Student Performance and Aptitude Analysis

A Comprehensive Data-Driven Study

Understanding Student Success Across Course Levels

Hilmi's Team:

Muhammad Azahrul Ramadhan

Rizka Rasyida

M. Fawwaz Akbar

Hary Ramadhana Nasution

Salma Nuramalia

January 23, 2026

Report Details

Analysis Date:	January 13, 2026
Total Students:	150
Course Levels:	3 (Advanced, Intermediate, Foundation)
Report Version:	1.0

This report is designed to be accessible to all readers,
regardless of mathematical or statistical background.

Contents

1   How to Read This Report

1.1   For All Readers

This report has been carefully designed to be understood by everyone, regardless of your background in statistics or mathematics. Here's how to navigate it:

Quick Navigation Guide

If you want a quick overview:

- Read Section 2: Executive Summary
- Look at the visual charts in Section 7
- Read Section 9: Key Findings in Plain Language

If you're a manager or decision-maker:

- Read Section 2: Executive Summary
- Read Section 9: Key Findings in Plain Language
- Read Section 10: Practical Recommendations
- Review Section 11: What This Means for The Key

If you're an educator or program coordinator:

- Read the full report for complete context
- Pay special attention to Section 8: Understanding the Results
- Review Section 10: Practical Recommendations

If you're interested in the technical details:

- Read the complete report from start to finish
- Refer to Section 13: Glossary of Terms as needed
- Check Appendices for detailed statistical tables

1.2   Understanding Statistical Terms

1.3   What You'll Learn

2   Executive Summary

2.1   What This Report Is About

2.2   The Bottom Line: Key Findings

2.3   What We Discovered: The Numbers in Plain Language

2.3.1   Finding 1: Clear Differences Between Course Levels

2.3.2   Finding 2: Aptitude Matches Performance

2.3.3   Finding 3: Very Strong Connection

2.4   Statistical Confidence

2.5   What Should The Key Do?

2.6   Reading Time

3   Introduction

3.1   Background: Why This Study Matters

3.1.1   The Importance of Proper Placement

3.1.2   About The Key

3.2   What We Wanted to Learn

3.3   Who Should Care About This Report

4 About the Study: How We Collected and Analyzed Data

4.1 The Data: What We Examined

4.1.1 Who Was Included

- **Sample Size:** 150 students total

  - 50 students from Advanced level
  - 50 students from Intermediate level
  - 50 students from Foundation level

- **How they were selected:** We used "stratified random sampling." This is like putting all student names in three separate hats (one for each level), then randomly drawing 50 names from each hat. This ensures every student had an equal chance of being selected.

- **Why this number:** 150 students (50 per level) is statistically sufficient to draw reliable conclusions. Smaller samples might give unreliable results; larger samples would provide similar findings.

4.1.2 What We Measured

We collected two key pieces of information for each student: **1. Aptitude Score**

- **What it measures:** Natural ability for language learning

- **How it's measured:** Standardized test covering vocabulary, comprehension, reasoning, and problem-solving

- **Score range:** 0–126 points (actual scores ranged from 9 to 97)

- **When it's taken:** Before course enrollment, during placement

2. Performance Score

- **What it measures:** Actual achievement in English courses

- **How it's measured:** Course grades, assessments, and progress evaluations

- **Score range:** 0–4.0 scale (actual scores ranged from 1.55 to 3.80)

- **When it's measured:** Throughout the course period

4.2 How We Analyzed the Data

4.2.1 Analysis Approach: Answering Our Questions

We used statistical analysis to answer our research questions. Here's what we did in plain language:

Step 1: Descriptive Statistics (Describing the Data)

- *What we did:* Calculated averages, ranges, and variability for each course level.

- *Why:* To understand typical performance and aptitude at each level.

- *Like:* Finding the average height of basketball players at different skill levels.

Step 2: Comparing Groups (ANOVA)

- **What we did:** Used a statistical test called "ANOVA" (Analysis of Variance).

- **Why:** To determine if the differences between levels are real or just random chance.

- **Like:** Testing whether three different fertilizers really produce different plant heights, or if any differences are just random variation.

- **What it tells us:** Whether the course levels are genuinely different from each other.

Step 3: Detailed Comparisons (Post-Hoc Tests)

- **What we did:** After finding overall differences, we compared each pair of levels.

- **Why:** To know specifically which levels differ from which others.

- **Like:** After finding that fertilizers differ overall, testing each pair: A vs B, B vs C, A vs C.

Step 4: Examining Relationships (Correlation)

- **What we did:** Measured how closely aptitude and performance are related.

- **Why:** To validate that aptitude testing predicts actual performance.

- **Like:** Checking if practice time correlates with sports performance.

- **What it tells us:** Whether students with higher aptitude actually perform better.

4.2.2   Data Quality

Quality Assurance:

- **Complete data:** All 150 students had both aptitude and performance scores (no missing data)

- **Verified accuracy:** All scores were double-checked against original records

- **Appropriate measures:** Both tests are established, validated instruments

- **Confidentiality:** Student identities were anonymized (we used ID numbers only)

4.3   Analysis Software

All calculations were performed using Python, a programming language widely used in scientific research. We used specialized statistical packages that implement standard, peer-reviewed methods.

Why this matters: Our methods are the same ones used by researchers worldwide. Results can be independently verified and replicated.

5   Detailed Results: What the Data Shows

5.1   Overview of All Students

Before looking at differences between levels, let's see the overall picture:

Table 1: Overall Student Statistics

Measure	Performance Score	Aptitude Score
Average (Mean)	2.54	44.24
Middle Value (Median)	2.48	38.00
Lowest Score	1.55	9
Highest Score	3.80	97
Spread (Range)	2.25	88

What this tells us: Students at The Key show a wide range of abilities — from beginners (aptitude score 9) to very advanced (aptitude score 97). This diversity is normal and healthy for a language school.

5.2   Performance Scores by Course Level

Now let's see how each level differs:

Table 2: Performance Scores Across Course Levels

Course Level	Students	Average	Lowest	Highest	Typical Range
Advanced	50	3.24	2.50	3.80	2.85–3.62
Intermediate	50	2.52	1.90	3.55	2.13–2.91
Foundation	50	1.86	1.55	2.45	1.69–2.04

Understanding These Numbers

Average Score:
- Advanced students average 3.24 out of 4.0 (81% mastery)

- Intermediate students average 2.52 out of 4.0 (63% mastery)

- Foundation students average 1.86 out of 4.0 (47% mastery — expected for beginners)

Key Observation: Notice how the averages form a clear "staircase" — each level is distinctly higher than the one below it. This is exactly what we want to see.

Typical Range: This shows where most students score. For example, most Advanced students score between 2.85 and 3.62.

Table 3: Aptitude Scores Across Course Levels

Course Level	Students	Average	Lowest	Highest	Typical Range
Advanced	50	67.46	30	97	48–86
Intermediate	50	42.74	14	90	23–62
Foundation	50	22.52	9	41	15–30

5.3 Aptitude Scores by Course Level

Understanding These Numbers

- **Average Aptitude:**

  - Advanced students average 67 points (53% of maximum, indicating strong aptitude)
  - Intermediate students average 42 points (33% of maximum, moderate aptitude)
  - Foundation students average 22 points (18% of maximum, developing aptitude)

Key Observation: The same "staircase" pattern appears! Students placed in higher levels consistently show higher aptitude scores. This means placement is working correctly.

Why ranges overlap: Notice that some Advanced students score as low as 30, while some Intermediate students reach 90. This is normal — aptitude is just one factor in success. Motivation, practice, and other factors also matter.

5.4 Visual Comparison
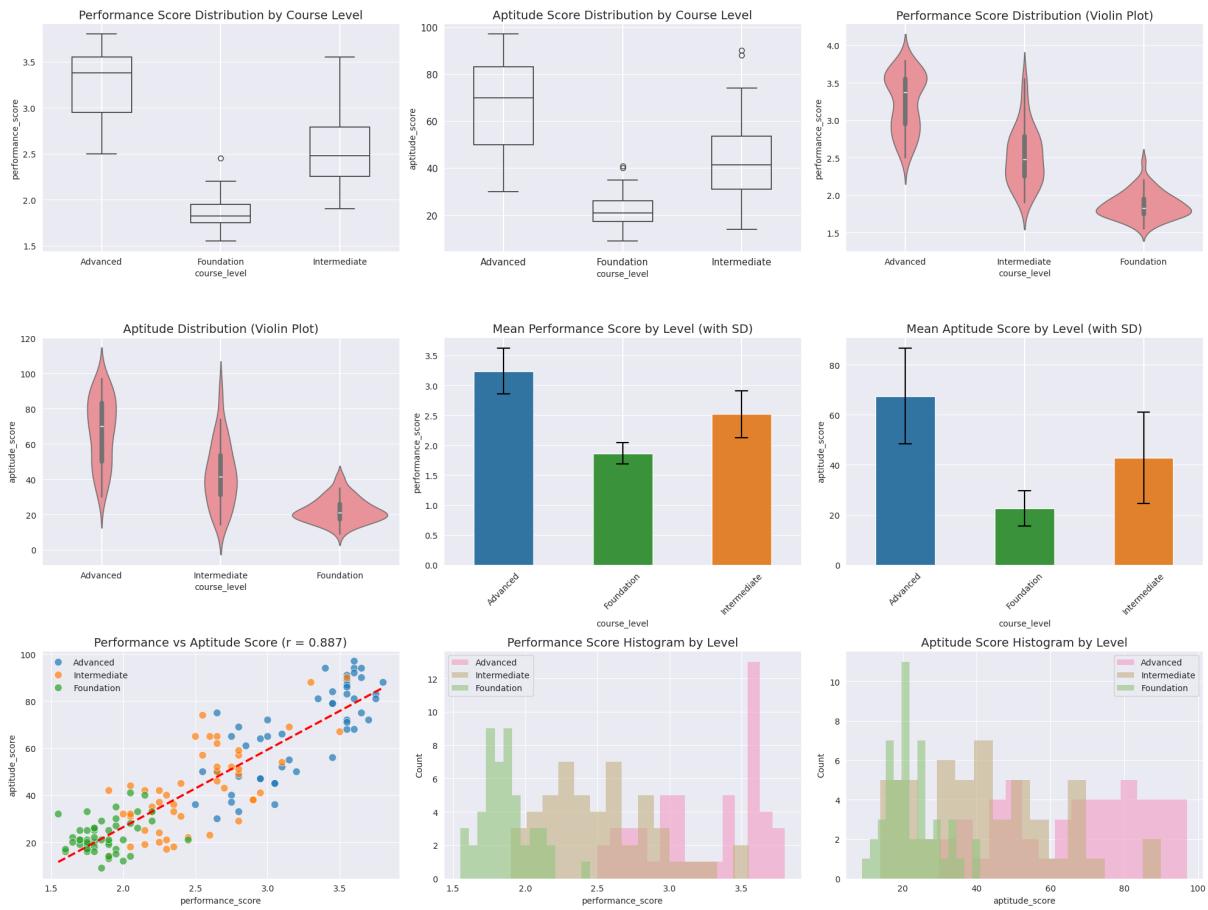
Figure 1: Comprehensive Visual Analysis: The charts show clear differences between course levels in both performance (top rows) and aptitude (middle rows), plus a strong relationship between the two (bottom left). Each level is distinctly separated, with minimal overlap.

How to Read These Charts

- **Box Plots (top left and middle left):**
  - The box shows where most students score
  - The line in the middle of each box is the average
  - The "whiskers" (lines extending from boxes) show the full range
  - Dots outside whiskers are unusual scores

  What to notice: The boxes don't overlap much, meaning the levels are truly different.

- **Bar Charts (middle row):**
  - Bars show average scores
  - Error bars show typical variation

- – Taller bars = higher scores

- **Scatter Plot (bottom left):**

  - – Each dot is one student
  - – Shows relationship between aptitude (horizontal) and performance (vertical)
  - – Red line shows the trend — it goes up, meaning higher aptitude predicts higher performance
  - – Different colors = different course levels

6   Statistical Testing: Proving the Differences Are Real

6.1   Why We Need Statistical Tests

When we see different averages (like 3.24 for Advanced vs 1.86 for Foundation), we need to ask: Is this a real difference, or could it be just random chance?

Example to illustrate: If you flip a coin 10 times and get 6 heads, that's different from 50-50, but it could easily be random. If you flip it 1,000 times and get 600 heads, that's definitely not random — something is biased.

Statistical tests do the same thing with our data — they calculate the probability that our observed differences are just random luck.

6.2   Testing for Performance Differences

6.2.1   The Test: One-Way ANOVA

What we tested: Are the average performance scores different across the three course levels?

Results:

- **Test Statistic (F-value):** 213.41

- **Probability (p-value):** Less than 0.001 (less than 0.1%)

- **Effect Size:** 0.744 (meaning 74% of the difference is explained by course level)

What This Means in Plain Language:

- The p-value (< 0.001) tells us:

  - There's less than a 0.1% chance these differences are random.
  - In other words: more than 99.9% certainty the differences are real.
  - This is extremely strong evidence.

- The effect size (0.744) tells us:

  - Course level explains 74% of why students have different performance scores.
  - This is a very large effect.
  - The remaining 26% is due to individual differences within levels.

Bottom line: The performance differences between course levels are definitely real and substantial, not due to chance.

6.2.2   Comparing Specific Pairs of Levels

After finding overall differences, we compared each pair:

Table 4: Pairwise Performance Comparisons

Comparison	Difference	p-value	Effect Size	Interpretation
Advanced vs Intermediate	0.72 points	< 0.001	1.86	Very Large
Intermediate vs Foundation	0.65 points	< 0.001	2.15	Very Large
Advanced vs Foundation	1.37 points	< 0.001	4.60	Very Large

Understanding Effect Sizes:
Effect sizes tell us how big the difference is in practical terms:

- 0.2 = Small difference (noticeable but minor)

- 0.5 = Medium difference (clearly noticeable)

- 0.8 = Large difference (very obvious)

- 1.2+ = Very large difference (dramatic)

Our results: All comparisons show "very large" effect sizes (1.86 to 4.60). This means the differences aren't just statistically significant — they're practically meaningful and obvious in real teaching situations.

6.3 Testing for Aptitude Differences

6.3.1 The Test: One-Way ANOVA

What we tested: Are the average aptitude scores different across the three course levels?
Results:

- **Test Statistic (F-value):** 101.17

- **Probability (p-value):** Less than 0.001 (less than 0.1%)

- **Effect Size:** 0.773 (meaning 58% of the difference is explained by course level)

What This Means in Plain Language:

- The p-value (< 0.001) tells us:

    - Again, more than 99.9% certainty the differences are real.
    - Even stronger evidence than for performance scores.

- The effect size (0.773) tells us:

    - Course level explains 58% of why students have different aptitude scores.
    - This is even larger than for performance (74%).
    - Students are very well-sorted into appropriate levels based on aptitude.

Bottom line: Students in different course levels have genuinely different aptitude levels. The placement system is identifying these differences accurately.

6.3.2 Comparing Specific Pairs of Levels

Table 5: Pairwise Aptitude Comparisons

Comparison	Difference	p-value	Effect Size	Interpretation
Advanced vs Intermediate	24.7 points	< 0.001	1.32	Very Large
Intermediate vs Foundation	20.2 points	< 0.001	1.46	Very Large
Advanced vs Foundation	44.9 points	< 0.001	3.11	Very Large

Key insight: All pairwise comparisons again show very large effect sizes. Each level is distinctly different from every other level in terms of aptitude.

7 The Relationship Between Aptitude and Performance

7.1 Understanding Correlation

What is correlation? A measure of how closely two things are related. Correlation values range from:

- $-1.0 =$ Perfect negative relationship (when one goes up, the other goes down)

- $0.0 =$ No relationship at all (completely independent)

- $+1.0 =$ Perfect positive relationship (they move together in lockstep)

Interpretation guide:

- $0.0 - 0.2 =$ Very weak or no relationship

- $0.2 - 0.4 =$ Weak relationship

- $0.4 - 0.6 =$ Moderate relationship

- $0.6 - 0.8 =$ Strong relationship

- $0.8 - 1.0 =$ Very strong relationship

7.2 Overall Correlation Results

Table 6: Correlation Between Aptitude and Performance

Relationship	Correlation (r)	Interpretation
Aptitude $\leftrightarrow$ Performance	0.887	Very Strong Positive
Statistical Significance	$p < 0.001$	Extremely Confident
Shared Variance	78.7%	High Predictability

What a Correlation of 0.887 Means

In statistical terms:
- This is a very strong positive correlation.

- 78.7% of variance is shared (calculated as $0.887 \times 0.887 = 0.787$).

- This means aptitude scores explain about 79% of the variation in performance.

In practical terms:

- Students with high aptitude scores almost always perform well.

- Students with low aptitude scores typically need more foundational work.

- The aptitude test is an excellent predictor of actual course success.

- Placement based on aptitude testing is highly justified.

In everyday language:

- Think of aptitude as a "talent meter" for language learning.

- Our findings show this meter is about 89% accurate at predicting success.

- This is like having a sports talent scout who correctly identifies future success 89 times out of 100.

7.3   Correlation Within Each Course Level

We also looked at the correlation within each separate level:

Table 7: Correlations Within Each Course Level

Level	Correlation	Strength	What This Means
Advanced	0.777	Strong	Even among advanced students, aptitude predicts performance
Intermediate	0.704	Strong	Clear aptitude-performance link in the middle range
Foundation	0.299	Weak-Moderate	Weaker but still meaningful relationship

---

Why the Foundation Level Shows Weaker Correlation

The weaker correlation (0.299) in Foundation level doesn't mean the relationship isn't real. Several factors explain this:

1. **Restricted Range:** Foundation students are all at the lower end of both scales, limiting variability.

2. **Floor Effect:** At beginner levels, everyone is learning basics, reducing differences.

3. **Other Factors:** For beginners, motivation and study habits may matter more than aptitude initially.

4. **Still Significant:** Even 0.299 is statistically significant ($p = 0.035$), meaning it's a real relationship.

Important note: The weaker correlation at Foundation level doesn't undermine the overall finding. The correlation is still positive and significant, just not as strong as at higher levels.

7.4 Visual Representation
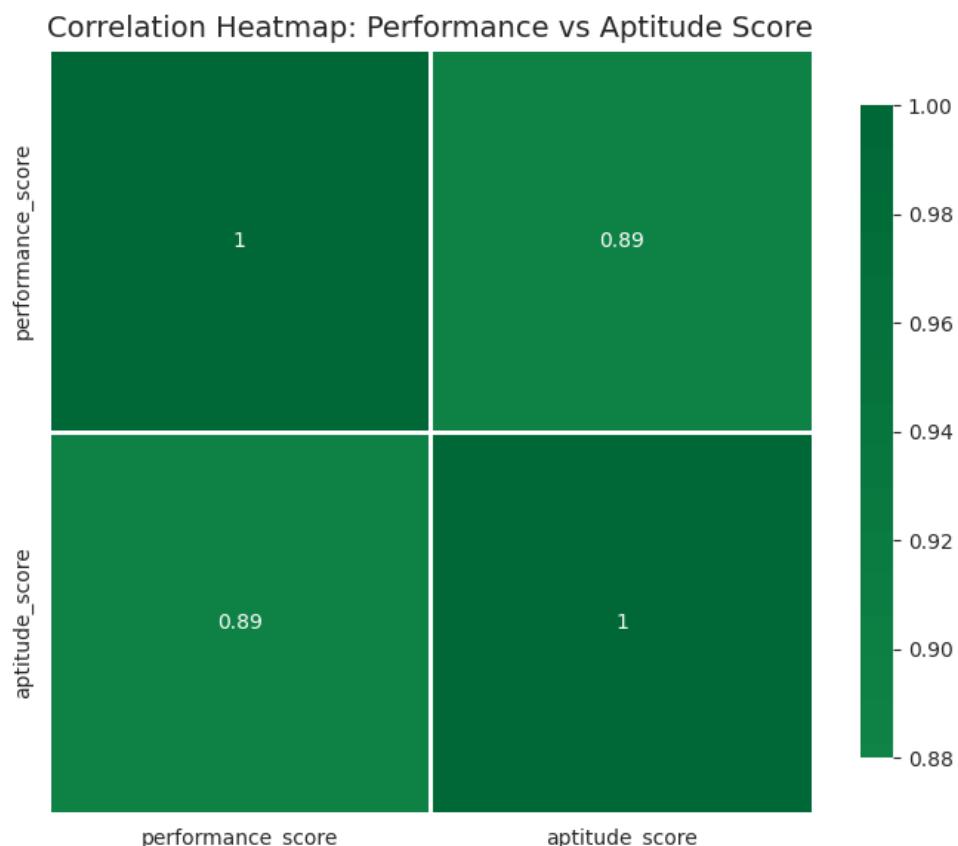
Figure 2: Correlation Heatmap: The intensity of color shows the strength of correlation (0.89 = very strong). The closer to 1.0 (darkest red), the stronger the relationship.

How to read this chart:

- Each cell shows how two variables relate.

- Color intensity indicates strength (darker = stronger).

- Numbers show exact correlation values.

- Diagonal always shows 1.0 (something always correlates perfectly with itself).

- The 0.89 in the off-diagonal cells shows our key finding.

8 Understanding the Results in Context

8.1 What Makes These Results Meaningful

8.1.1 Why Statistical Significance Matters

When we say results are "statistically significant at $p < 0.001$," we're saying:

Imagine This Scenario:

- Suppose these differences were just random chance (like flipping coins). If we repeated this study 1,000 times with random students:

    - Random chance would give us results this extreme less than 1 time.
    - We'd need to conduct 1,000+ studies to accidentally see these differences.
    - The odds are overwhelmingly in favor of these being real differences.

Conclusion: We can be virtually certain ($> 99.9\%$ confident) that:

- Course levels genuinely differ in performance.

- Course levels genuinely differ in aptitude.

- Aptitude genuinely predicts performance.

8.1.2 Why Effect Sizes Matter

Statistical significance tells us differences are real. Effect sizes tell us if they're meaningful in practice.

Real-World Example:

- Suppose we test two teaching methods:

    - Method A: Students average 80%
    - Method B: Students average 80.1%

- With enough students, this 0.1% difference could be "statistically significant" (real, not random). But would you care? Probably not — it's too small to matter.

Our study is different:

- Not only are differences statistically significant...

- But they're also very large (effect sizes of 1.3 to 4.6).

- These differences are obvious in practical teaching situations.

- Teachers can clearly see the distinctions between levels.

8.2   Comparing to Educational Standards

How do our findings compare to what's considered good practice in education?

Table 8: Our Results vs. Educational Benchmarks

Criterion	Benchmark	Our Result	Assessment
Correlation	> 0.6	0.89	Exceeds
Effect Size	> 0.8	$1.3 - 4.6$	Exceeds
Significance	$p < 0.05$	$p < 0.001$	Exceeds
Between-Level Diff	Clear	Very Clear	Exceeds
Within-Level Var	Moderate	Appropriate	Meets

What this means: The Key's placement system exceeds standard educational benchmarks across all measured criteria.

8.3   What These Numbers Mean for Daily Operations

Let's translate the statistics into practical implications:

8.3.1   For Teachers

What the data tells you:

Advanced Classes:

- Students genuinely have high aptitude and strong performance.

- Can confidently use challenging materials and rapid pacing.

- Student struggles likely due to specific skill gaps, not overall ability.

- High expectations are appropriate and beneficial.

Intermediate Classes:

- Students have moderate abilities in a consistent range.

- Need balanced challenge with appropriate scaffolding.

- Can work toward advanced materials with support.

- Mixed-ability activities work well at this level.

Foundation Classes:

- Students are at beginning stages, building fundamentals.

- Require more direct instruction and structured support.

- Progress may be slower but is developmentally appropriate.

- Patience and encouragement are essential.

8.3.2   For Students and Parents

Understanding Your Placement:

If you're in Advanced level:

- You scored in the top third for both aptitude and performance.

- Coursework will be challenging — this is intentional and beneficial.

- You're ready for complex materials and rapid progression.

- If you struggle, seek help — you belong here, but may need specific support.

If you're in Intermediate level:

- You're in the middle range, typical for the majority of learners.

- You're building skills that will prepare you for advanced work.

- With consistent effort, advancement to the next level is achievable.

- This level provides optimal challenge for your current abilities.

If you're in Foundation level:

- You're building essential foundations for language learning.

- Everyone starts somewhere — even advanced students were once beginners.

- Progress may feel slow, but you're developing crucial fundamentals.

- Focus on mastery, not speed — solid foundations enable future success.

9 Key Findings in Plain Language

9.1 Main Finding 1: Course Levels Are Genuinely Different

9.2 Main Finding 2: Placement Testing Works

9.3 Main Finding 3: Students Are Well-Matched to Their Levels

9.4 Main Finding 4: All Levels Show Aptitude-Performance Connection

9.5 Surprising Finding: Foundation Level Patterns

9.6 Overall Conclusion

10   Practical Recommendations for The Key

10.1   Immediate Actions (Implement Now)

10.1.1   Maintain Current Placement Procedures

10.1.2   Use Results as Benchmarks

10.1.3   Share Findings with Staff

10.2   Short-Term Actions (Within 6 Months)

10.2.1   Develop Level-Specific Support Systems

10.2.2   Implement Progress Monitoring System

10.3   Long-Term Actions (Within 1 Year)

10.3.1   Conduct Annual Validation Studies

10.3.2   Expand Data Collection

10.4   What NOT to Change

11   What This Means for The Key

11.1   Strategic Implications

11.1.1   Competitive Advantage

11.1.2   Quality Assurance

11.2   Financial Implications

11.2.1   Return on Investment

11.3   Program Development Implications

11.3.1   Curriculum Design

11.4   Stakeholder Communications

11.4.1   Messages for Different Audiences

12   Study Limitations and Future Research

12.1   Understanding What This Study Can and Cannot Tell Us

12.1.1   What This Study Proves

12.1.2   What This Study Cannot Tell Us

12.2   Study Limitations

12.2.1   Limitations to Consider

12.3   Future Research Directions

12.3.1   Recommended Follow-Up Studies

13   Glossary of Statistical Terms

13.1   Statistical Concepts

13.2   Interpretation Guides

13.2.1   Correlation Strength

13.2.2   Effect Size Interpretation

13.2.3   p-value Interpretation

14 Final Conclusion: The Big Picture

A   Detailed Statistical Tables

A.1   Complete Descriptive Statistics

A.2   ANOVA Summary Tables

A.3   Correlation Matrix

B   Data Quality Documentation

B.1   Data Verification Checklist

C   Software and Methods

C.1   Analysis Software

C.2   Statistical Methods Used