology and education out of school.	Offers insight into the social and family context of ICT use out of school.
Input	Access to ICT	Availability of ICT resources and tools for students outside the school environment.	Assesses opportunities for students to engage with ICT for learning beyond the school setting.
	Parental/Teacher Supervision in Using ICT	The encouragement and resources provided by parents and teachers that facilitate the use of ICT for learning at home.	Assesses the home learning context, as parental and teacher involvement can significantly influence students' ability and motivation to use ICT for educational purposes outside of school.
Process	ICT Use for Learning (Homework-Related Activities)	Extent of student use of ICT for homework-related learning activities outside the classroom.	Measures student engagement with ICT for structured educational purposes out of school.

	ICT Use for Learning (Self-Directed Learning Activities)	Extent of student use of ICT for self-directed learning activities outside the classroom.	Measures student engagement with ICT for unstructured educational purposes out of school.
	ICT Use for Leisure	Extent of student use of ICT for leisure activities outside the classroom.	Measures student engagement with ICT for leisure purposes out of school.
Output	Educational Outcomes	Impact of ICT on students' academic performance in areas such as reading, mathematics, and science out of school.	Evaluates the effectiveness of ICT in enhancing learning out of school.
	Student Well-being	Influence of ICT use on students' emotional and psychological well-being out of school.	Examines the broader effects of ICT on student welfare out of school.

Table 4: Key Constructs of the Multi-Level Framework of Technology Acceptance and Use (MLFTAU)

Dimension	Construct	Definition	Relevance
Individual	Performance Expectancy (PE)	The degree to which an individual believes that using the technology will help them to attain gains in job performance.	Assesses how students perceive the use of ICT in improving their learning outcomes.
	Effort Expectancy (EE)	The degree of ease associated with the use of the technology.	Assesses how the perceived ease of using ICT impacts students' willingness to engage with it for educational purposes.
	Social Influence (SI)	The degree to which an individual perceives that important others believe they should use the new technology.	Assesses the impact of peer and parental influence on students' use of ICT in education.
	Facilitating Conditions (FC)	The degree to which an individual believes that an organisational and technical infrastructure exists to support the use of the technology.	Assesses the availability of ICT resources and support in schools that enable students to use ICT effectively.
	Hedonic Motivation	The pleasure or pain derived from using the technology.	Assesses how the enjoyment of using ICT influences students' engagement with technology for learning.
	Price Value	The individual's cognitive trade-off between the perceived benefits of the technology and its monetary cost.	Although this construct may be more relevant in consumer settings, it could be adapted to assess the perceived value of investing time and effort in using ICT for educational purposes.
	Habit	The extent to which individuals tend to perform behaviours automatically because of learning.	Assesses how habitual use of ICT outside of school settings influences students' use of ICT for learning.

Group	Group Norms	The shared expectations about behaviour within the group.	Assesses how classroom or school norms regarding ICT use impact students' attitudes and behaviours.
	Group Cohesion	The extent to which group members stick together and remain united in the pursuit of common goals.	Assesses how collaborative learning environments using ICT influence educational outcomes.
	Team Climate for Innovation	The shared perceptions of the group's practices, procedures, and behaviours that support innovation.	Assesses how a school's climate for innovation affects the adoption and effective use of ICT in education.
Organisation	Organisational Facilitating Conditions	The degree to which an organisation provides the infrastructure and support for the use of the technology.	Assesses how the school's ICT infrastructure and policies support or hinder the integration of ICT in education.
	Organisational Culture	The shared values, beliefs, and practices within an organisation that shape behaviour and attitudes toward technology use.	Assesses the school's culture regarding how technology can provide insights into the systemic factors influencing ICT integration in education.
	Leadership Support	The extent to which leaders within the organisation encourage and support the use of technology.	Assesses the role of school leadership in promoting or impeding the use of ICT for educational purposes.
Cross-Level Constructs	Cross-Level Interactions	The interactions between constructs at different levels that influence technology acceptance and use.	Assesses how individual attitudes interact with group norms or organisational culture to impact ICT use in educational settings.
Outcomes (Output)	Educational Outcomes	Impact of ICT on students' academic performance in areas such as reading, mathematics, and science out of school.	Evaluates the effectiveness of ICT in enhancing learning out of school.
	Student Well-being	Influence of ICT use on students' emotional and psychological well-being out of school.	Examines the broader effects of ICT on student welfare out of school.

The following table provides an example of how In-School Input variables are mapped to the research objectives. By focusing on key resources and conditions within schools—such as ICT availability, teacher training, and internet connectivity—this mapping demonstrates how each variable supports an in-depth understanding of ICT integration at the school level. This structured selection ensures that each variable is both theoretically grounded in the PISA ICT framework and practically relevant to evaluating in-school ICT infrastructure's role in educational outcomes.

Table 5: In-School ICT Input Variables: Alignment with Research Objectives on ICT Integration in Education

Framework Construct	Variable	Description	Data Source	Relevance to Research Objectives

Input (In-School)	ICT Availability	Number and quality of computers and digital devices available in classrooms	PISA 2022	Assesses the basic ICT infrastructure in schools, crucial for equitable access and effective ICT integration
	Teacher ICT Training	Level of teacher training or professional development in ICT skills	PISA 2022	Evaluates teacher readiness to implement ICT in teaching, influencing student engagement with technology
	Digital Learning Resources	Access to digital resources like educational software and online tools in schools	PISA 2022	Measures the variety of resources available to support diverse ICT-based teaching methods
	Internet Connectivity	Quality and speed of internet access within the school environment	PISA 2022	Determines the school's capacity to support internet-based learning, a critical component of ICT integration
	Technical Support for ICT	Availability of dedicated technical support staff for ICT maintenance	PISA 2022	Provides insight into the ongoing support for ICT, which affects sustainability and reliability of ICT resources

This systematic and rigorous variable selection protocol ensures that my analysis is anchored in empirical evidence, is methodologically sound, and aligns seamlessly with my research aims.

4.5 Analytical Methods

In the example analysis below, the plausible values and weight for the specified year are accessed from the lists, and the average mathematics performance is calculated by gender, demonstrating the practical use of the data and lists.

5 Load necessary library

library(intsvy)

6 Function to load cleaned PISA data for a specified year

load_cleaned_pisa_data <- function(year) { # Define file path dynamically for the processed folder with the correct naming format csv_path <- paste0("data/processed/", year, "/pisa", year, "_cleaned.csv") # Read the CSV pisa_data <- read_csv(csv_path) return(pisa_data) }

7 Specify the year of analysis

year <- "2022"

8 Load the cleaned PISA data for the specified year

pisa_data <- load_cleaned_pisa_data(year)

9 Access PVs and weight variable for the specified year

pvs <- pv_list[[year]] weight_var <- weights[[year]]

10 Ensure 'gender' is properly set as a factor in the data

if ("gender" %in% colnames(pisa_data)) { pisa_data$gender$ $<-factor(pisa_data$gender, $levels = c("Female", "Male")) } else { stop("The 'gender' variable is missing in the data.") }

11 Example analysis: Calculate average mathematics performance by gender

math_results <- pisa.mean.pv(pvlabel = pvs, by = "gender", weight = weight_var, data = pisa_data)

12 Display results

print(math_results)

The example below calculates the average mathematics performance by gender, using PVs and the appropria

```
# Load necessary library
library(intsvy)

# Define plausible values and weight variable
pvs_math <- c("PV1MATH", "PV2MATH", "PV3MATH", "PV4MATH", "PV5MATH")
weight_var <- "W_FSTUWT"

# Example analysis using plausible values and weights
# Average mathematics performance by gender
math_results <- pisa.mean.pv(pisa_data, pvlabel = pvs_math, by = "gender")

# Display results
print(math_results)
```

A range of statistical techniques will be applied to analyze the data.

Descriptive Statistics: To provide an overview of key variables related to ICT use and student performance.

Hierarchical Linear Modeling (HLM): This technique will be used to account for the nested structure of the PISA data (students within schools). HLM models will help analyze the relationships between ICT access, school characteristics, and student performance.

Weighted Linear Regression: This will assess the impact of ICT use on student outcomes while accounting for the survey's sampling design.

Plausible Values (PVs): PVs will be used to represent student performance across multiple dimensions, accounting for uncertainty in test scores. The EdSurvey package will be utilized to handle PVs in analysis.

12.1   Model Validation

Validation techniques, such as cross-validation and sensitivity analysis, will be employed to ensure the robustness of the models. This includes testing for assumptions, identifying potential biases, and assessing model stability.

13   Addressing Validity and Reliability

The study ensures both validity and reliability through carefully designed data handling procedures, the use of well-established data sources, and robust analytical techniques. Several strategies are employed to maintain the highest standards of academic rigor:

13.1   Validity

Content Validity: The selection of variables is closely aligned with the CIPO model and the Multi-Level Framework of Technology Acceptance and Use (MLFTAU), ensuring that the constructs being measured accurately reflect the key dimensions of ICT integration and educational outcomes.

Construct Validity: All key variables, such as ICT access, ICT utilization, and student performance, are derived from reliable and validated instruments within the PISA datasets and national databases, providing confidence that these constructs are measured as intended.

Internal Validity: Potential confounding variables (e.g., socio-economic status, school infrastructure) are controlled for using appropriate statistical techniques such as Hierarchical Linear Modeling (HLM) and weighted regression, minimizing the impact of extraneous factors on the results.

External Validity: The use of national-level PISA data and a well-defined sampling framework ensures that findings are generalizable to the broader population of Thai students and schools, as well as potentially applicable to other contexts with similar educational environments.

13.2   Reliability

Measurement Reliability: The PISA dataset, along with other national data sources, undergoes thorough validation and testing to ensure consistency and accuracy in the measurement of variables. The use of Plausible Values (PVs) further strengthens reliability by accounting for the uncertainty inherent in student performance data.

Reproducibility: All steps of the analysis are fully documented in this RMarkdown document, ensuring that the code, data processing steps, and statistical analyses are transparent and reproducible. This includes the use of the EdSurvey package to manage the complex sampling design and PVs, ensuring that future researchers can replicate the analysis.

Analytical Techniques: Established statistical methods (e.g., HLM, weighted linear regression) are applied rigorously, with models validated through techniques such as cross-validation and sensitivity analysis to check for consistency and robustness.

13.3 Documentation and Transparency

All R code, data manipulations, and analysis steps are documented in this RMarkdown file, ensuring that the research process is transparent and easily replicable by others. The document adheres to best practices for reproducible research, including version control through GitHub, ensuring that any future updates or revisions are traceable.

By maintaining strict adherence to these principles, the study produces results that are both accurate and generalizable, with confidence in the reliability and validity of the findings. The combination of robust data sources, appropriate statistical techniques, and clear documentation ensures the integrity and reproducibility of the research.

14 Ethical Considerations

This research adheres to the highest standards of ethical conduct as outlined by international and institutional guidelines for research involving human data. Several key ethical principles guide the study to ensure respect for individuals, data confidentiality, and the broader sociocultural context:

14.1 Data Confidentiality and Privacy

The data used in this study, including individual-level PISA data, are anonymized to protect the identity and privacy of students and schools. No personally identifiable information (PII) is used in the analysis, and data are handled in accordance with national and international data protection regulations (e.g., GDPR where applicable).

Strict access control measures are in place to ensure that the data is only accessed by authorized personnel for the purposes of this research. The dataset is stored securely, and all analyses are conducted in a manner that prevents the identification of individual respondents.

The use of Plausible Values (PVs) in the PISA dataset enhances privacy by ensuring that no single score can be directly attributed to a specific student, thereby preserving confidentiality.

14.2 Transparency and Honesty in Reporting

All findings are reported accurately and without manipulation or bias. The methodology is fully transparent, allowing for the reproducibility and scrutiny of results. Potential limitations or biases in the data are acknowledged and addressed, ensuring that the findings contribute meaningfully to both academic knowledge and policy discussions.

In line with best practices, the study commits to sharing anonymized data, code, and methodology transparently via platforms such as GitHub, allowing for peer review and replicability of the research.

14.3 Respect for Sociocultural Context

The study recognizes the importance of Thailand's unique sociocultural and educational landscape. Special attention is paid to ensuring that the research does not reinforce inequities or biases inherent in the data, particularly with respect to disadvantaged and marginalized groups.

Ethical considerations include the contextual sensitivity of findings, ensuring that policy recommendations are culturally appropriate and do not disproportionately disadvantage any group. The study aims to contribute positively to the development of Thailand's education system by offering evidence-based, inclusive policy suggestions.

14.4   Compliance with Ethical Review Boards

The study has been reviewed and approved by the Internal Review Board (IRB) of the researcher's institution, ensuring that it meets ethical standards regarding data use, privacy, and the handling of sensitive information.

In line with institutional and international ethical guidelines (e.g., APA and the Declaration of Helsinki), the research adheres to principles of respect for persons, beneficence, and justice, ensuring that all steps in the research process protect the dignity and rights of individuals represented in the data.

14.5   Policy and Societal Implications:

The study is designed with careful consideration of its potential policy implications. Recommendations made as a result of the findings aim to improve educational outcomes in Thailand and to close the digital divide, particularly for marginalized communities. Ethical responsibility is taken to ensure that these recommendations are grounded in evidence and are not used in ways that could negatively impact vulnerable populations.

By adhering to these ethical principles, the research not only ensures compliance with data protection and privacy standards but also contributes positively to educational policy and practice in Thailand in a responsible and culturally sensitive manner.