

PATHWAYS TO PROFICIENCY: AN EXPLORATION OF HOW COGNITIVE AND NON-COGNITIVE VARIABLES CONTRIBUTE TO THE DEVELOPMENT OF ACHIEVEMENT AND PROFICIENCY OUTCOMES

Megan C. Masters, PhD
University of Maryland College Park
Romance Languages LEARN Workshop
Loyola Graduate Center, Columbia, MD
March 26, 2019
8:40 a.m.

Language Program Reviews

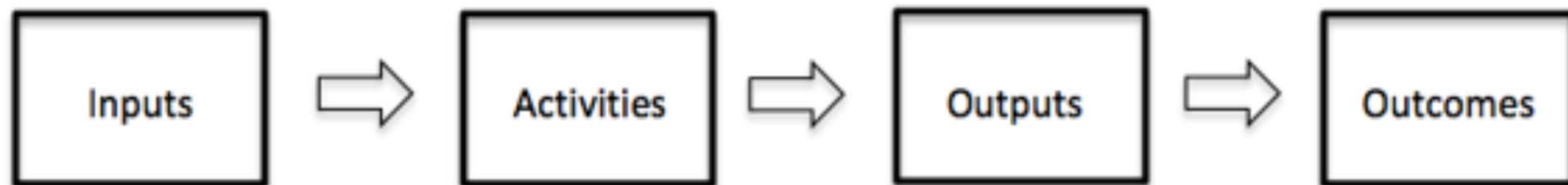
- Increased evaluation, assessment, and accreditation demands across all levels of higher education, particularly during times of fiscal austerity
- Considerable variation in strategies and methods employed when documenting program processes and outcomes (Ross, 2003; Watanabe, 2009)
 - Early L2 program evaluation research focused on the summative evaluation process
 - Main purpose was to determine the superiority of a given pedagogical approach (Genesee, 1985; Keating, 1963; Schreier & Wertheimer, 1964; Smith, 1970, Swain & Lapkin, 1982)
 - Later research called for including more diverse types of program evidence, such as interviews, surveys, and questionnaires (Alderson, 1992; Beretta, 1986, 1992, a,b; Frechtling, 2007; Kiley & Rea-Dickens, 2005; Lynch, 1996; Norris 2006, 2009; Rea-Dickens & Germaine, 1998) (for a more detailed discussion see also Watanabe, 1999; Ross, 2003).
 - Most recent research have applied the use of data mining strategies as an alternative to quasi-experimental research designs (Bloomfield et al., 2012; 2013; Mackey, 2014; Ross, 2011; Wagner, 2014).
- Goal of most L2 program reviews is to systematically document how well a given intervention, product, or system is working (Brousseau & Champagne, 2011; Frechtling, 2007)
- Absence of literature detailing best practices for using empirical models to document L2 learner progress

Program Theory

- Program Theory is “the set of assumptions about the manner in which the program relates to the social benefits it is expected to produce and the *strategy* and tactics the program has adopted to achieve its goals and objectives” (Callow-Heusser et al., 2005, p.38)
- Defines, a-priori, each of the causal links expected to occur between project start and goal attainment
- Allows the process of program review to play an integral role in the development of a program theory

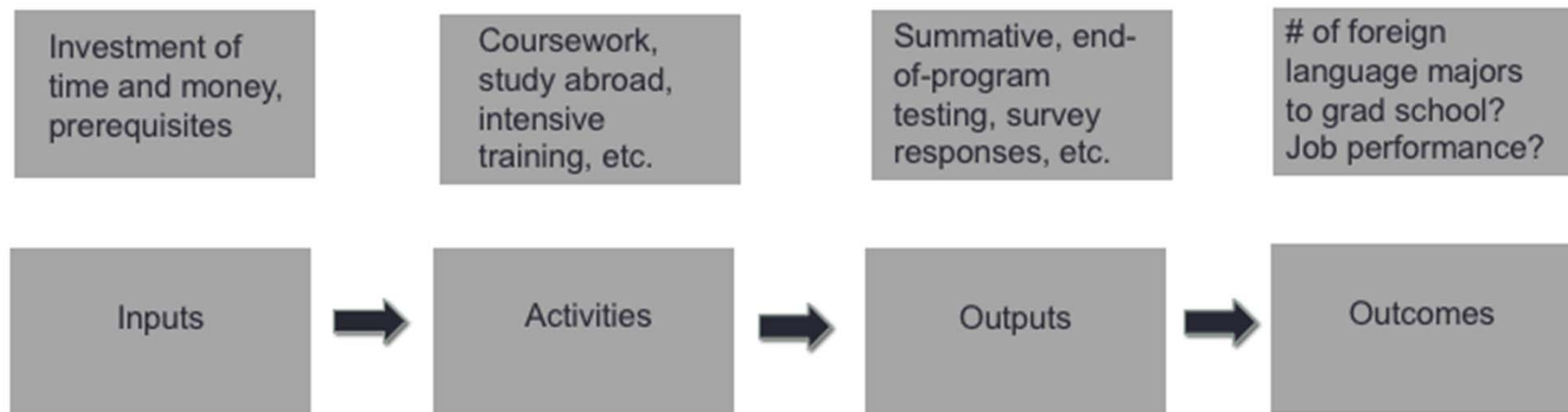
Logic Modeling

- A method of making explicit the assumptions associated with each stage of a given program
- The purpose of a logic model is to make clear, via visual representation, the underlying rational or “logic” of a program as well as the expected causal relationships between each phase
- “Knowing what causal factors are being targeted first is essential to assessing whether an activity is appropriately targeted, identifying appropriate indicators of change, and writing sound objectives” (Renger and Titcomb, 2002, p. 494)
- When well defined can be used as a substitute for classical experimental study using random assignment (Weiss, 1997, 2000)



Logic Modeling and L2 Instructional Programs

- Large-scale L2 instructional programs are typically comprised of both localized indicators of student achievement (e.g., GPA scores) and standardized summative assessments (e.g., proficiency tests)




Logic Modeling and L2 Instructional Programs

- Defense Language Institute Foreign Language Center (DLIFLC)
- For enlisted personnel to attend, must meet a minimum performance threshold on sub-sections of the Armed Services Vocational Battery (ASVAB)
- Those who meet minimum ASVAB requirements qualify to take the Defense Language Aptitude Battery (DLAB)
- Prior to language instruction, students take an Introduction to Language Studies (ILS) course
 - Five-day course designed to prepare and motivate students
- As part of this course, students take a Language Preference Self-Assessment
 - I am here to study this language, which is...
 - Not my choice. I would prefer to do something else rather than study a foreign language.
 - Not my choice. I am not motivated to study the assigned language.
 - Not my choice, but I am still motivated to study the assigned language.
 - Based on my second or third choice.
 - Based on my first choice (Lett, n.d.)

Logic Modeling and L2 Instructional Programs

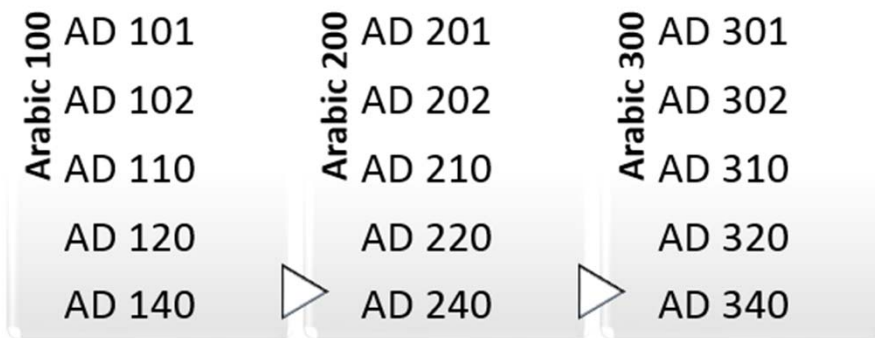
- USG has developed various LDC frameworks that are characterized as:
 - Comprehensive
 - Efficient
 - High-Stakes
- Approaches to grouping languages within a hierarchy have been informed by early SLA research as well as by observations made by expert L2 instructors or training specialists
- Upon admission, students assigned to one of over 20 languages within the LDC framework
 - Courses last between 26 and 64 weeks
 - Training length dependent upon the category to which a language is assigned (Cat I-IV)
 - Five days per week
 - Seven hours per day
 - Two-three hours of homework per night



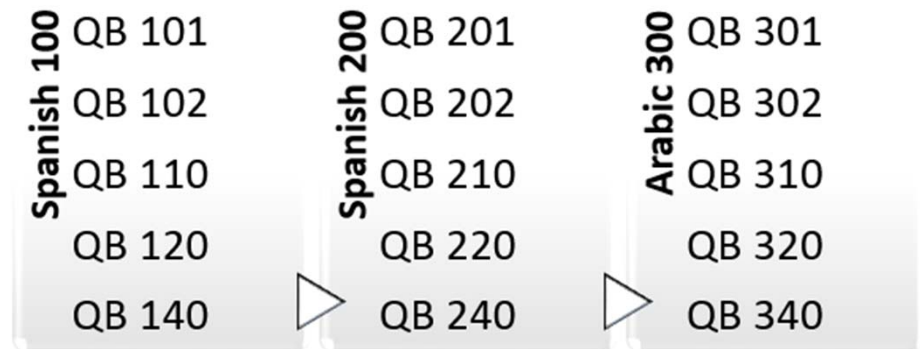
	Weeks Instruction	Languages
Category IV	64	MSA, Egyptian Arabic, Iraqi Arabic, Levantine Arabic, Sudanese Arabic, Mandarin Chinese, Japanese, Korean, Pashto
Category III	48	Hebrew, Hindi, Persian Farsi, Russian, Serbian/Croatian, Tagalog, Turkish, Urdu
Category II	35	German, Indonesian
Category I	26	French, Portuguese, Spanish

Learner Progression Through DLIFLC Coursework

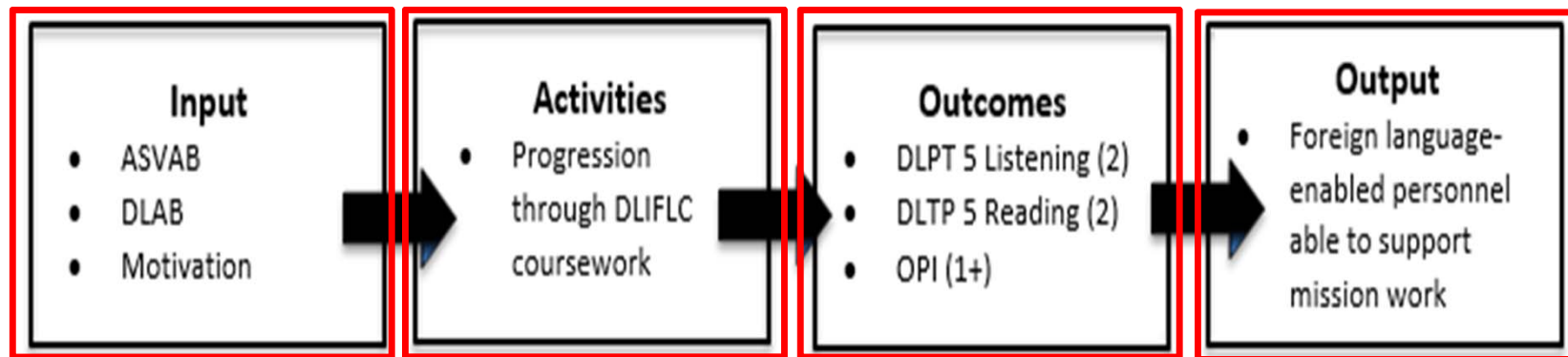
Arabic Coursework



Spanish Coursework



Logic Modeling and L2 Instructional Programs



The purpose of a logic model is to make clear, via visual representation, the underlying rationale, or “logic” of program, as well as the expected causal relationships between each phase

Study Motivation and Research Questions

- Previous studies have focused on the role of cognitive and non-cognitive factors in predicting ultimate program outcomes and/or growth, maintenance, or loss in FL proficiency test scores over one's career (Boomfield et al., 2012; 2013, Lett & O'Mara, 1990; Mackey, 2015; Ross, 2011; Wagner, 2014)
- Few research studies have examined how cognitive and non-cognitive variables predict both course achievement and end-of program outcomes
- Despite widespread adoption of LDC framework, few empirical validation studies
- Three main research questions:
 - RQ1: How do differences in general aptitude, language-specific aptitude, and motivation predict learners' success as they progress through coursework?
 - RQ2: How do differences in course achievement predict learners' end-of-program proficiency test scores?
 - RQ3: How do differences in general aptitude, language-specific aptitude, predict learners' end-of-program proficiency test scores?
 - RQ4: Are differing paths observed for learners studying languages within different difficulty categories? (e.g., a Category I language vs. a Category IV language)?

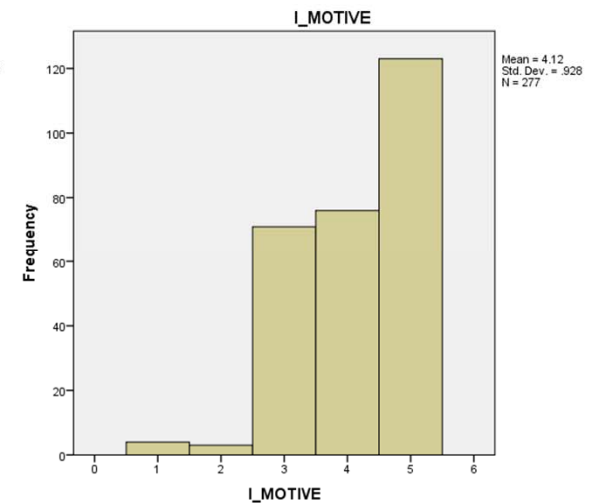
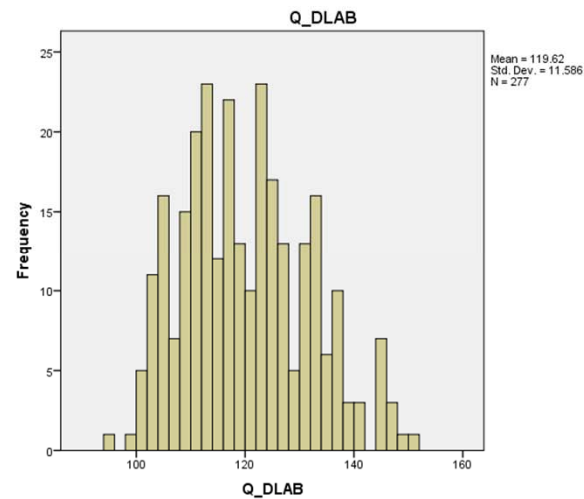
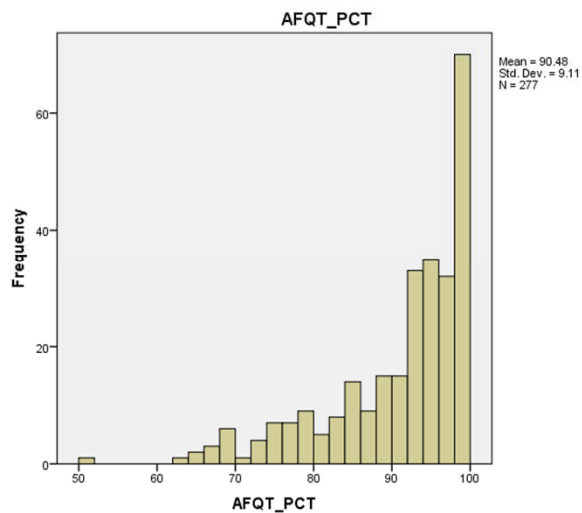
Dataset

- Extant data from previously conducted research at the University of Maryland Center for Advanced Study of Language
- Data from four different systematically maintained databases were restructured, merged, and coded
- Learners with incomplete longitudinal records and students who had been recycled or re-languaged were not included in the analysis

Descriptive Statistics (Arabic)

- Existing learner data from March 22, 2001-June 16, 2011

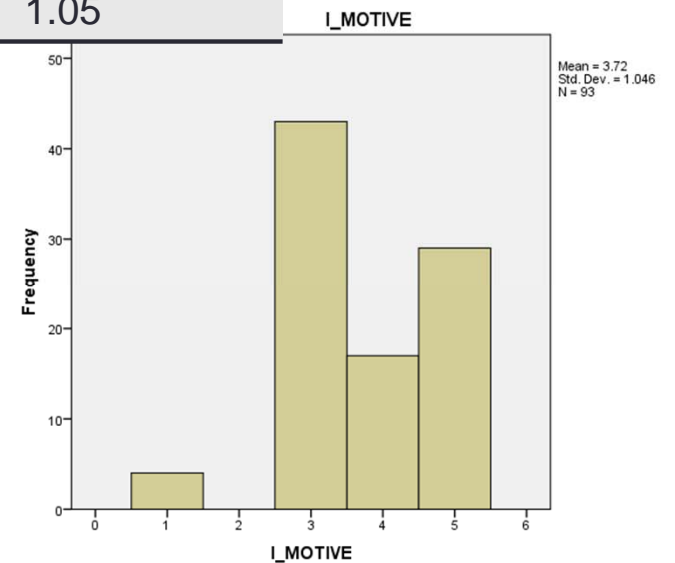
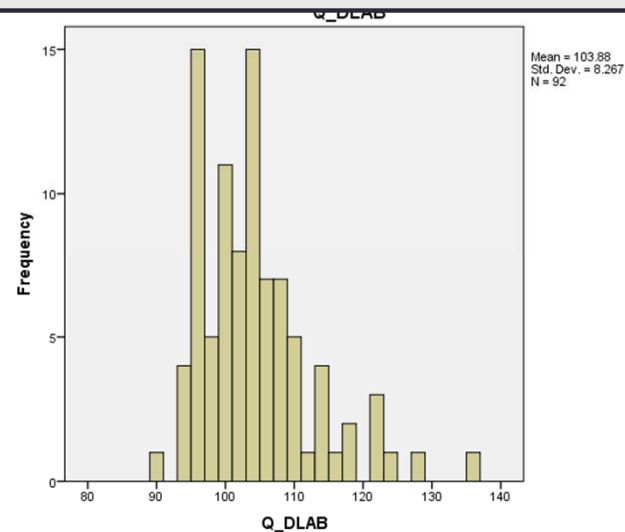
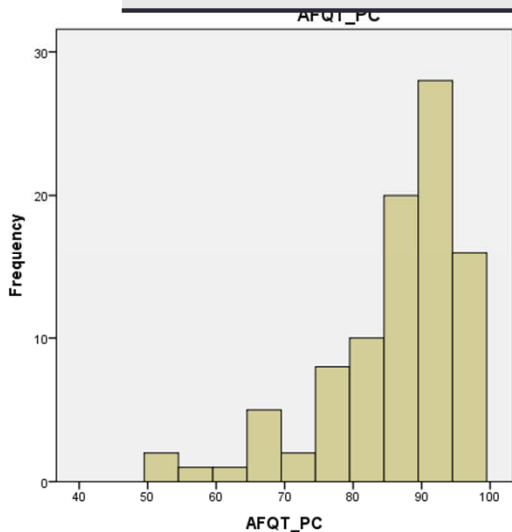
	N	Minimum	Maximum	Mean	Std. Deviation
AFQT_PCT	277	51	99	90.48	9.11
Q_DLAB	277	95	150	119.62	11.59
Motivation	277	1	5	4.12	0.928



Descriptive Statistics (Spanish)

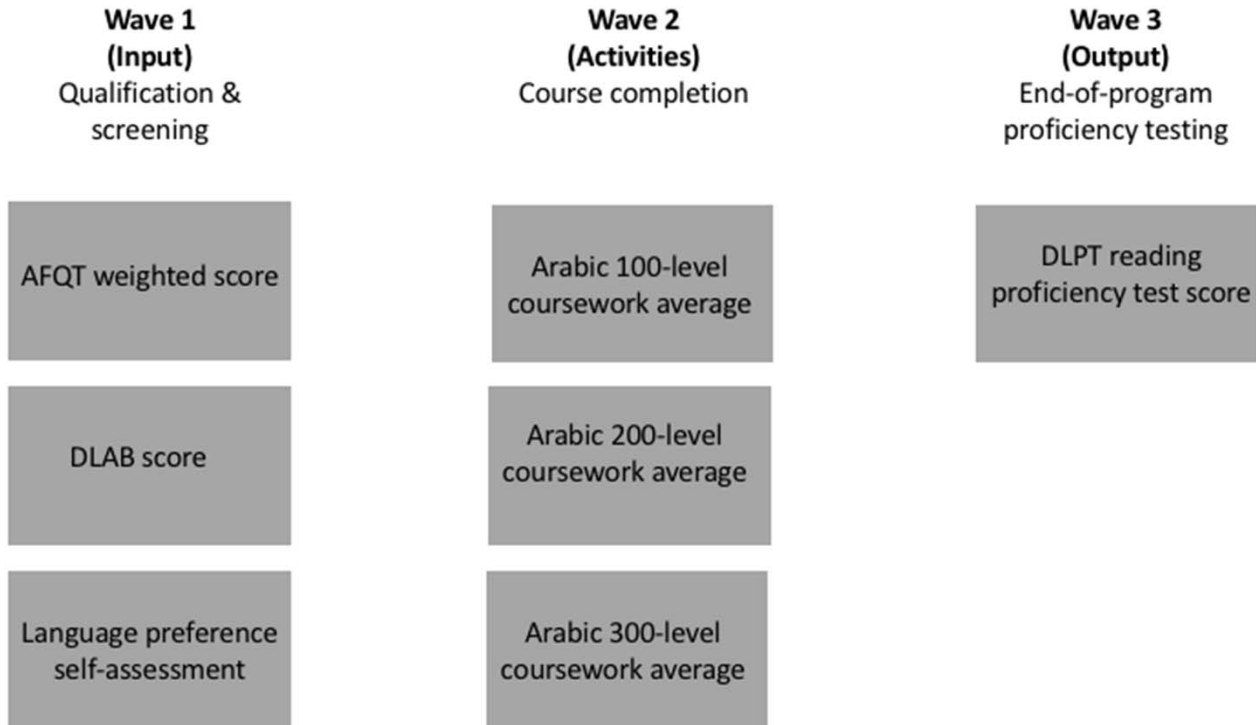
- Existing learner data from February 12, 2009 –March 04, 2010

	N	Minimum	Maximum	Mean	Std. Deviation
AFQT_PC T	93	52	99	86.30	10.22
Q_DLAB	93	90	135	103.88	8.27
Motivation	93	1	5	3.72	1.05

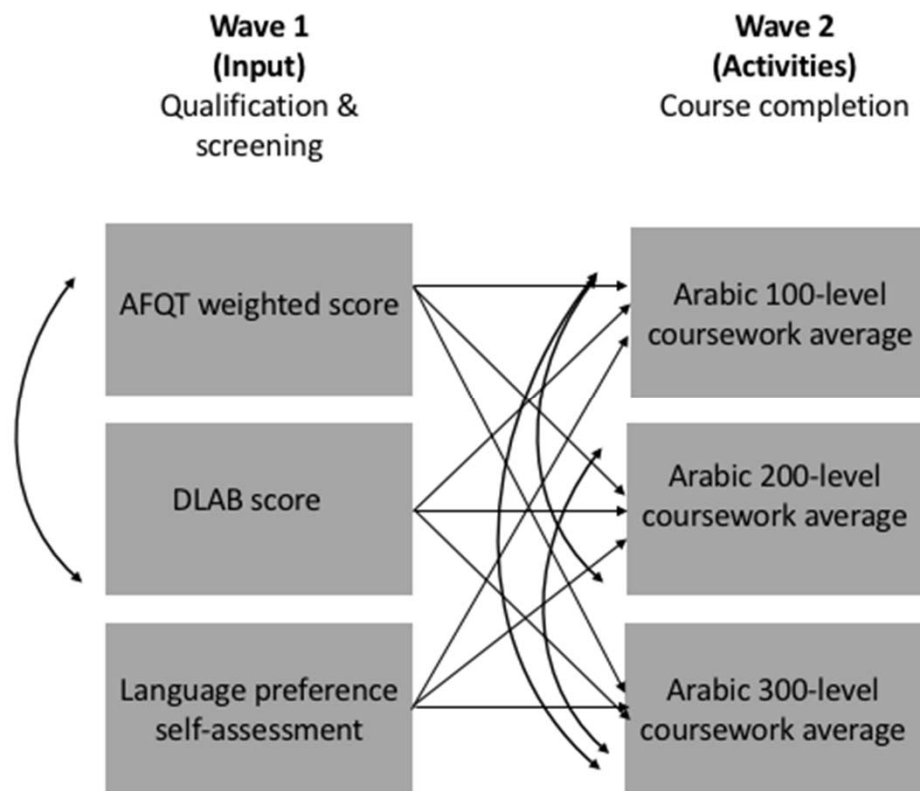


Research Design: Panel Study

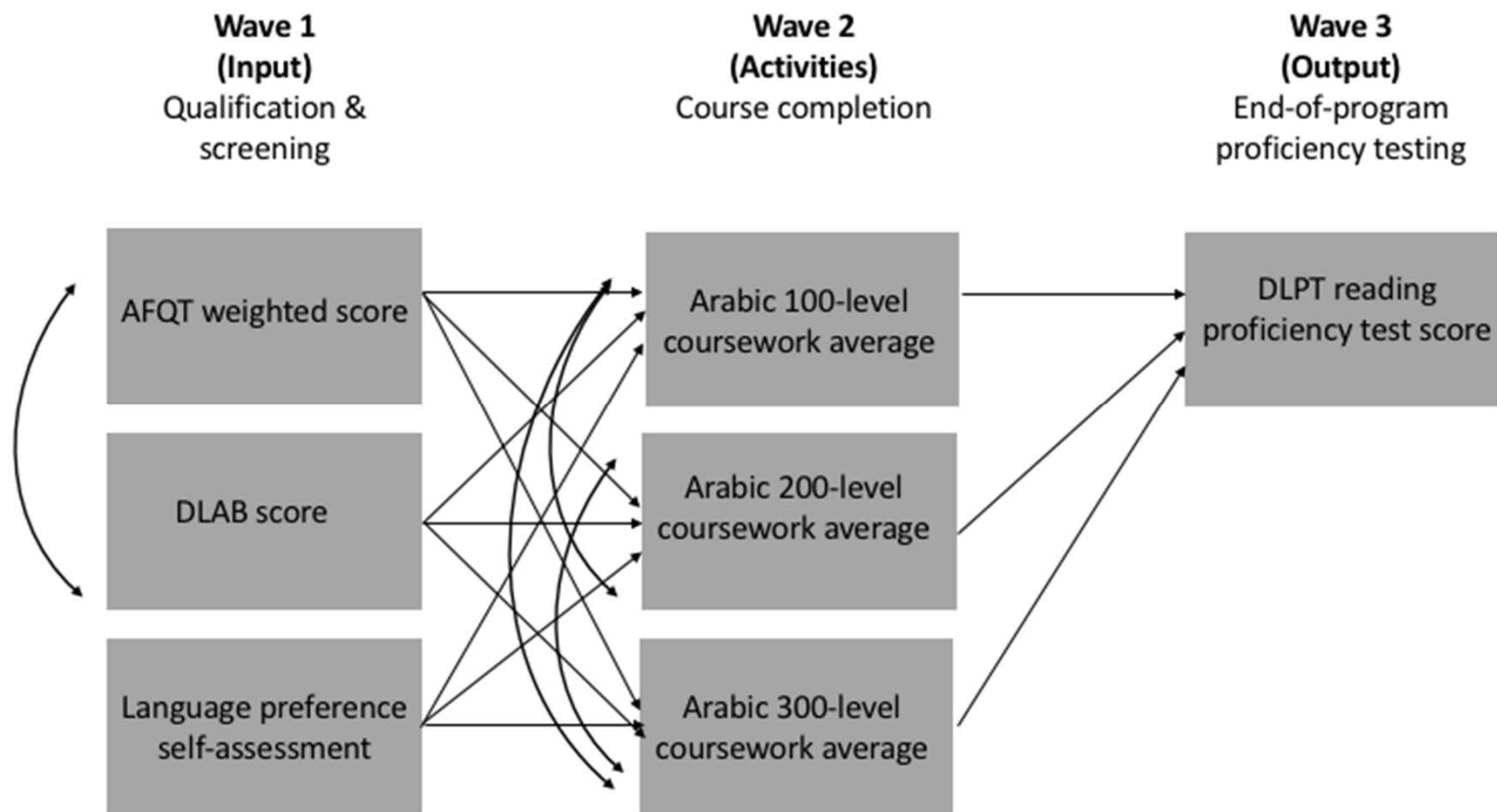
Intensive foreign language instruction takes place over **64 weeks**.



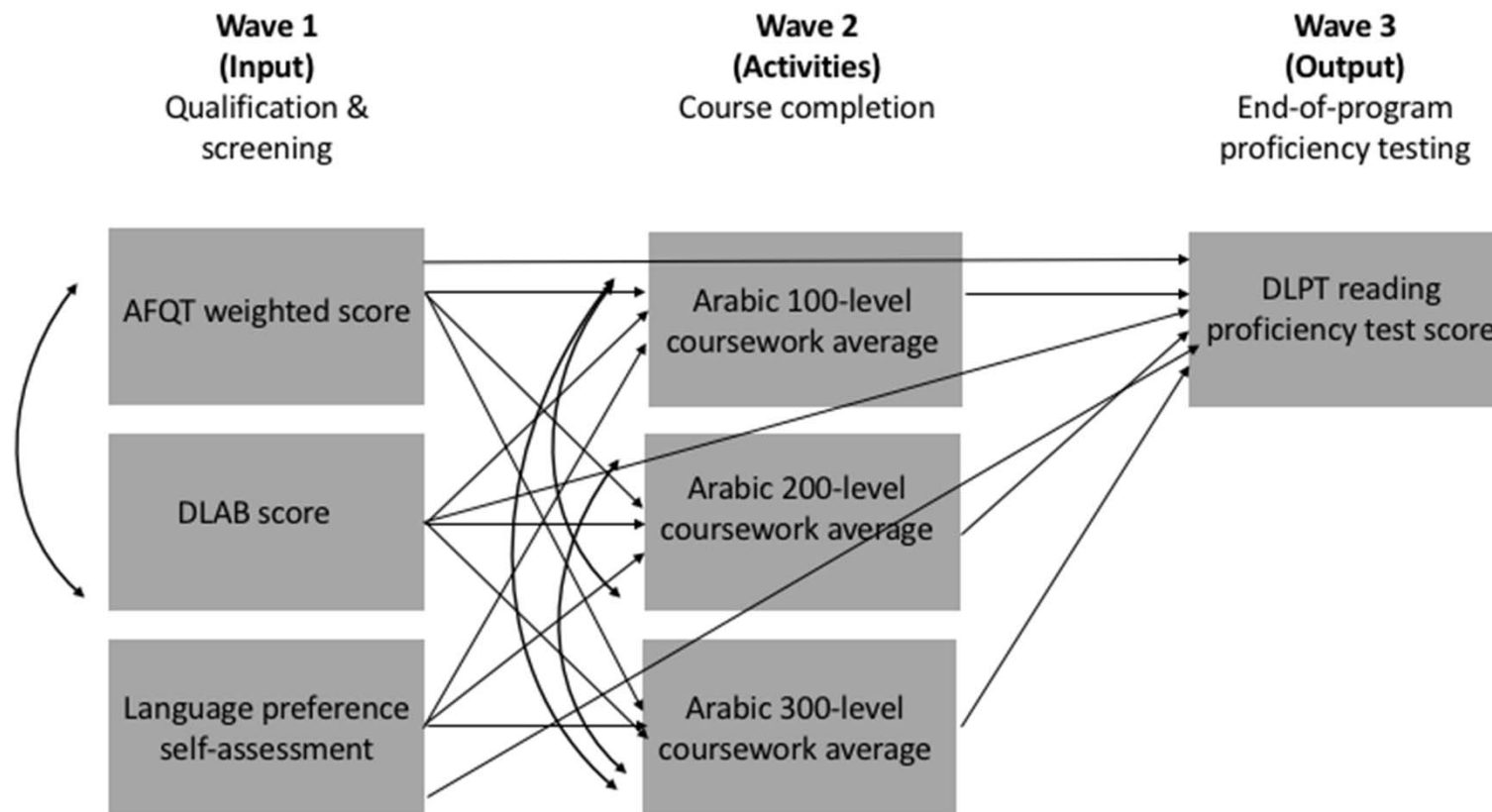
Predicted Results (Wave 1 to Wave 2)



Predicted Results (Wave 2 to Wave 3)



Predicted Results (Wave 1 to Wave 3)



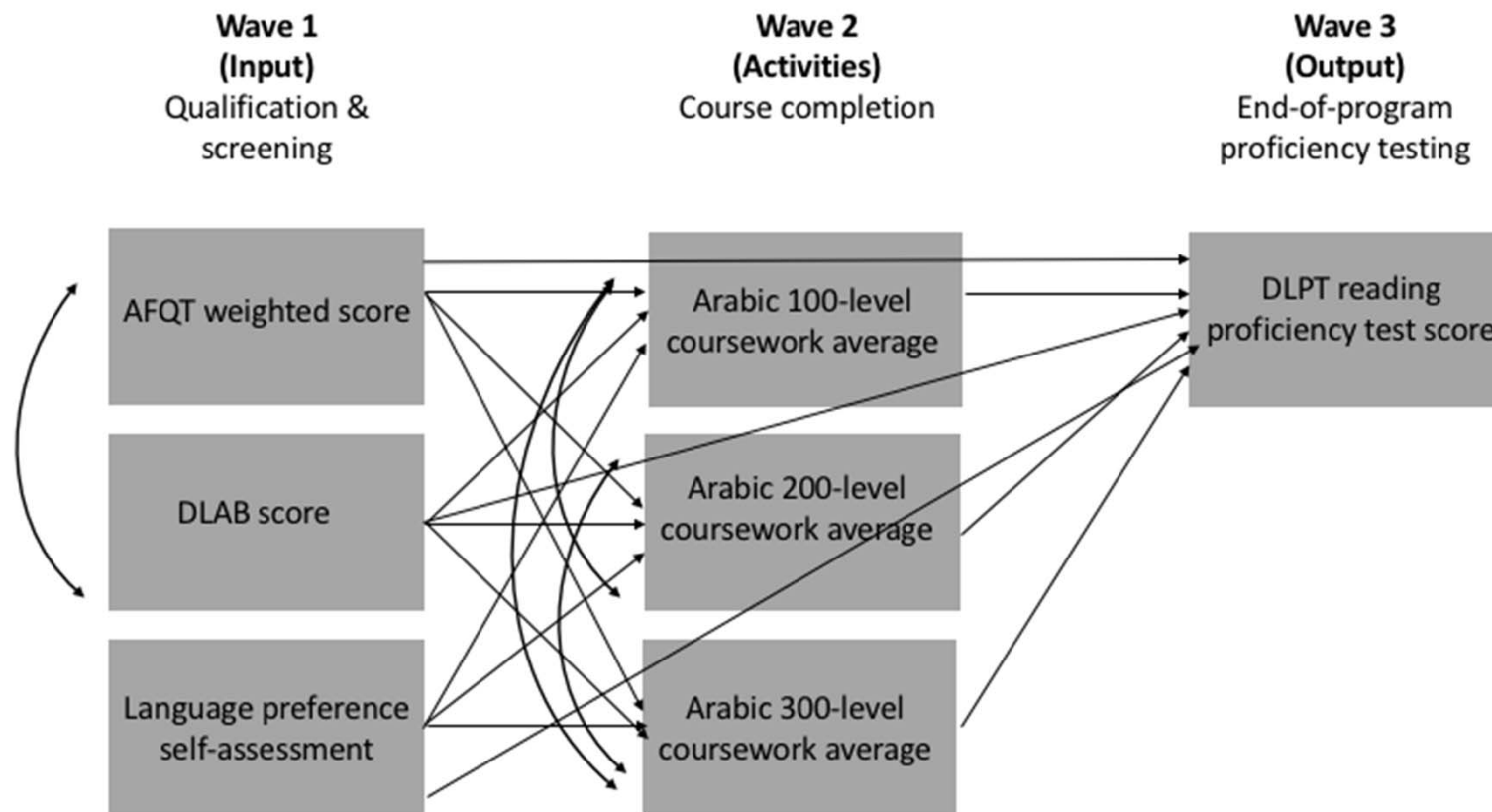
Method of Analysis: Path Analysis

- ❖ Allows for the simultaneous estimation of both indirect and direct causal influences between upstream and downstream variables
- ❖ Purpose is to interpret a-priori specified, theory defined, hypothesized causal relationships specified within a path diagram (Byrne, 2001 ;Hancock and Mueller, 2011)
 - ❖ Output from the analysis is analyzed in terms of model fit indices
 - ❖ Chi Square Statistics (p value > 0.05)
 - ❖ Confirmatory Fit Indices (CFI > 0.90)
 - ❖ Root Mean Square Error of Approximation (< 0.06)
 - ❖ Model is compared to both fully saturated and independent models
- Path diagrams are graphical representations of the structural equations underlying each hypothesized relationship between measured variables
- Each endogenous variable expressed as a function of all elements having a direct structural effect or covariation with it
- Final model contains only the significant pathways

Method of Analysis: Recursive Path Analysis

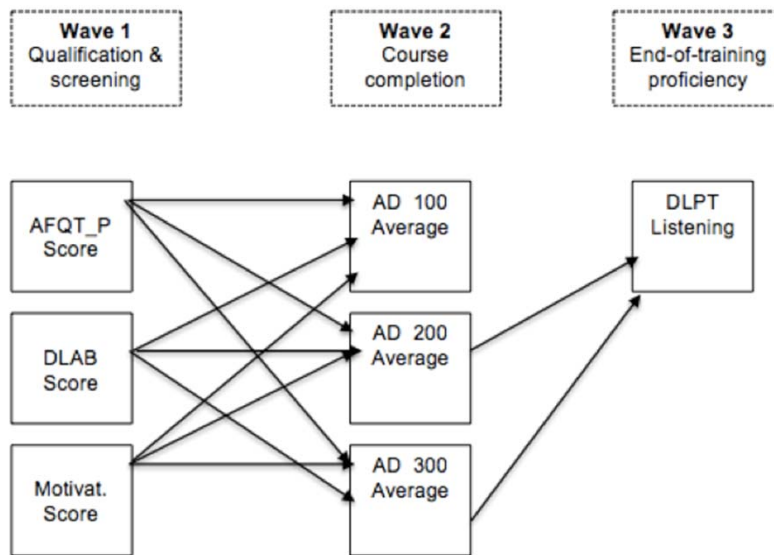
Variable	Structural Equation
AD 100	= constant + $\beta(\text{AFQT})$ + $\beta(\text{DLAB})$ + $\beta(\text{AFQT}*\text{DLAB})$ + $\beta(\text{Motivation})$ + error
AD 200	= constant + $\beta(\text{AFQT})$ + $\beta(\text{DLAB})$ + $\beta(\text{AFQT}*\text{DLAB})$ + $\beta(\text{Motivation})$ + error
AD 300	= constant + $\beta(\text{AFQT})$ + $\beta(\text{DLAB})$ + $\beta(\text{AFQT}*\text{DLAB})$ + $\beta(\text{Motivation})$ + error
DLPT L	= constant + $\beta(\text{AFQT})$ + $\beta(\text{DLAB})$ + $\beta(\text{AFQT}*\text{DLAB})$ + $\beta(\text{Motivation})$ + $\beta(\text{AD100})$ + $\beta(\text{AD200})$ + $\beta(\text{AD300})$ + $\beta(\text{AD100}*\text{AD200})$ + $\beta(\text{AD100}*\text{AD300})$ + $\beta(\text{AD200}*\text{AD300})$ + error
DLPT R	= constant + $\beta(\text{AFQT})$ + $\beta(\text{DLAB})$ + $\beta(\text{AFQT}*\text{DLAB})$ + $\beta(\text{Motivation})$ + $\beta(\text{AD100})$ + $\beta(\text{AD200})$ + $\beta(\text{AD300})$ + $\beta(\text{AD100}*\text{AD200})$ + $\beta(\text{AD100}*\text{AD300})$ + $\beta(\text{AD200}*\text{AD300})$ + error
DLPT S	= constant + $\beta(\text{AFQT})$ + $\beta(\text{DLAB})$ + $\beta(\text{AFQT}*\text{DLAB})$ + $\beta(\text{Motivation})$ + $\beta(\text{AD100})$ + $\beta(\text{AD200})$ + $\beta(\text{AD300})$ + $\beta(\text{AD100}*\text{AD200})$ + $\beta(\text{AD100}*\text{AD300})$ + $\beta(\text{AD200}*\text{AD300})$ + error

Results: Do the data fit the model-implied theory?

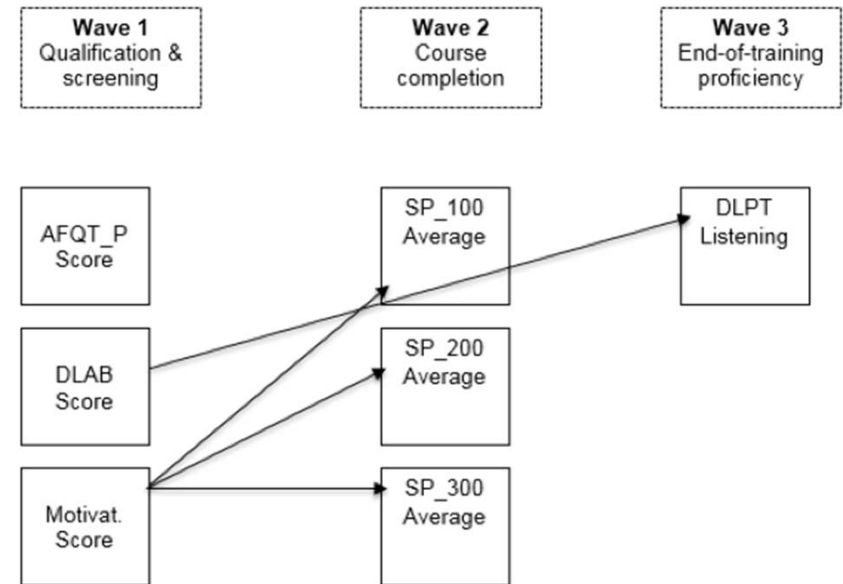


Results: Listening

Arabic

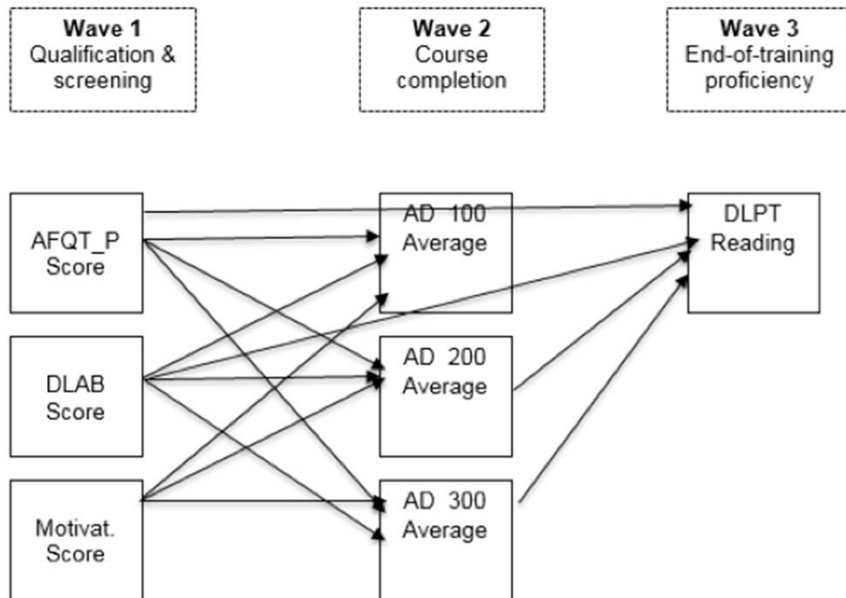


Spanish

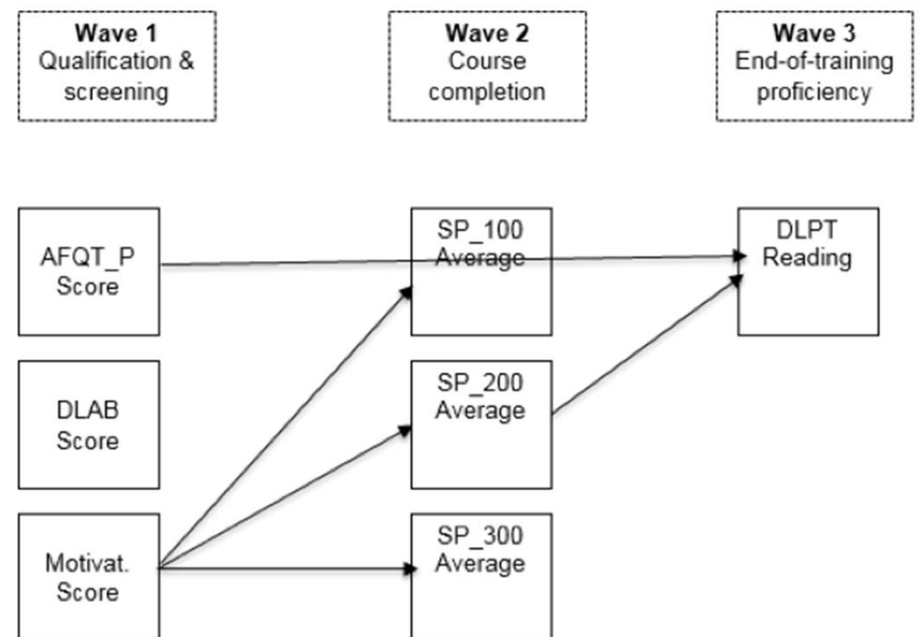


Results: Reading

Arabic

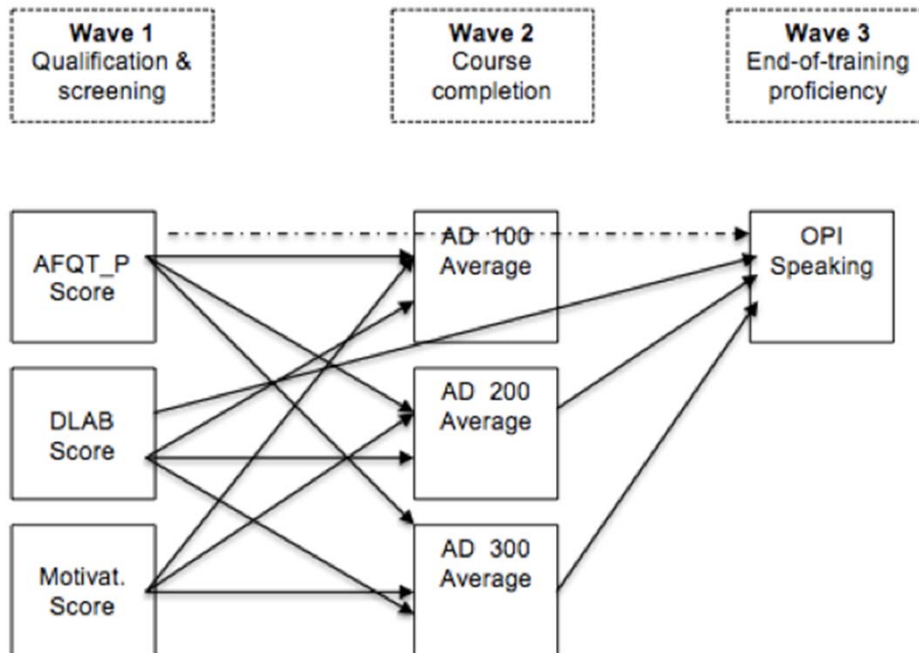


Spanish

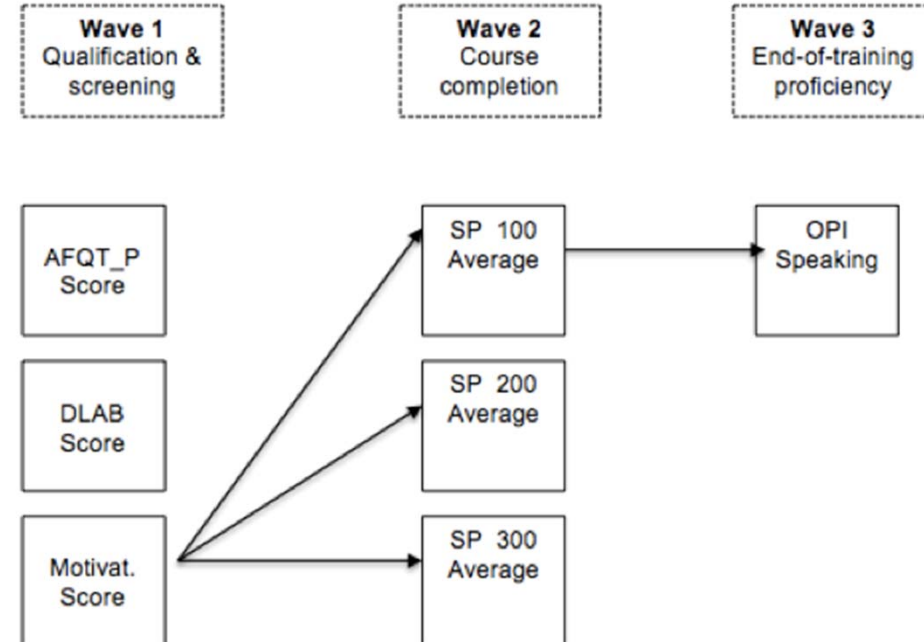


Results: Speaking

Arabic



Spanish



Discussion

- **RQ1: How do differences in general aptitude, language-specific aptitude, and motivation predict learners' success as they progress through coursework?**

Arabic

- Wave 1→Wave 2 (All Skills)
 - All cognitive and non-cognitive variables predict learners' success throughout 100-, 200-, and 300-level coursework
 - DLAB strongest predictor of 100-level coursework
 - Motivation strongest predictor of 200- and 300-level coursework
 - The influence of AFQT and DLAB scores strongest for 100-level coursework and wanes as students progress through 300-level coursework
 - Motivation variable exhibits the opposite pattern, showing weakest influence for 100-level coursework and strongest influence for 300-level coursework

Spanish

- Wave 1→Wave 2 (All Skills)
 - Only the motivation variable predicted 100-, 200-, and 300-level average course outcomes
 - Motivation variable exhibited a V-shaped influence as students progressed through coursework, but was strongest in predicting 300-level average Spanish course outcomes

Discussion

- **RQ2: How do differences in general aptitude, language-specific aptitude and motivation predict learners' end-of-program proficiency test scores?**

Arabic

- Listening
 - Wave 1: None
 - Wave 2: AVG_200 (0.368) and AVG_300 (0.319)
- Reading
 - Wave 1: AFQT (0.143) and DLAB scores (0.117)
 - Wave 2: AVG_200 (0.280) and AVG_300 (0.272)
- Speaking
 - Wave 1: **AFQT (-0.110)** and DLAB scores (0.148)
 - Wave 2: AVG_200 (0.316) and AVG_300 (0.147)

Spanish

- Listening
 - Wave 1: None
 - Wave 2: None
- Reading
 - Wave 1: AFQT (0.143) and DLAB scores (0.117)
 - Wave 2: AVG_200 (0.489)
- Speaking
 - Wave 1: **AFQT (-0.110)** and DLAB scores (0.148)
 - Wave 2: AVG_100 (0.431)

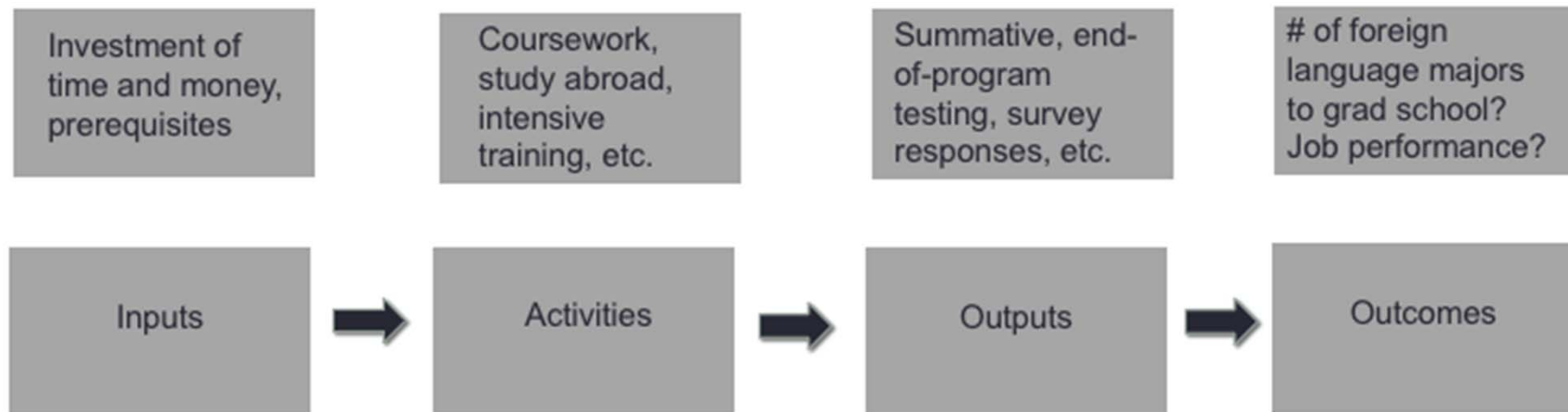
Discussion

- RQ3: Are differing paths observed for learners studying languages within different difficulty categories?

Language	Skill	Sig. Paths from Wave 1 to 2 (RQ1)	Sig. paths from Wave 2 to 3 (RQ3)	Sig. paths from Wave 1 to 3
Arabic	Listening	AFQT, DLAB, Motivation	AD_200, AD_300	None
	Reading	AFQT, DLAB, Motivation	AD_200, AD_300	AFQT, DLAB
	Speaking	AFQT, DLAB, Motivation	AD_200, AD_300	-AFQT, DLAB
Spanish	Listening	Motivation	None	DLAB
	Reading	Motivation	SP_200	AFQT
	Speaking	Motivation	SP_100	None

Implications: Logic Model

- Logic models useful for defining the main components of each phase of a given program
- Lend themselves well to the use path analysis
 - Allows for predicted steps between program input, activities, outcomes, and output to be modeled empirically

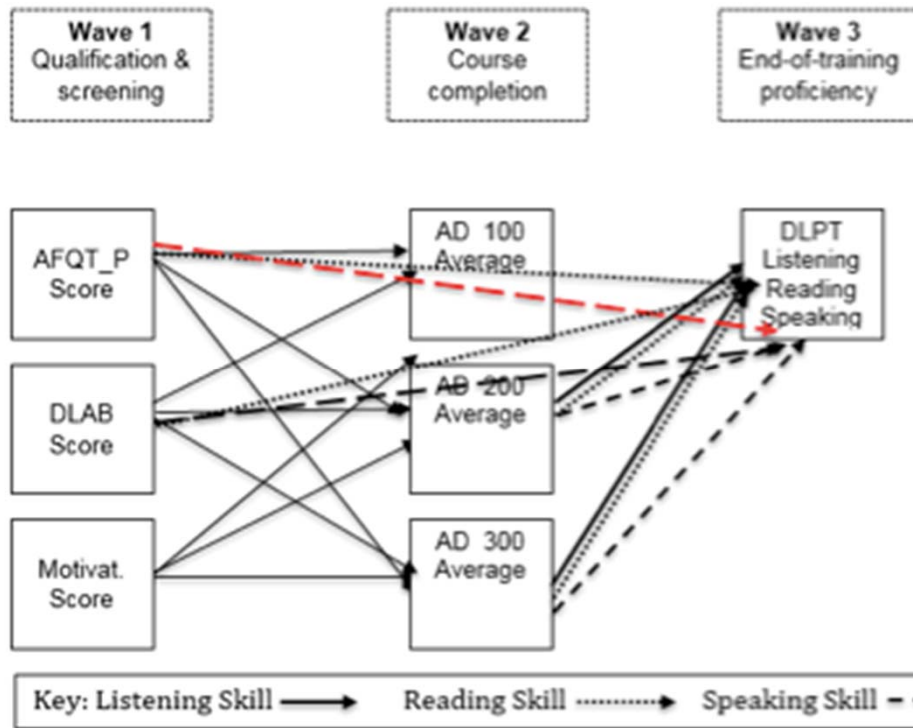


Implications

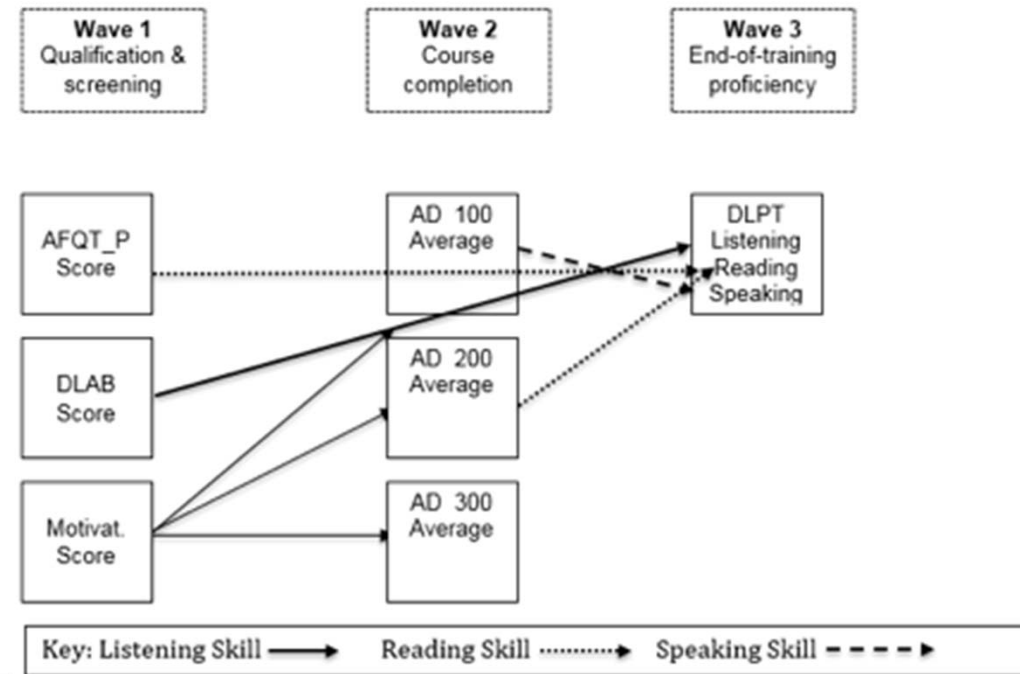
- Cognitive variables found to play a key role in predicting coursework success for a Cat IV language, but not a Cat I language
- For both Cat I and Cat IV languages, motivation found to play a significant role in predicting coursework success, but not in predicting ultimate outcomes
- Motivation plays a significant role in predicting coursework success, particularly as learners progress through coursework
- Variability in the predictive influence of course achievement-related variables on DLPT outcomes
- Lack of homogeneity in the significant paths found across skills within the same language and between languages

Implications

Arabic Path Model



Spanish Path Model



Limitations

- **Restricted Sample:** All learners have already been twice selected from their ASVAB and DLAB scores, which will impact the predictive power of these variables (Lett & O'Mara, 1990)
 - Scores for general ability and aptitude not included in current sample
 - Negatively skewed AFQT, DLAB, Motivation variables (Arabic)
- **Contextual Differences:** Despite the general homogeneity of the language learning context at DLIFLC, it is unlikely that the programs will be pedagogically equivalent within or across Arabic and Spanish courses (Lett & O'Mara, 1990)
- **Test Version Differences:** Arabic dataset contains DLPT IV to DLPT 5 test version differences for the reading and listening outcome measures
- **Analyses not separated by skill:** Need to examine curricular differences
- **Existing Data:** Analyses constrained to data available within existing datasets; also assumes accuracy of data entry
- **Complete Cases:** Analyses constrained to data from only those learners who completed entire program of study. To what extent are the results influenced by potential hidden effects of non-random attrition across learners?

Conclusions

- Well-designed logic modeling can provide feedback to both teachers and students on the learning process itself as well as to educational administrators concerning general program coherence
- Results of path analytic framework provide initial empirical validation to the LDC system
 - Role of cognitive variables varied by language category and by skill
 - Learners' motivation plays an increasingly important role in academic achievement
 - Significant heterogeneity in the development of proficiency between the Arabic and Spanish languages
- Future research should explore variation for languages grouped within the same difficulty category
 - Important implications of this research since the difficulty categorization of a given language impacts total instructional time
 - Career-long foreign language skill maintenance would also be important to examine as well

Thank you!

- Mentors and Advisors
 - Steven J. Ross, PhD, Advisor
 - Catherine J. Doughty, PhD
 - Jared Linck, PhD
- Colleagues:
 - Amber Bloomfield, PhD
 - Martyn Clark, PhD
 - Alia Lancaster, PhD
- My family

Questions?