

Mosher, A., Slagter, P. J., & Surface, E. A. (2010, November). *CEFR and ACTFL guidelines: Correlating the rubrics and descriptors*. Symposium presented at the ACTFL 2010 Annual Convention and World Languages Expo, Boston, MA.

# **CEFR and ACTFL Guidelines: Correlating the Rubrics and Descriptors**

Art Mosher  
University of Dayton

Peter J. Slagter  
University College Utrecht

Eric A. Surface  
SWA Consulting Inc.



**NOVEMBER 2010**

APPROVED FOR PUBLIC RELEASE;  
DISTRIBUTION UNLIMITED

## **Copyright Notice**

This document and its content is copyright ©1997-2010 of SWA Consulting Inc. All rights reserved.

Any redistribution or reproduction of part, or the entire document in any form is prohibited except for: (1) you may print or download to a local hard disk extracts for your personal and non-commercial use only, and (2) you may copy the content to individual third parties for their personal use, but only if you acknowledge the website and author(s) as the source of the material. You may not, except with our express written permission, distribute or commercially exploit the content, nor may you transmit it or store it on any other website or other form of electronic retrieval system.



# CEFR and ACTFL Guidelines: Correlating the Rubrics and Descriptors

November 19, 2010

ACTFL Annual Convention 2010

Dr. Eric A. Surface

Dr. Peter Jan Slagter

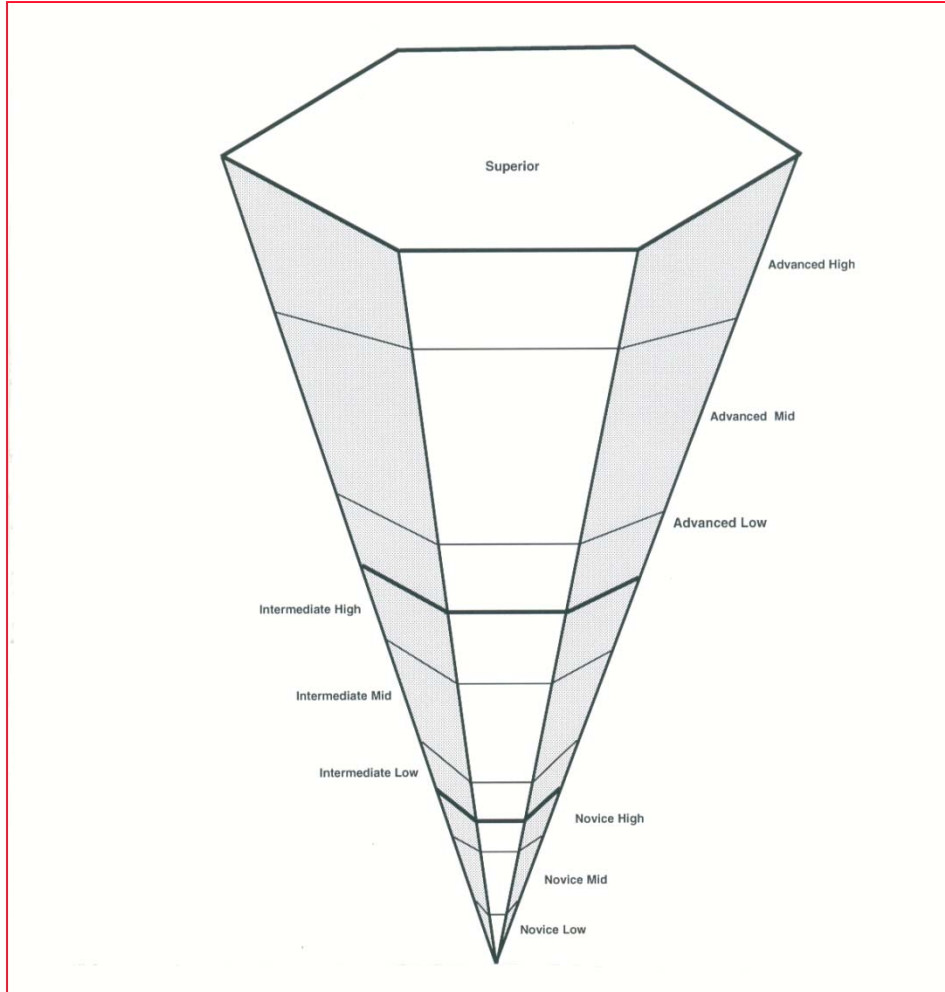


- Study Objective
- About CEFR and ACTFL Scales
- Methodology
- Preliminary Results by Research Question
- Future Directions



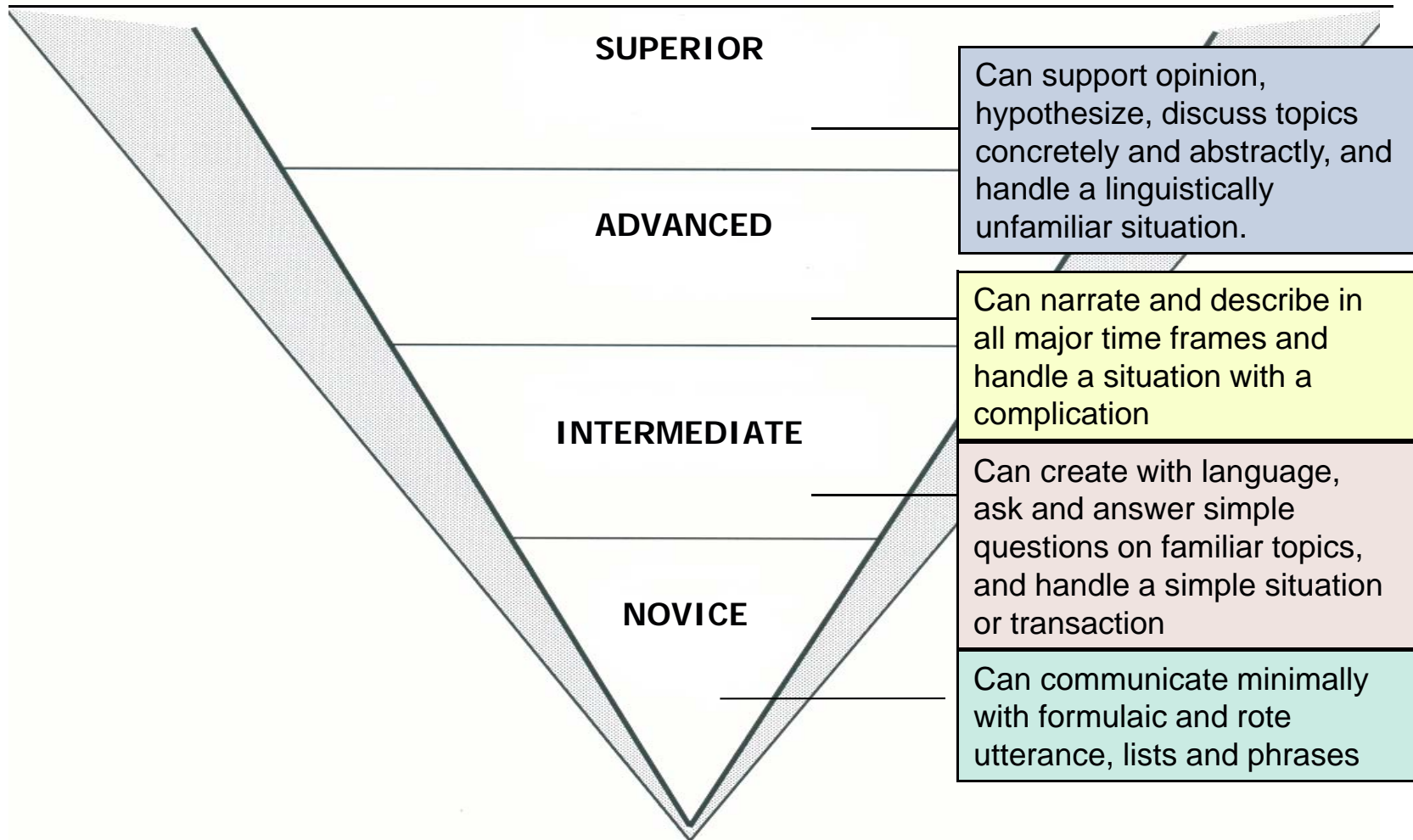
- To compare perceived and actual foreign language speaking and writing proficiency of university students:
  - *Research Question 1:* What is the relationship between *perceived* CEFR proficiency (measured by ‘can-do’ ratings) and *perceived* ACTFL (measured by ‘can-do’ ratings) proficiency?
  - *Research Question 2:* What is the relationship between *perceived* CEFR proficiency (measured by ‘can-do’ ratings) and *actual* ACTFL proficiency (measured by OPIc)?
- Utility of ‘can-do’ ratings as a proxy of actual proficiency
  - *Note: Meta-analysis on relationship of perceived and actual proficiency to presented at 2011 AAAL*

Level	Description
<b>Basic User A1 and A2</b>	
A1 “Breakthrough”	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type...
A2 “Waystage”	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment)...
<b>Independent User B1 and B2</b>	
B1 “Threshold”	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken...
B2 “Vantage”	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialization...
<b>Proficient User C1 and C2</b>	
C1 “Effective Operational Proficiency”	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes...
C2 “Mastery”	Can understand with ease virtually everything heard or read. Can summarize information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation...



- Hierarchy of global tasks
- Four major levels
- Major levels divided into sublevels

# ACTFL Scale



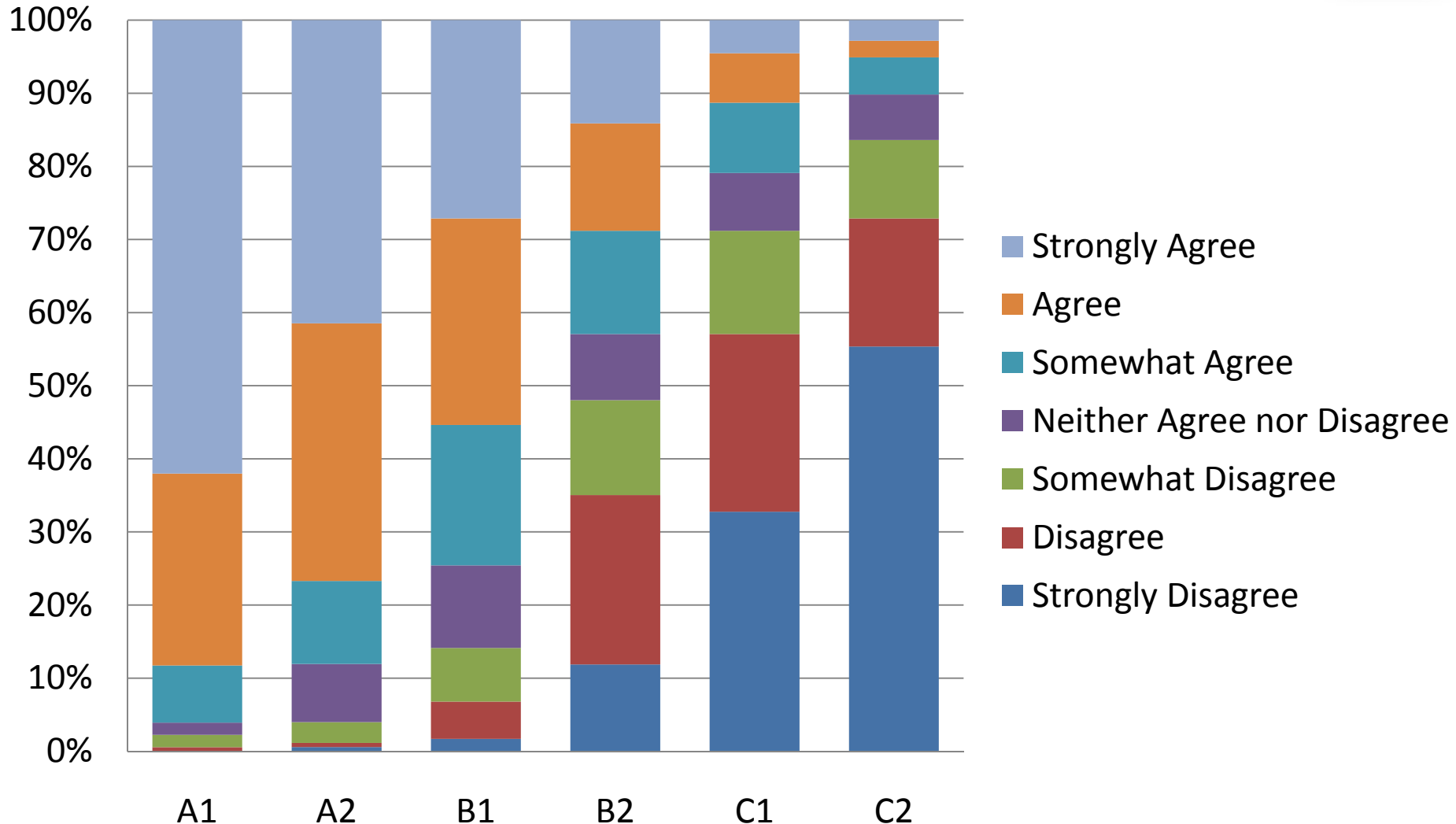


# Speaking Proficiency

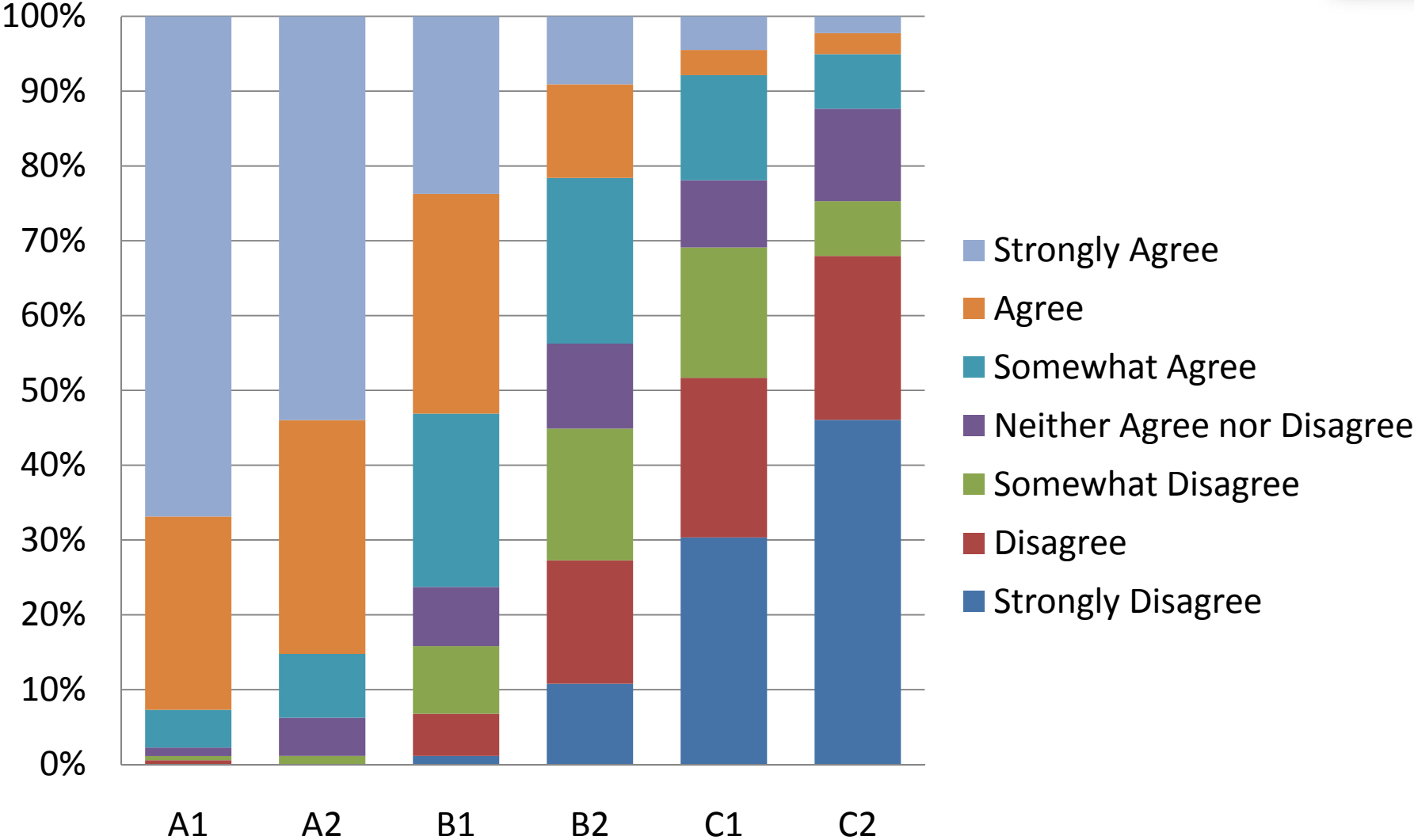


- Follow-up on similar studies presented since 2008
- 237 students from a European university enrolled in Spanish as a Foreign Language
  - Data collected in Spring and Fall from 2007-2010
  - 152 respondents with CEFR ‘can-do’ ratings and ACTFL speaking ‘can-do’ ratings
    - CEFR Spoken Interaction, **CEFR Spoken Production**
  - 31 respondents with CEFR ‘can-do’ ratings and ACTFL OPIc scores
    - CEFR Spoken Interaction, **CEFR Spoken Production**

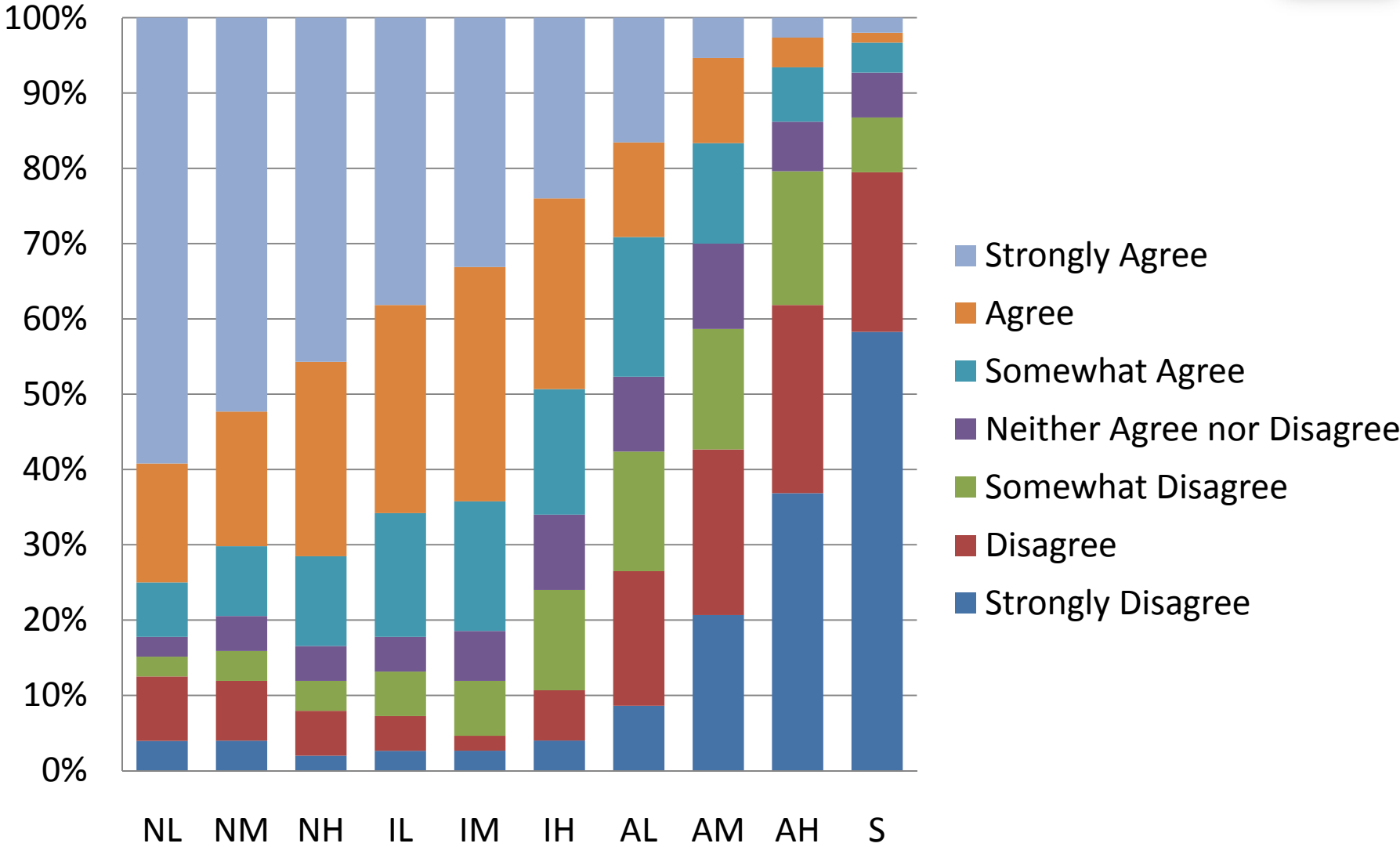
# Perceived CEFR Spoken Interaction Proficiency Distribution



# Perceived CEFR Spoken Production Proficiency Distribution



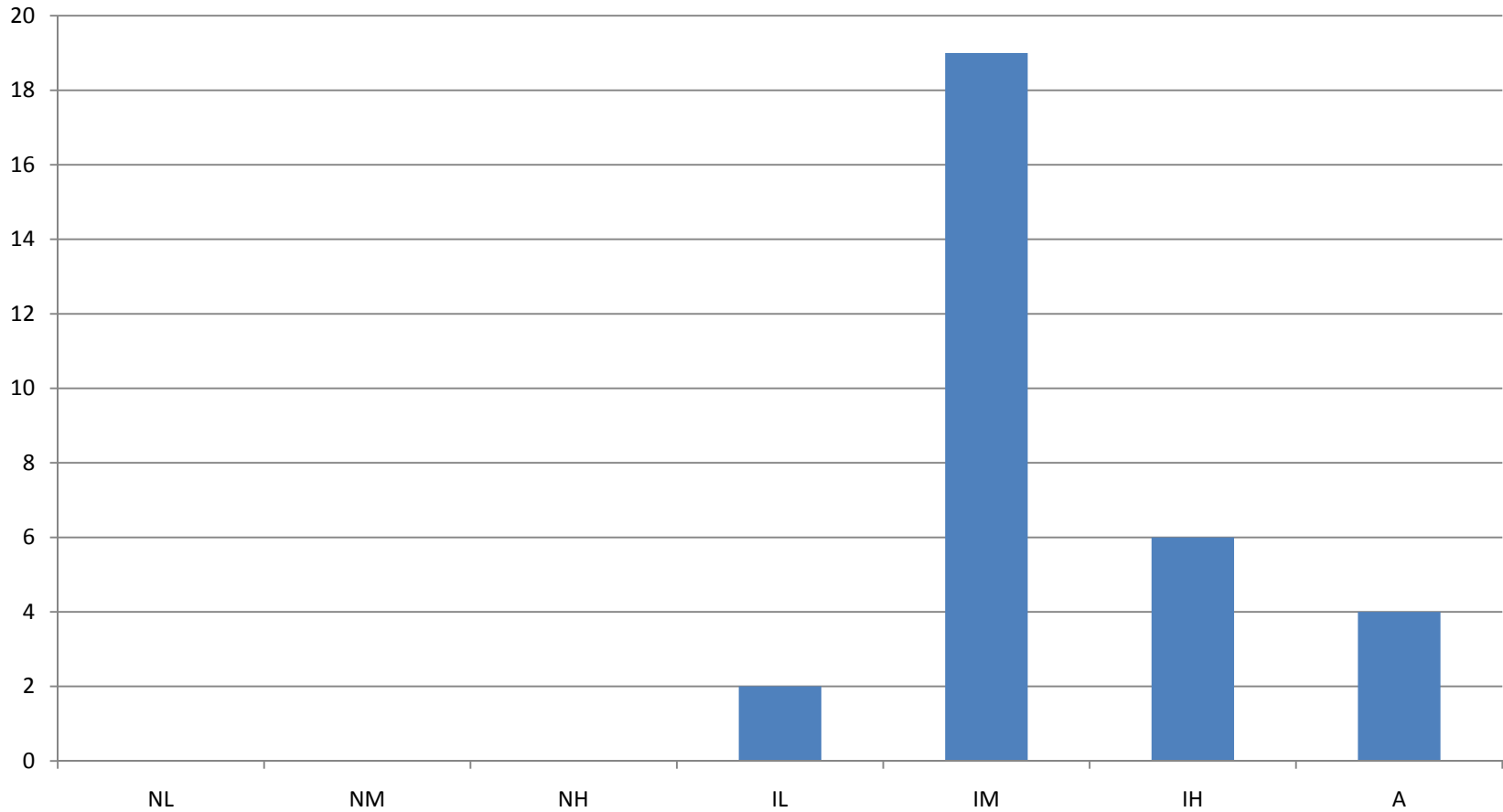
# Perceived ACTFL Speaking Proficiency Distribution



# ACTFL OPIc Rating Distribution



OPIc Rating ( $n = 31$ )



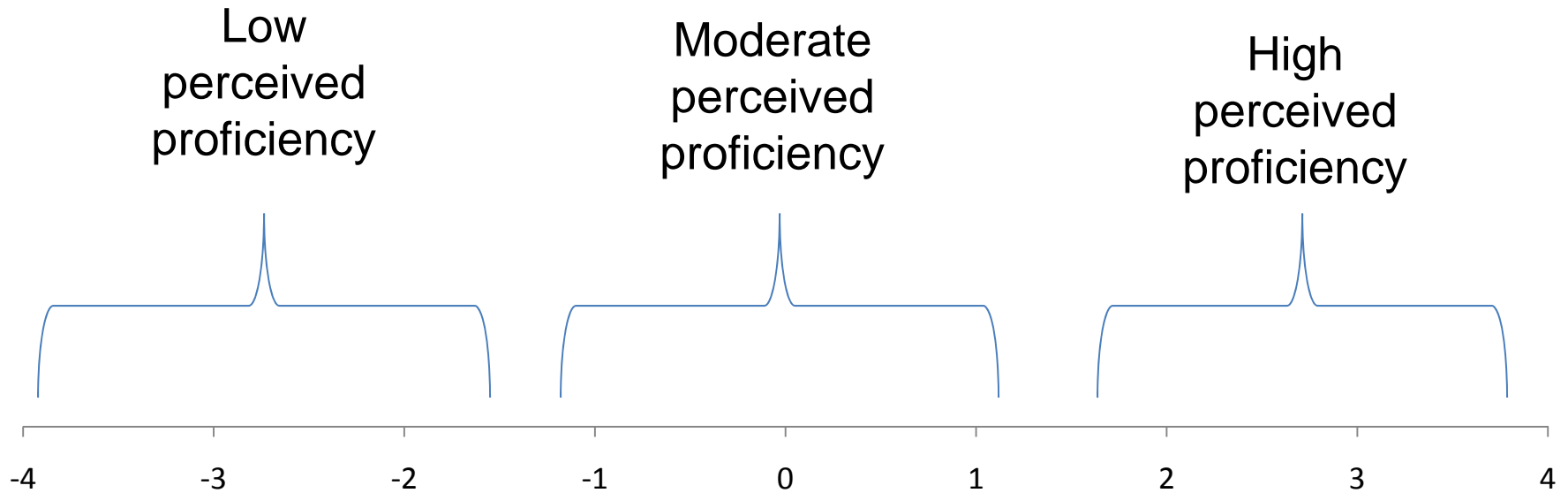


- What is the relationship between *perceived* CEFR proficiency (measured by ‘can-do’ ratings) and *perceived* ACTFL proficiency (measured by ‘can-do’ ratings)?
  - RQ1: CEFR Spoken Production ↔ ACTFL Speaking



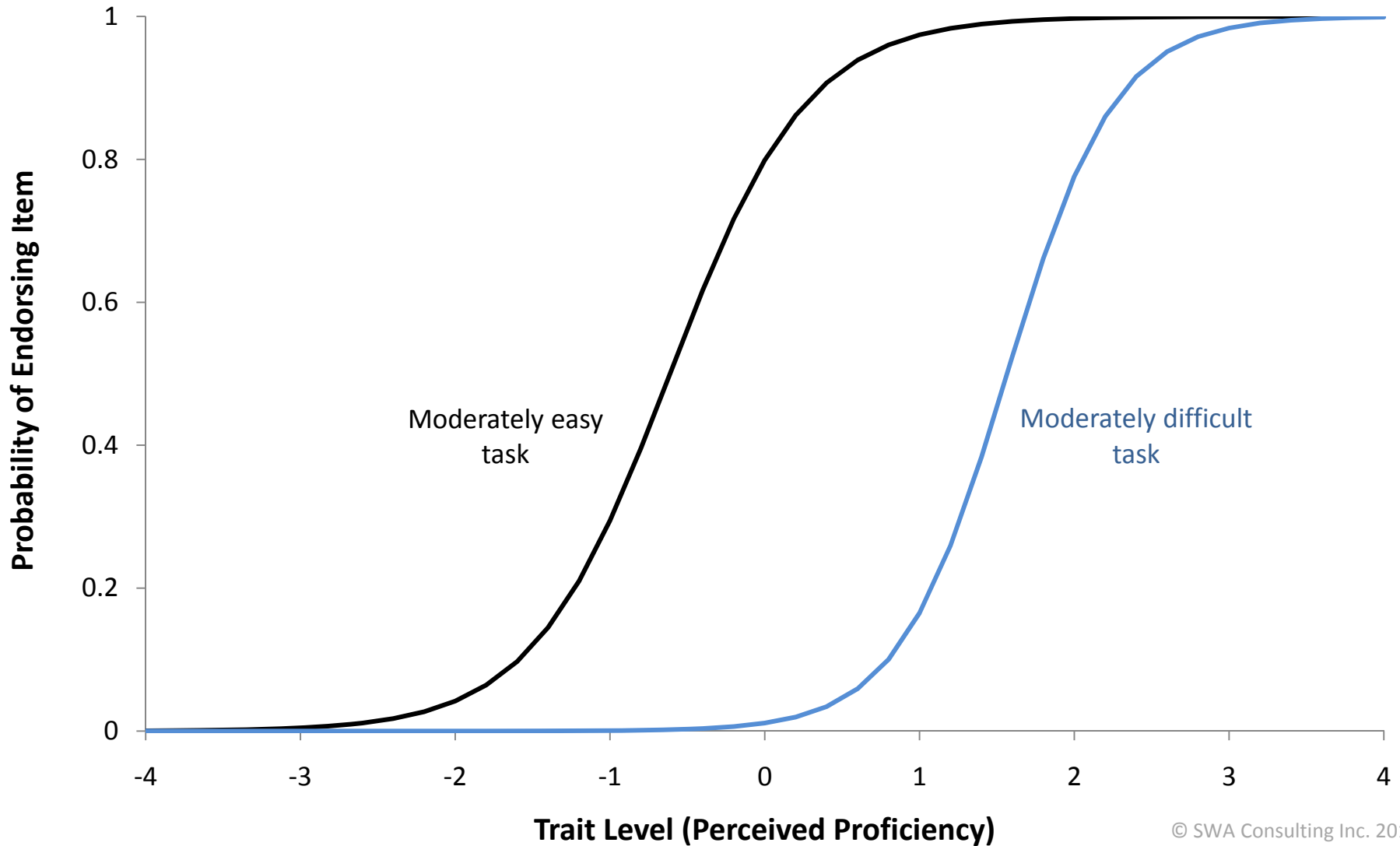
- Placed items and respondents' perceived proficiency on same metric using item response modeling (IRT).
  - Results interpreted with caution given the small sample
- Looked at how items from different frameworks map onto the perceived capability construct.

# Introduction



**“Trait Level” (Perceived Proficiency)**

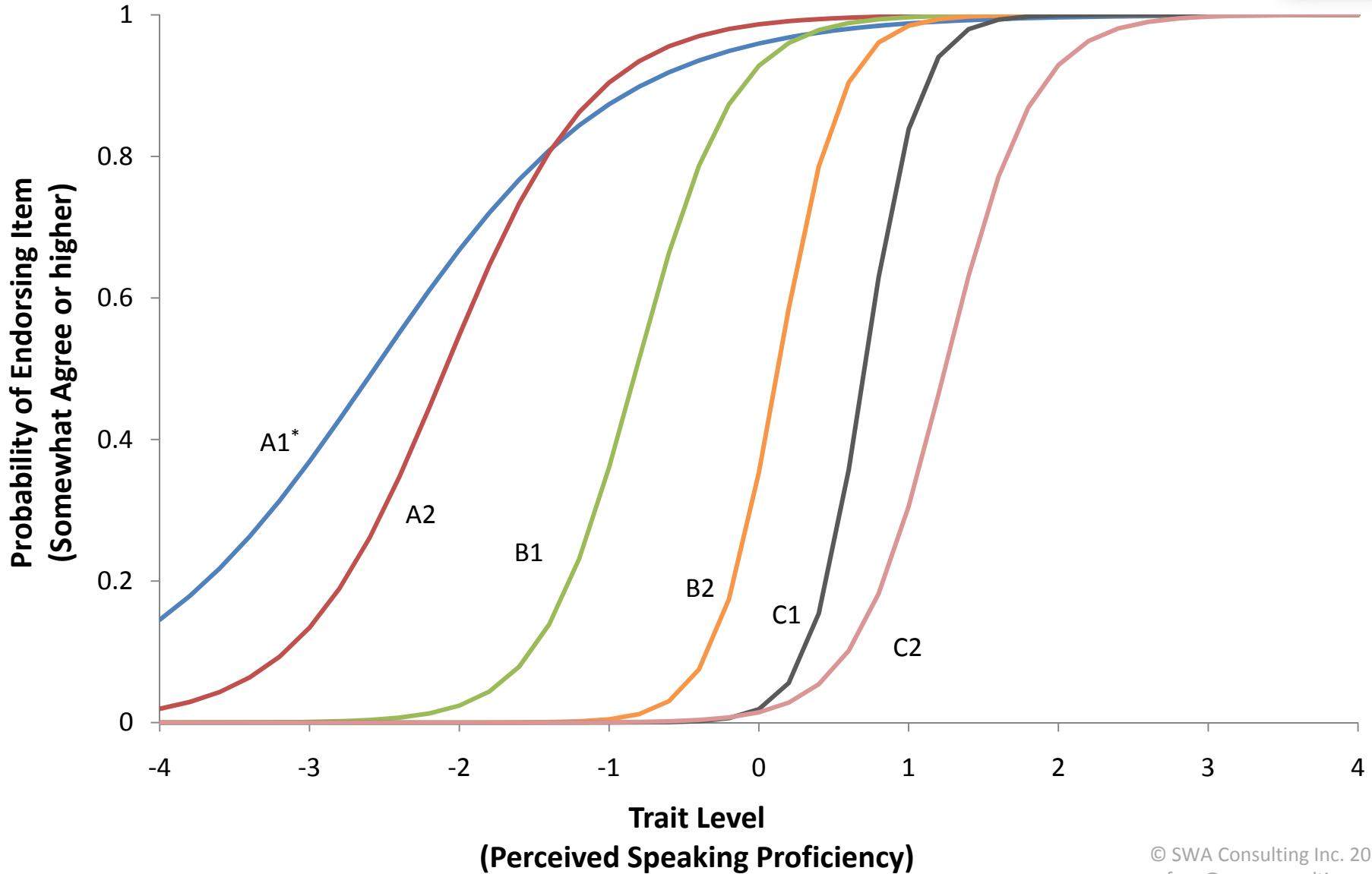
# Introduction (cont.)





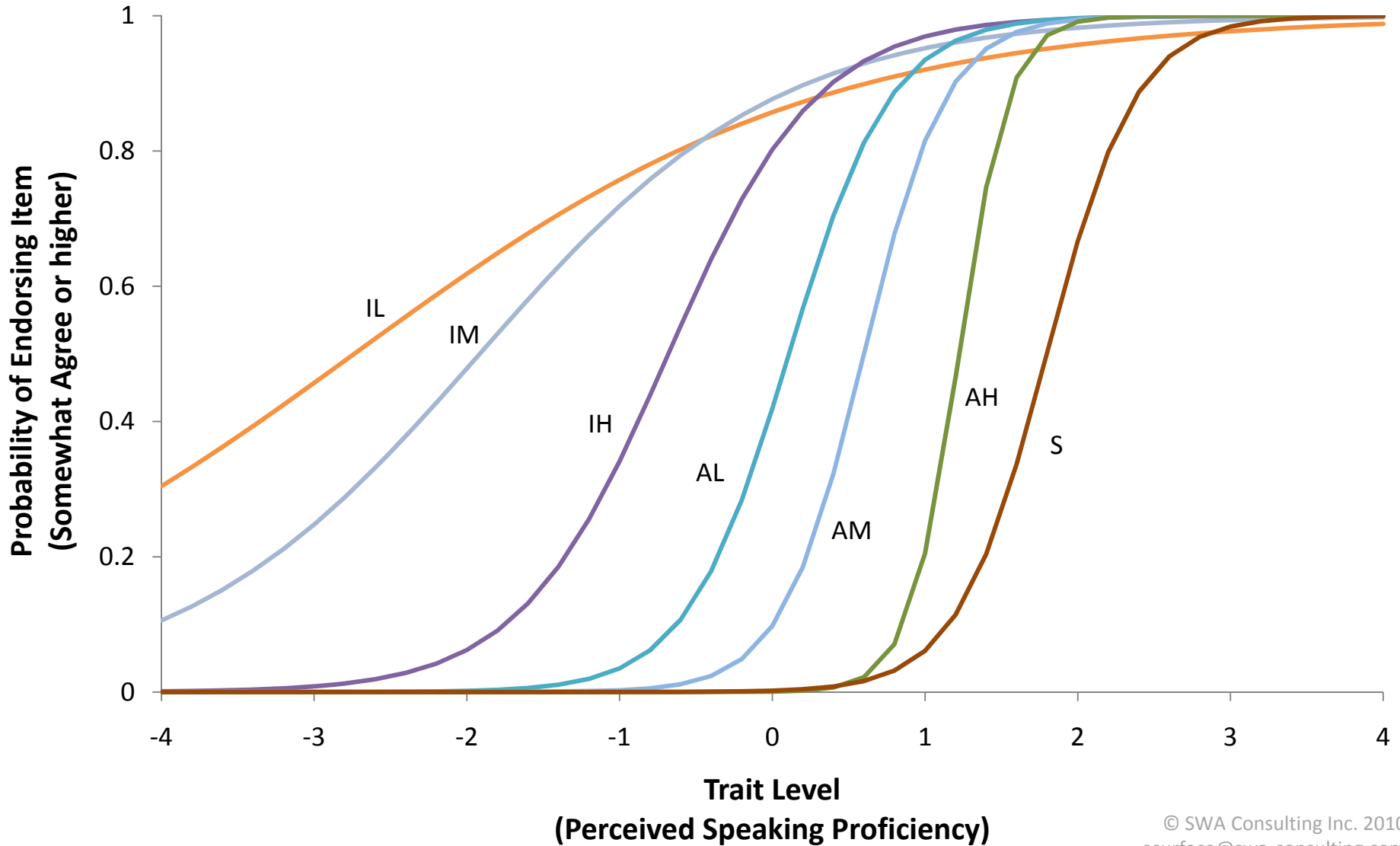
# RQ1: CEFR Spoken Production ↔ ACTFL Speaking

# CEFR Spoken Production 'Can-Do' Statements



\* Endorse = "Agree" or higher

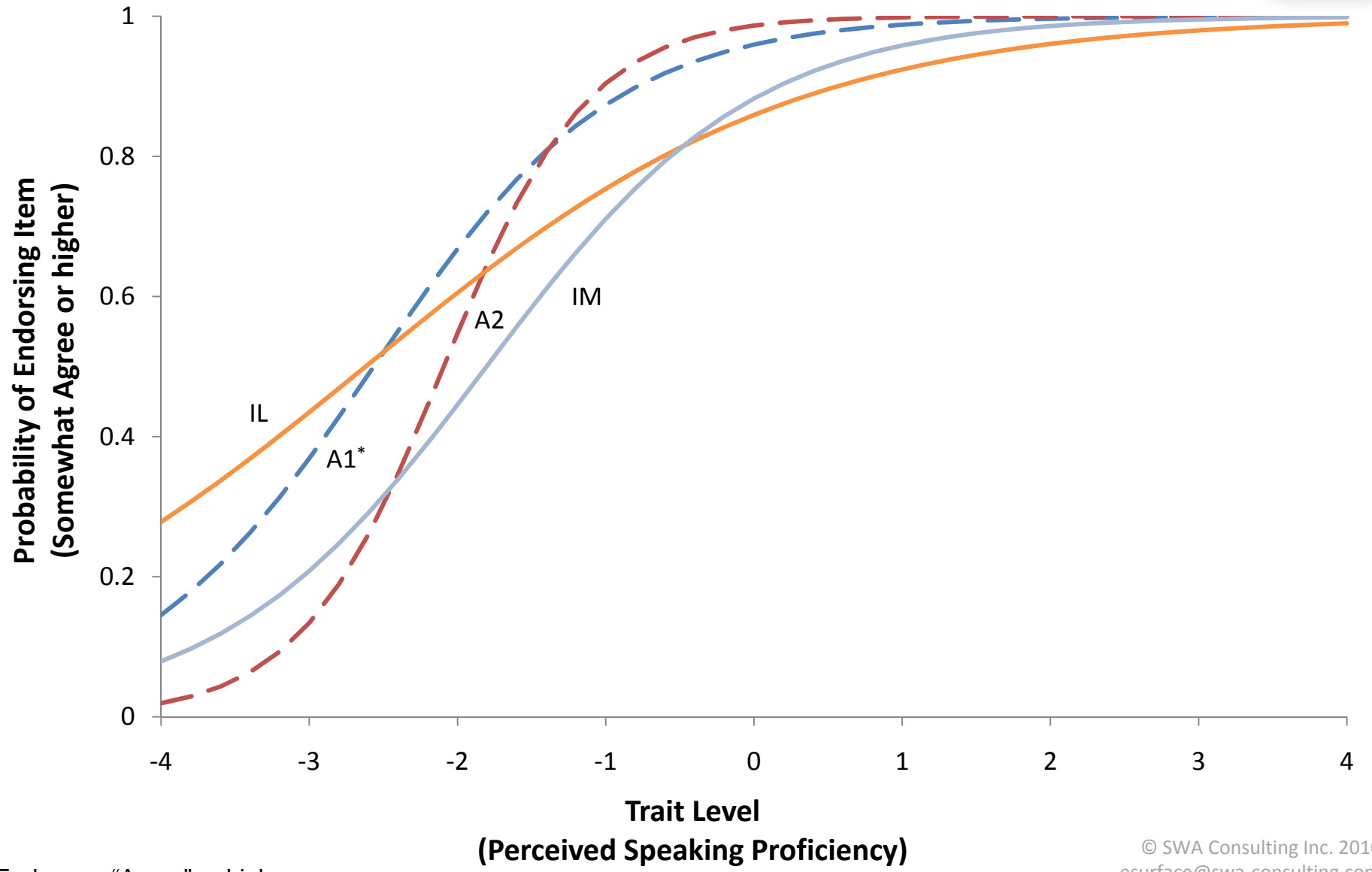
# ACTFL Speaking 'Can-Do' Statements



# Framework Comparisons: CEFR Spoken Production ↔ ACTFL Speaking



Universiteit Utrecht

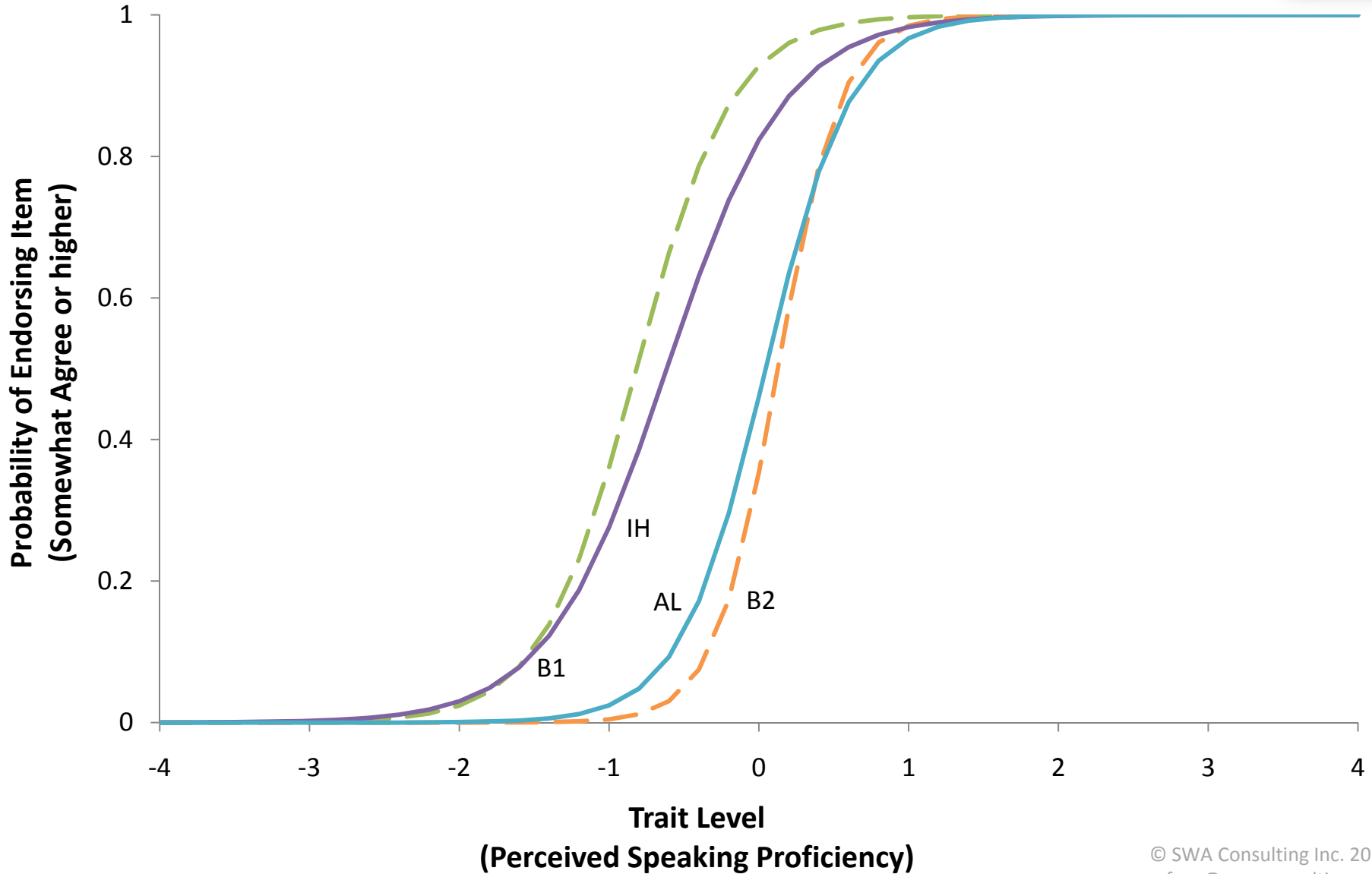


\* Endorse = "Agree" or higher

# Framework Comparisons: CEFR Spoken Production ↔ ACTFL Speaking



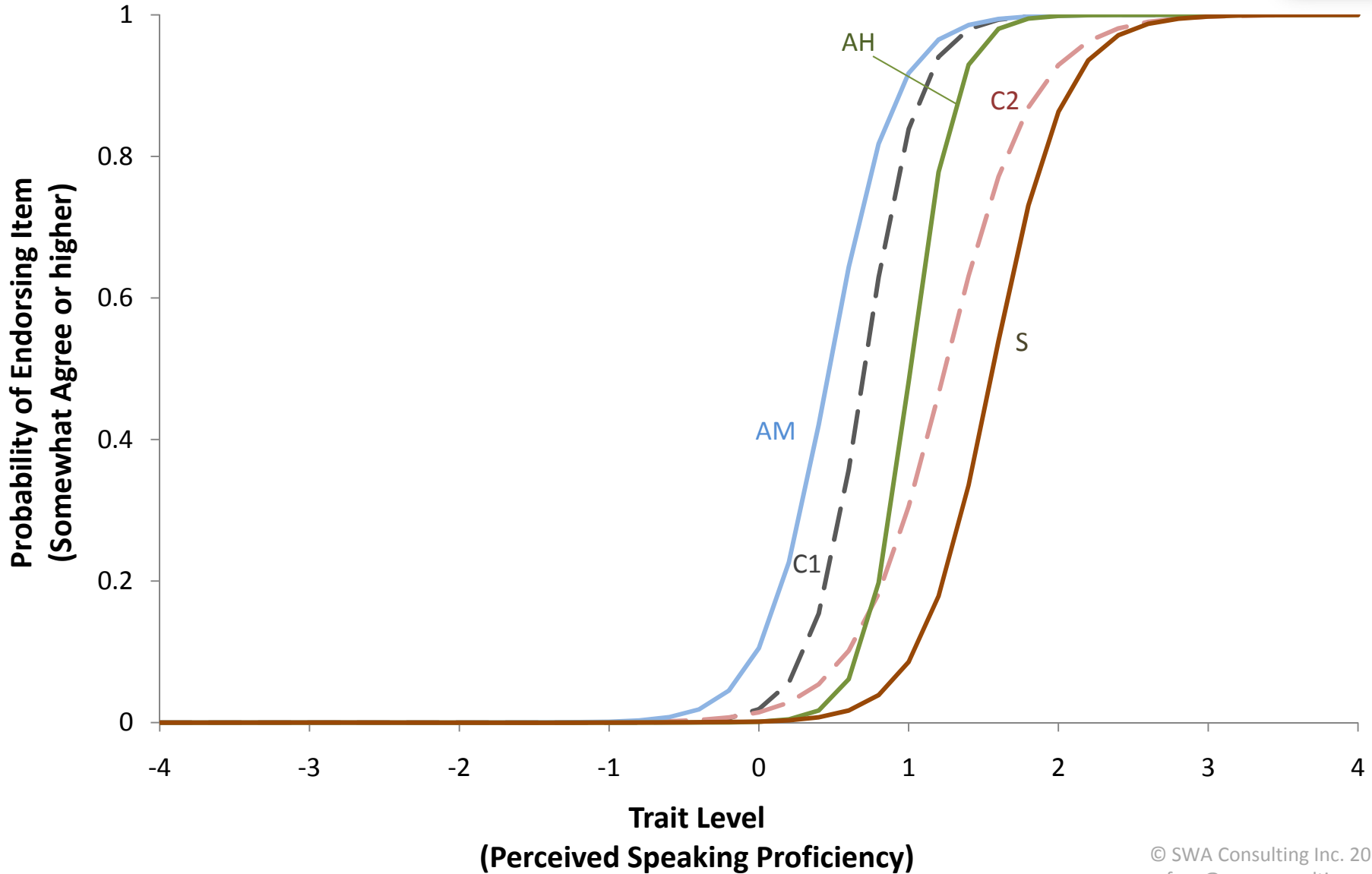
Universiteit Utrecht



# Framework Comparisons: CEFR Spoken Production ↔ ACTFL Speaking



Universiteit Utrecht



# Perceived Speaking Proficiency Across Frameworks



Universiteit Utrecht



## Perceived CEFR Spoken Production



## Perceived ACTFL Speaking

*Note.* Spacing is intentional.

<sup>a</sup> Novice items could not be modeled because all respondents endorsed them.

# Perceived Speaking Proficiency Across Frameworks



Universiteit Utrecht

SWA

Consulting

## Perceived CEFR Spoken Interaction



## Perceived ACTFL Speaking

*Note.* Spacing is intentional.

<sup>a</sup> Novice items could not be modeled because all respondents endorsed them.



# RQ2: CEFR Spoken Production

## ↔ ACTFL OPIc



- What is the relationship between *perceived* CEFR proficiency (measured by ‘can-do’ ratings) and *actual* ACTFL proficiency (measured by OPIc)?
  - RQ2: CEFR Spoken Production ↔ ACTFL OPIc



- Placed CEFR items and respondents' perceived proficiency on same metric using item response modeling (IRT).
  - Results interpreted with caution given the small sample
- Looked at how actual speaking proficiency on the OPIc mapped onto the perceived capability constructs (i.e., interaction and **production**).

$n = 22$   
Spearman  $R = .43$  ( $p < .05$ )

OPIc Rating

AL

IH

IM

IL

-2

-1

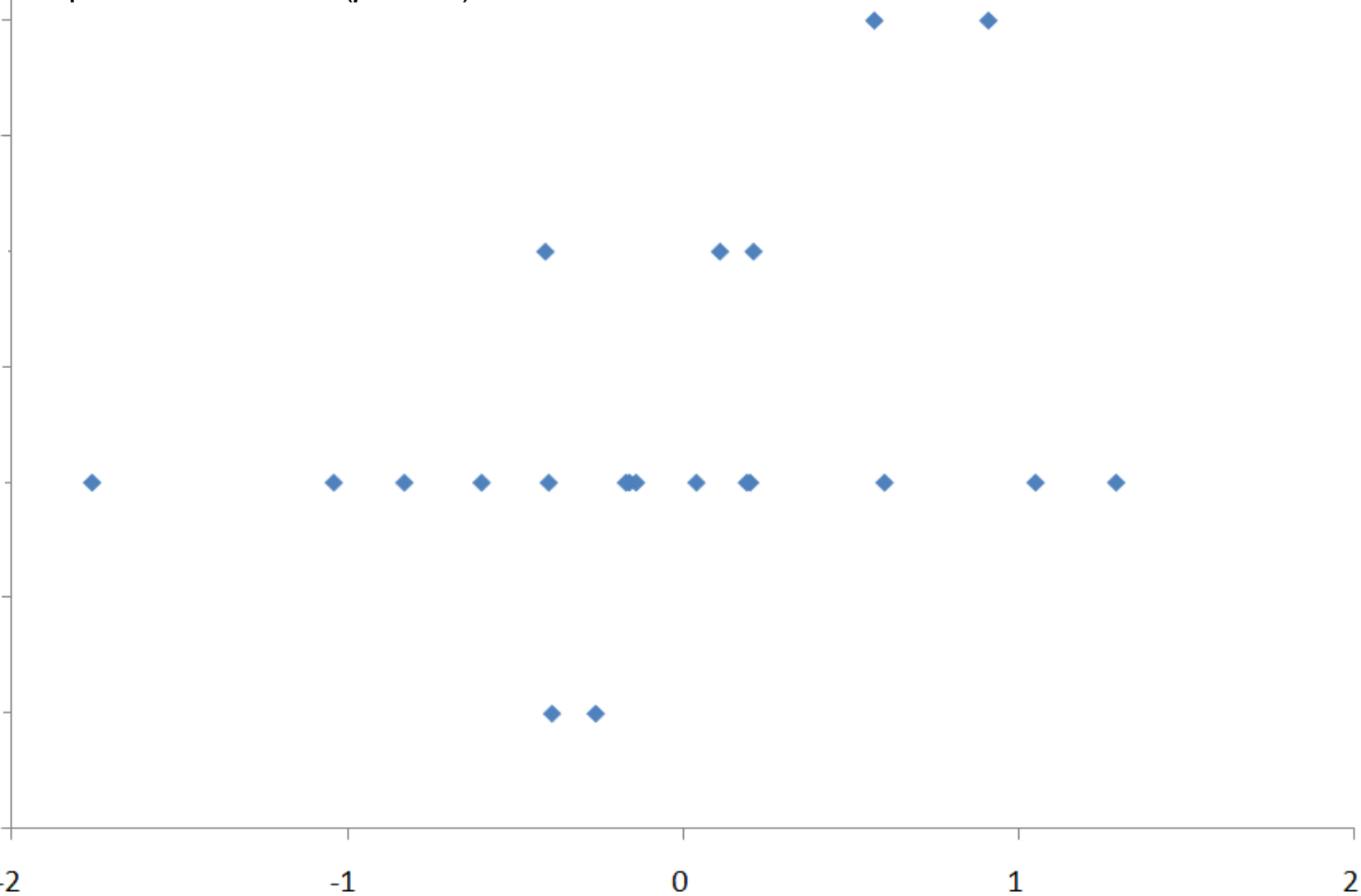
0

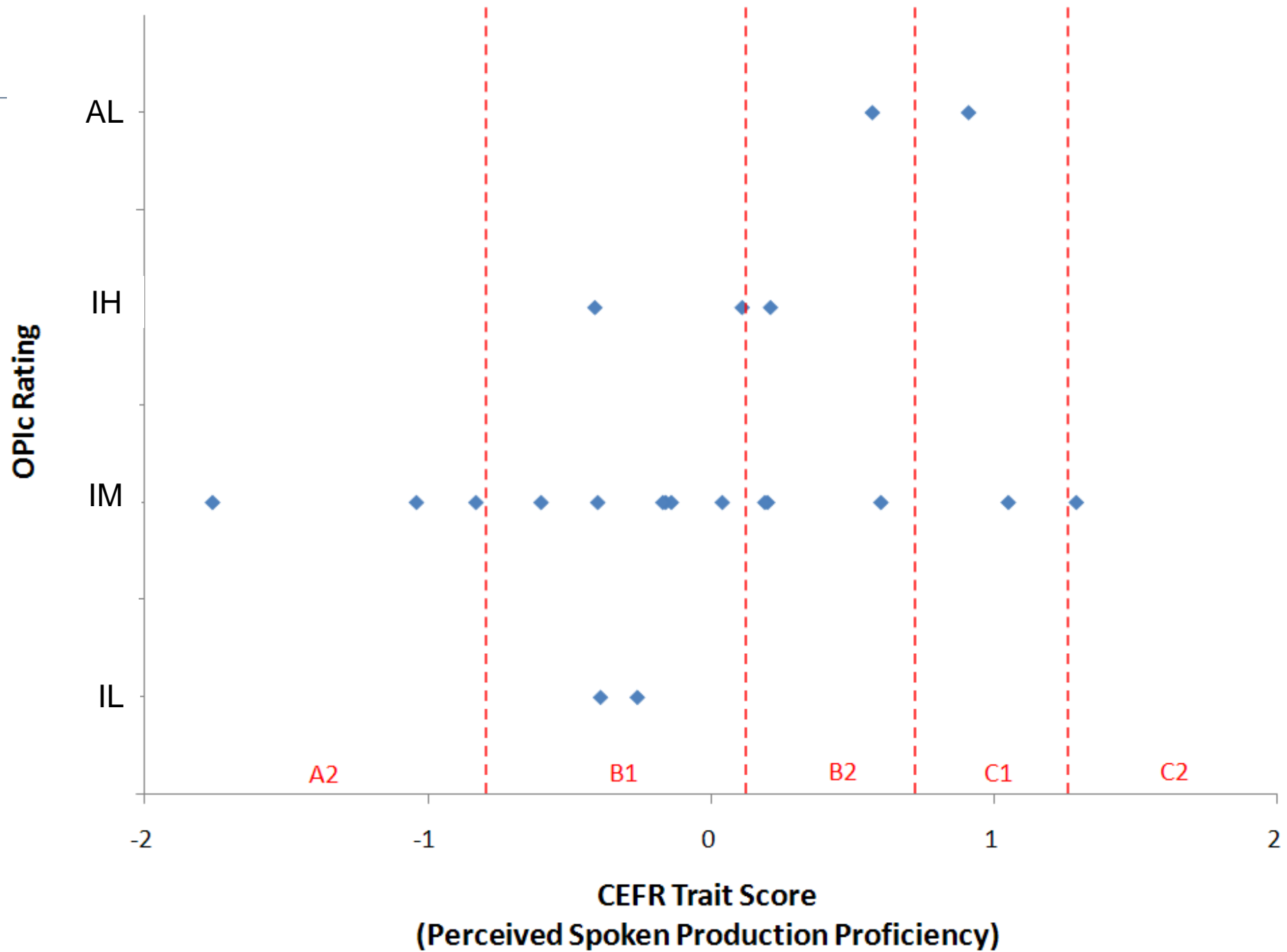
1

2

CEFR Trait Score

(Perceived Spoken Production Proficiency)





# Perceived Spoken Production Proficiency and Actual OPIC Rating



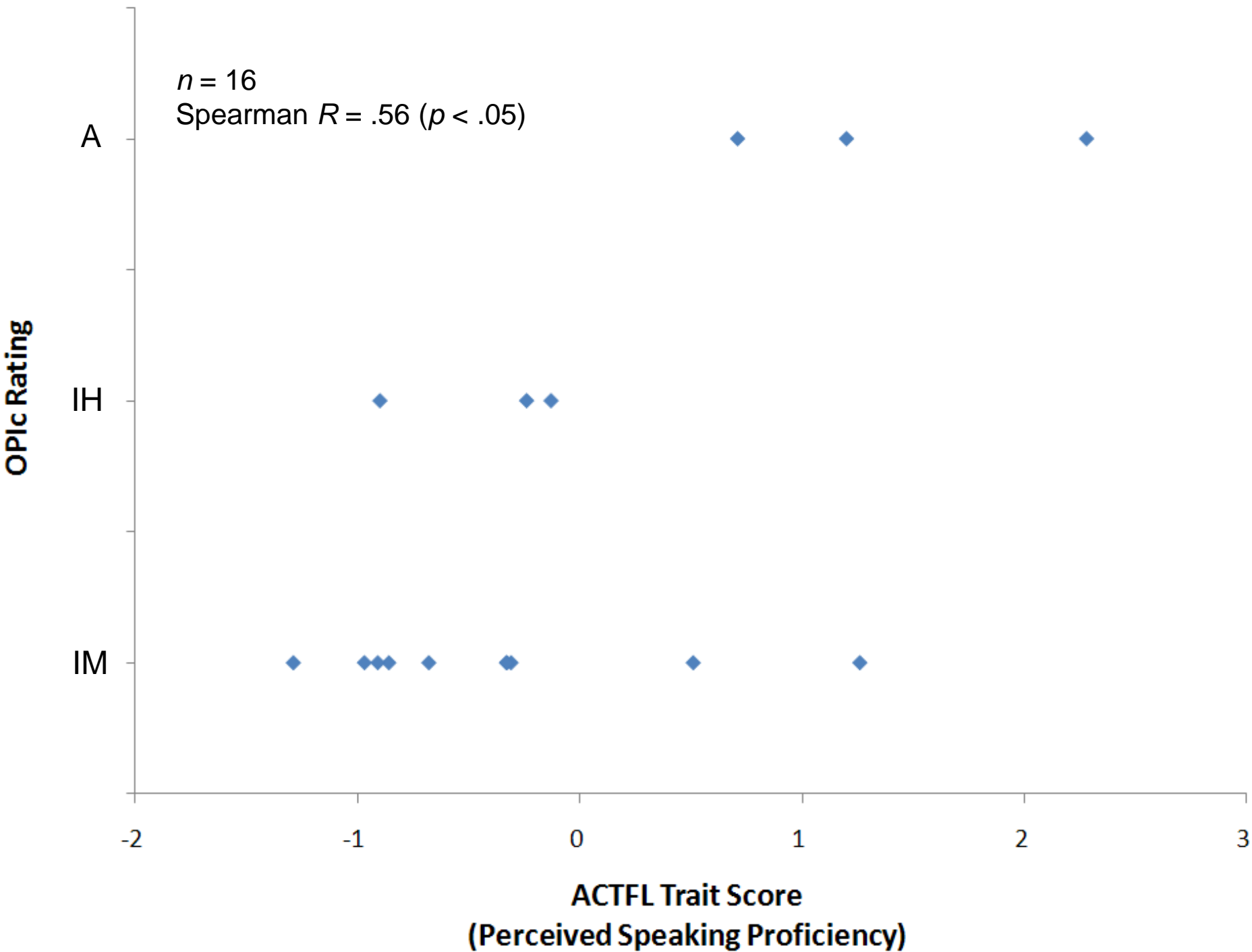
## Actual OPIC Rating

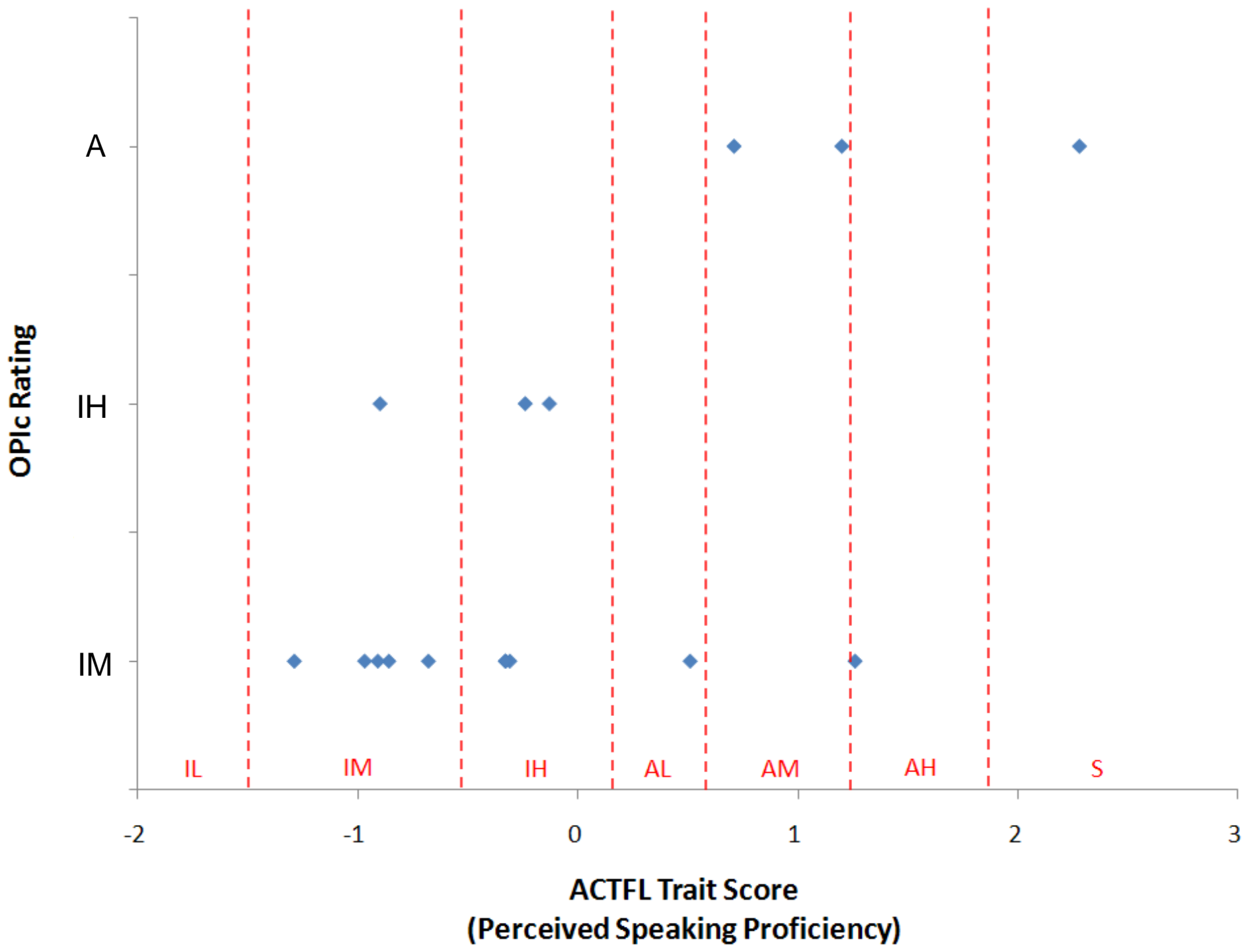
		IL	IM	IH	A	Total
Perceived Spoken Production Proficiency (CEFR)	A1	-	-	-	-	0
		-	-	-	-	0%
	A2	-	4	-	-	4
		-	29%	-	-	18%
	B1	2	8	3	-	13
		100%	57%	100%	-	59%
	B2	-	2	-	2	4
		-	14%	-	67%	18%
	C1	-	-	-	-	0
		-	-	-	-	0%
	C2	-	-	-	1	1
		-	-	-	33%	5%
<i>Total</i>		2	14	3	3	22
		100%	100%	100%	100%	100%

% = percentages for column



# **Perceived ACTFL Speaking Proficiency vs. Actual ACTFL Speaking Proficiency**





# Perceived Spoken Interaction Proficiency and Actual OPIC Rating



## Actual OPIC Rating

		IL	IM	IH	A	Total
Perceived Speaking Proficiency (ACTFL)	IL	-	-	-	-	0
		-	-	-	-	0%
	IM	-	5	1	-	6
		-	50%	33%	-	38%
	IH	-	3	2	-	5
		-	30%	67%	-	31%
	AL	-	1	-	-	1
		-	10%	-	-	6%
AM	-	-	-	2	2	
	-	-	-	67%	13%	
AH	-	1	-	-	1	
	-	10%	-	-	6%	
S	-	-	-	1	1	
	-	-	-	33%	6%	
<i>Total</i>		0	10	3	3	16

% = percentages for column



# Writing Proficiency

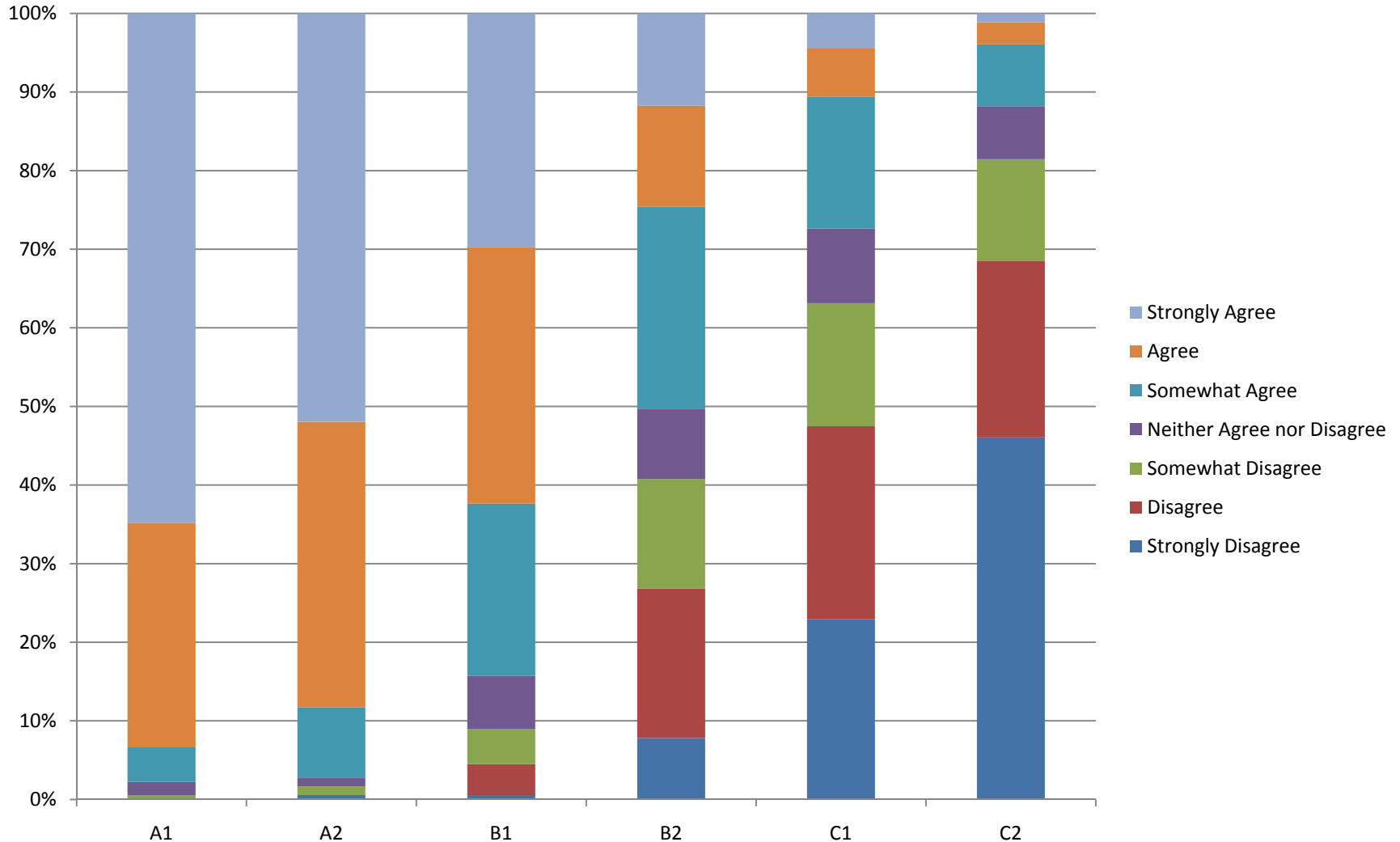


# Perceived CEFR Writing Proficiency vs. Perceived ACTFL Writing

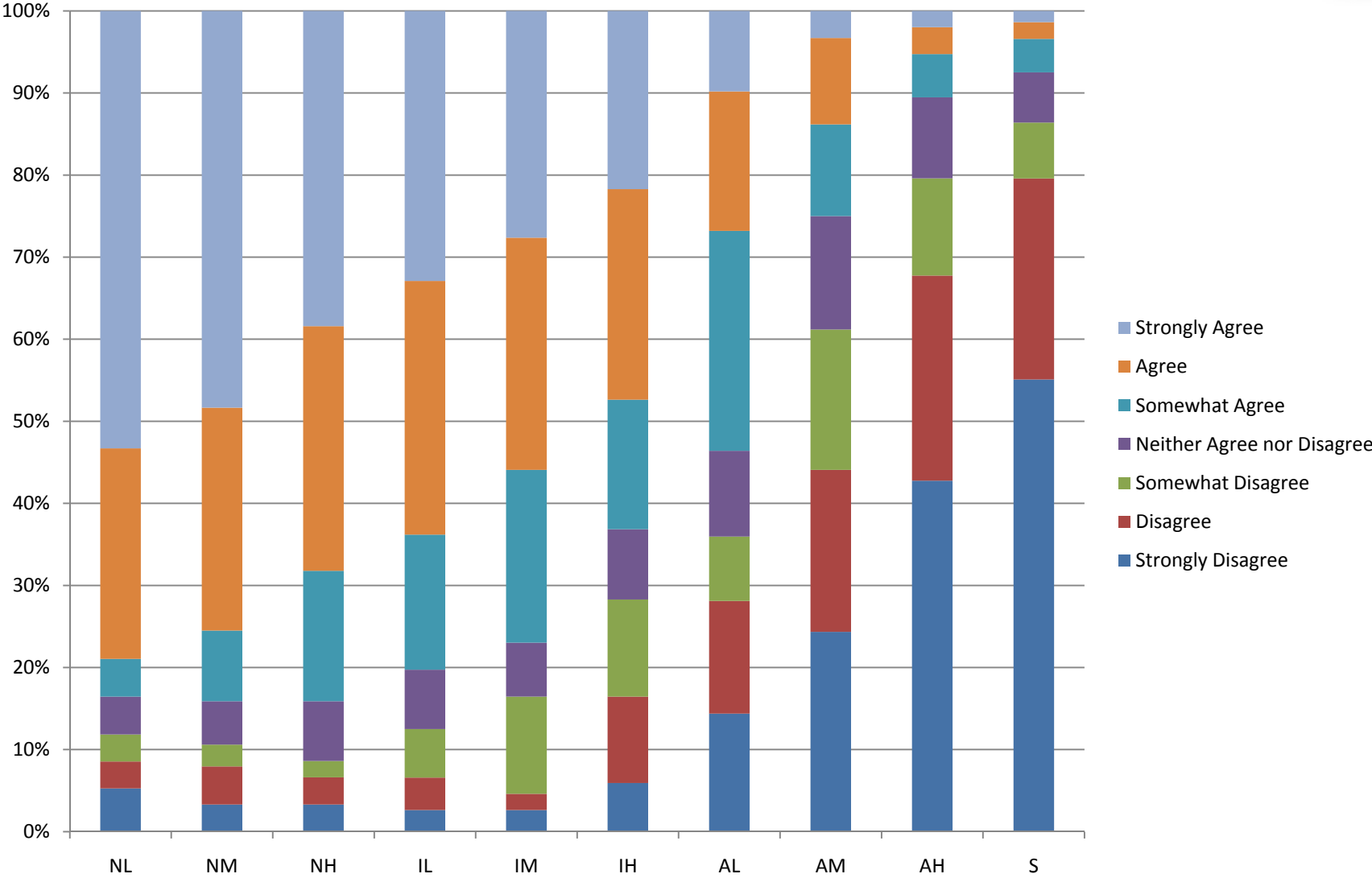


- Follow-up on similar studies presented since 2008
- 237 students from a European university enrolled in Spanish as a Foreign Language
  - Data collected in Spring and Fall from 2007-2010
  - 153 respondents with CEFR Writing ‘can-do’ ratings and ACTFL Writing ‘can-do’ ratings
  - 99 respondents with CEFR Writing ‘can-do’ ratings and ACTFL WPT scores

# Perceived CEFR Writing Proficiency



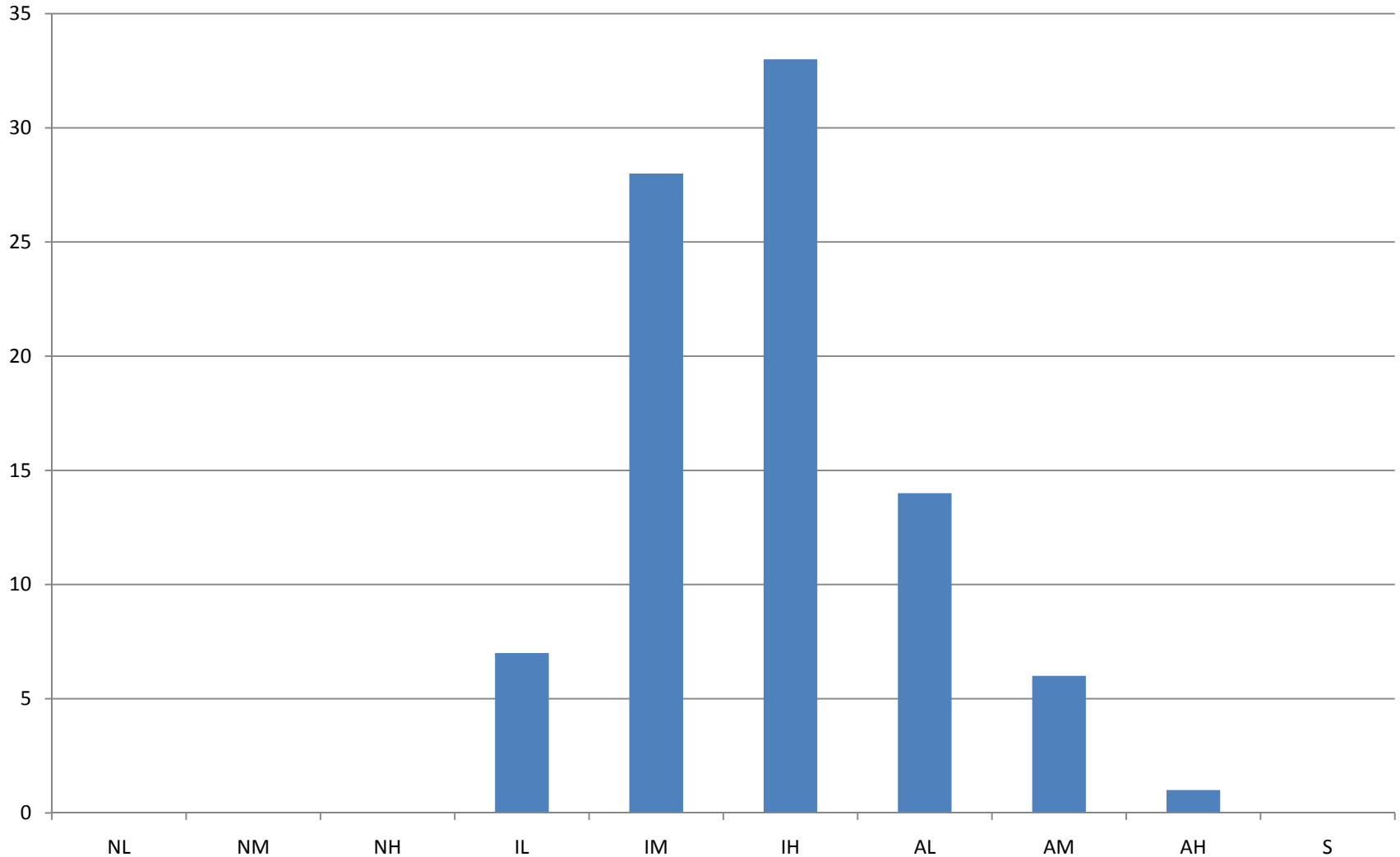
# Perceived ACTFL Writing Proficiency



# ACTFL WPT Ratings



Universiteit Utrecht



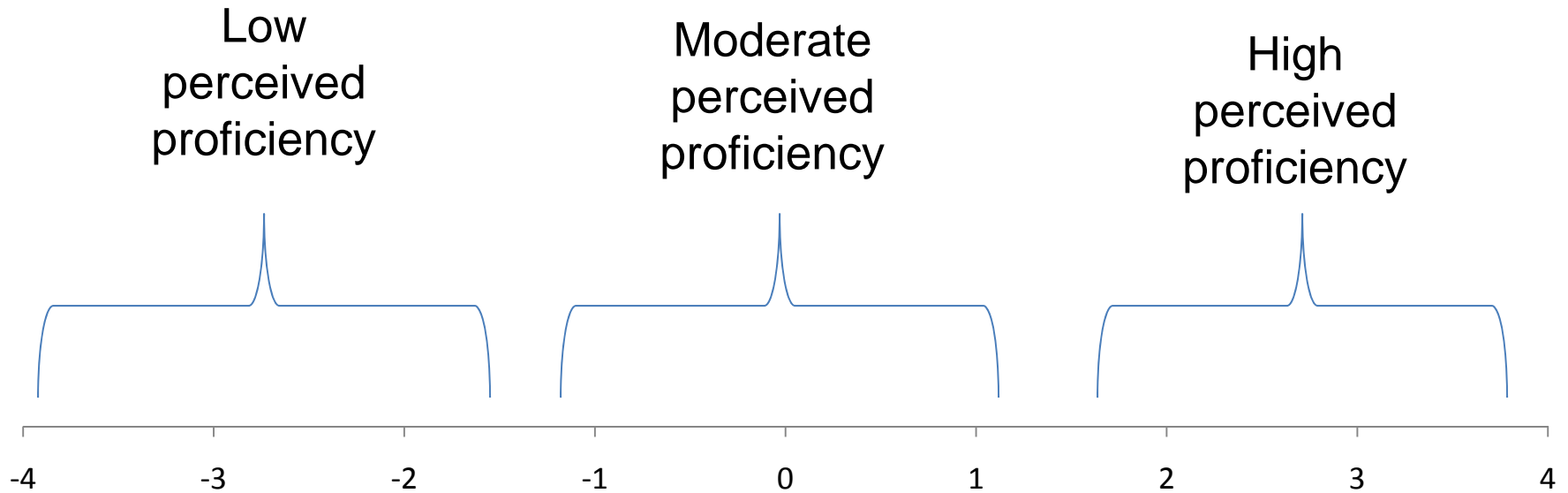


- What is the relationship between *perceived* CEFR writing proficiency (measured by ‘can-do’ ratings) and *perceived* ACTFL writing proficiency (measured by ‘can-do’ ratings)?



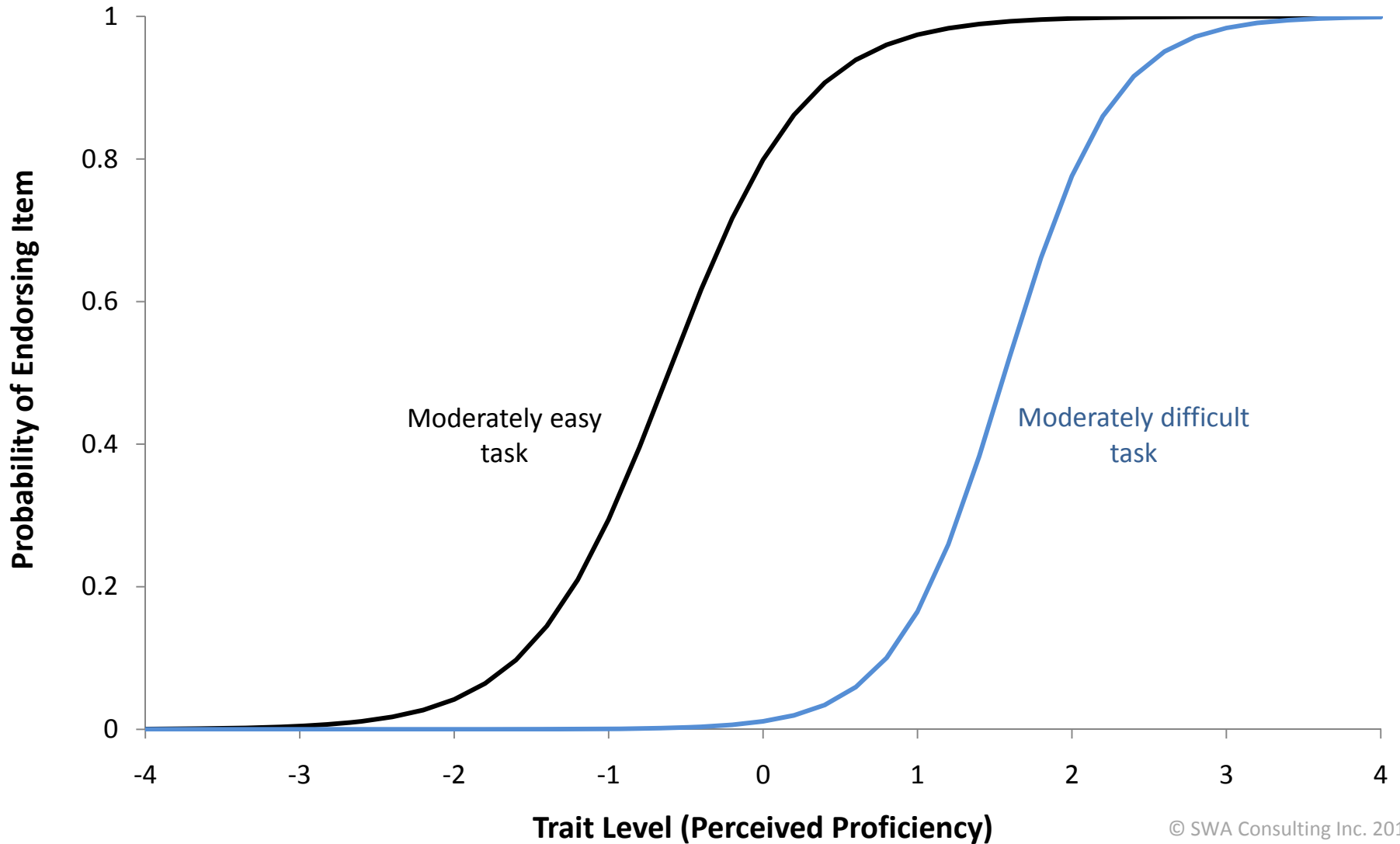
- Placed items and respondents' perceived proficiency on same metric using item response modeling (IRT).
  - Results interpreted with caution given the small sample
- Looked at how items from different frameworks map onto the perceived capability construct.

# Introduction

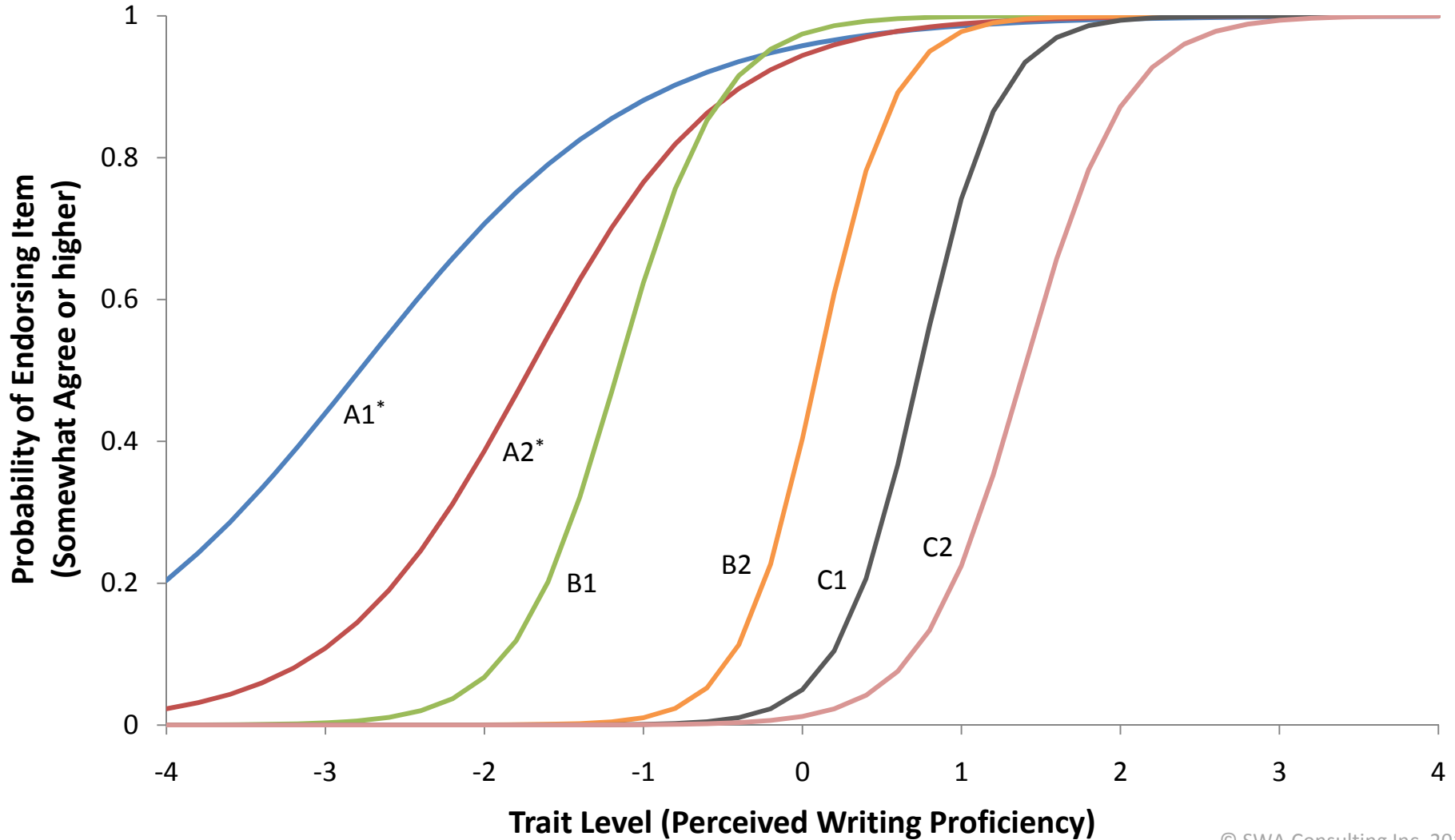


**“Trait Level” (Perceived Proficiency)**

# Introduction (cont.)

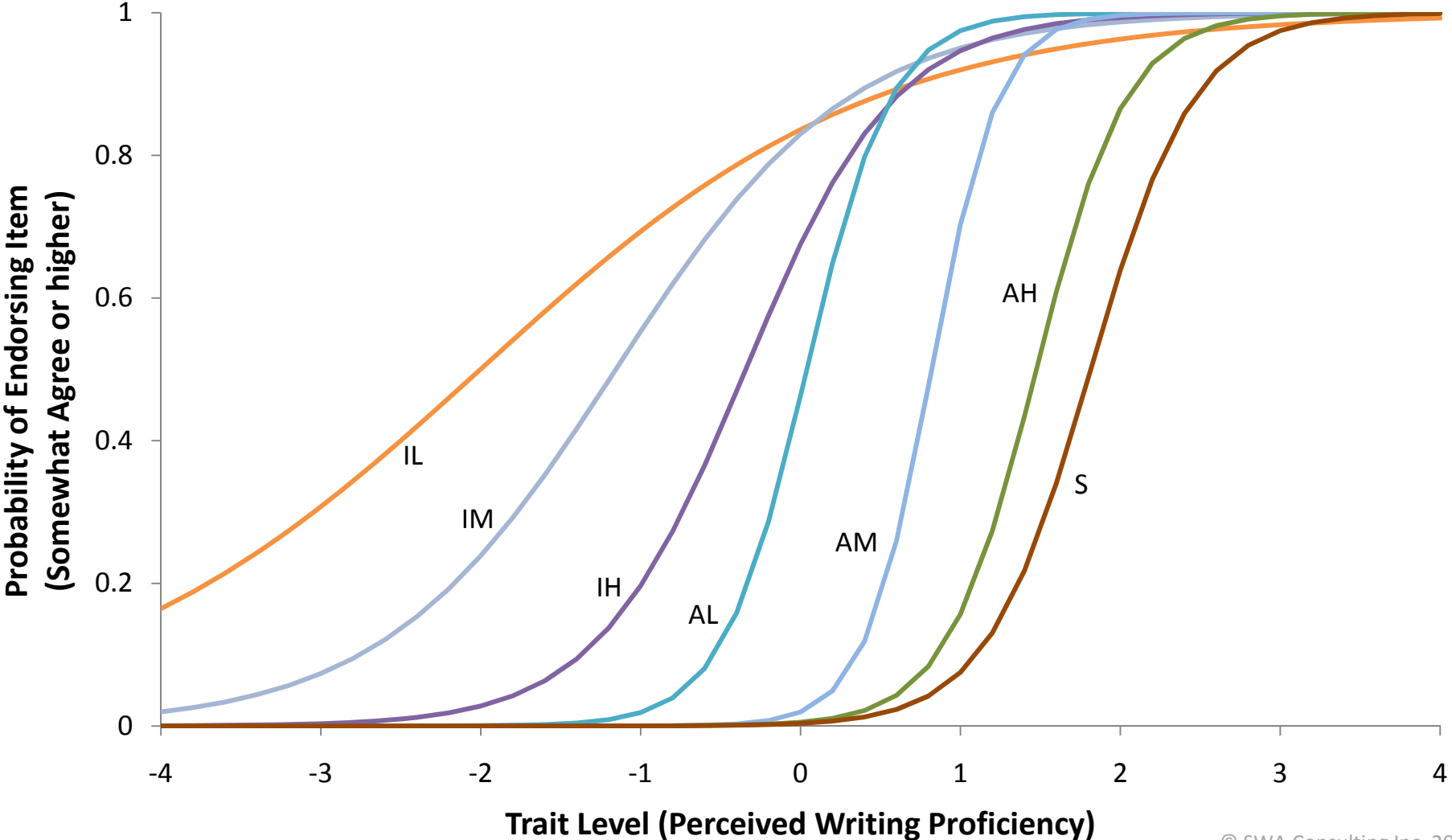


# CEFR Writing 'Can-Do' Statements

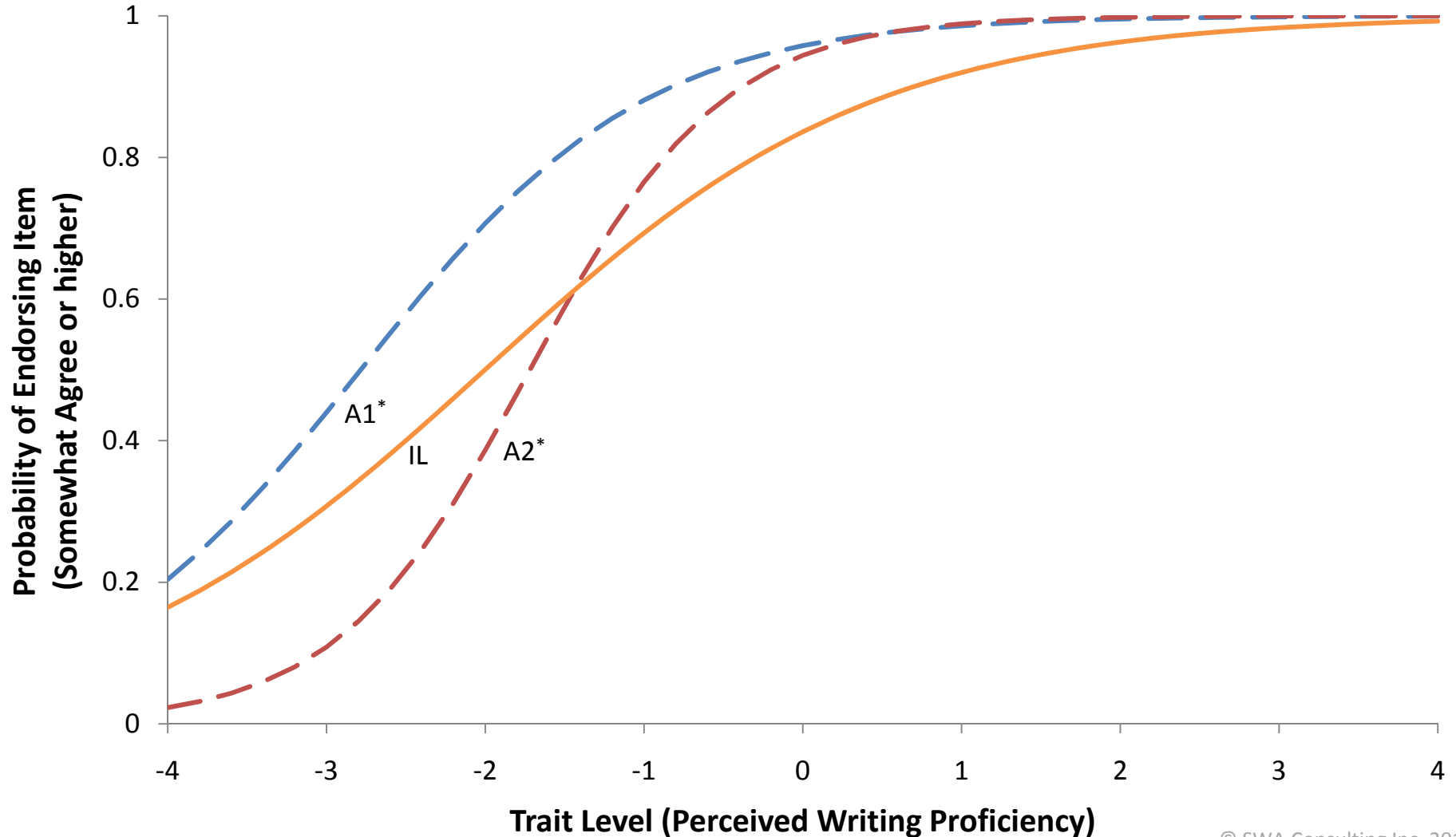


\* Endorse = "Agree" or higher

# ACTFL Writing 'Can-Do' Statements

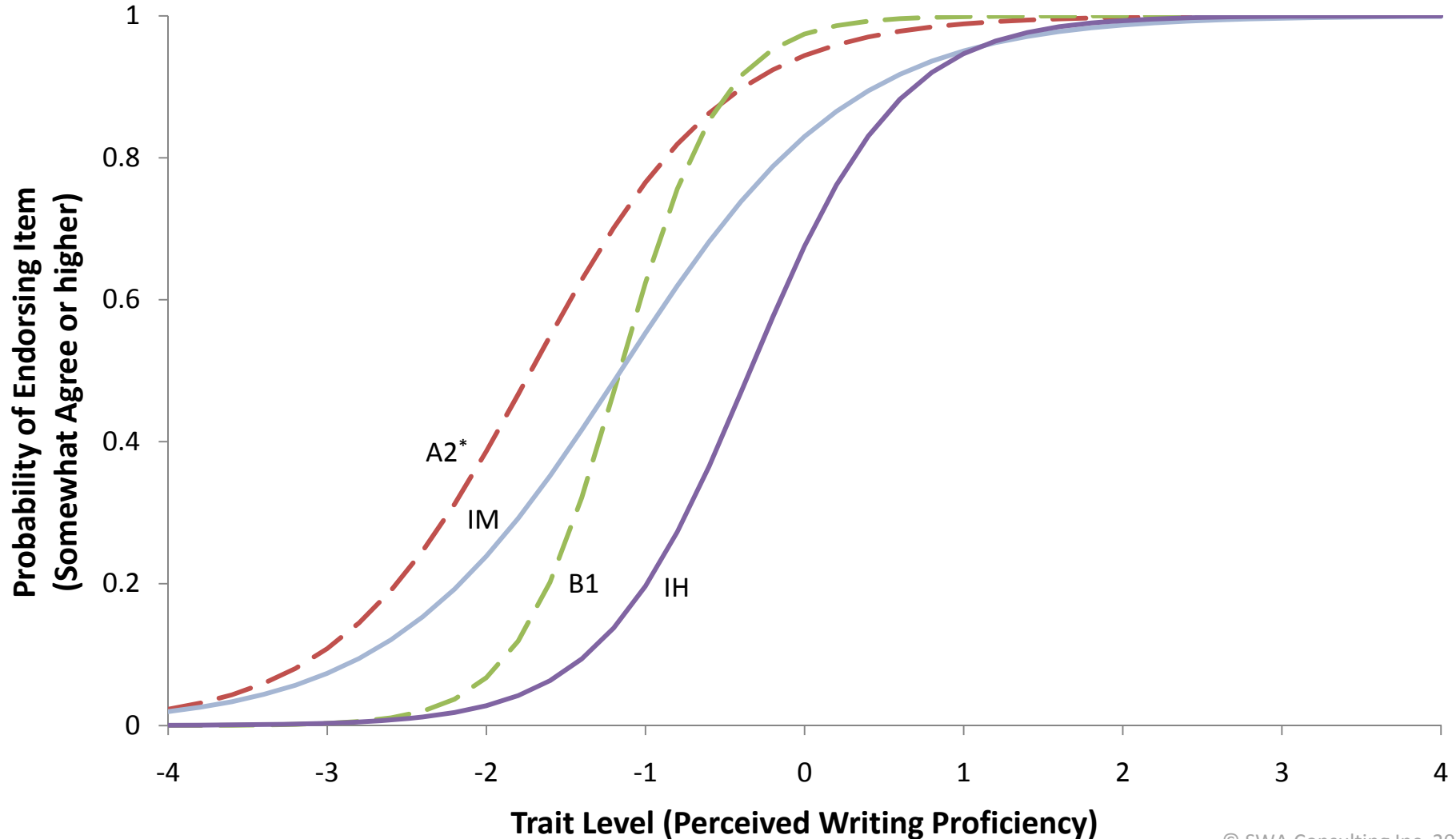


# Framework Comparisons



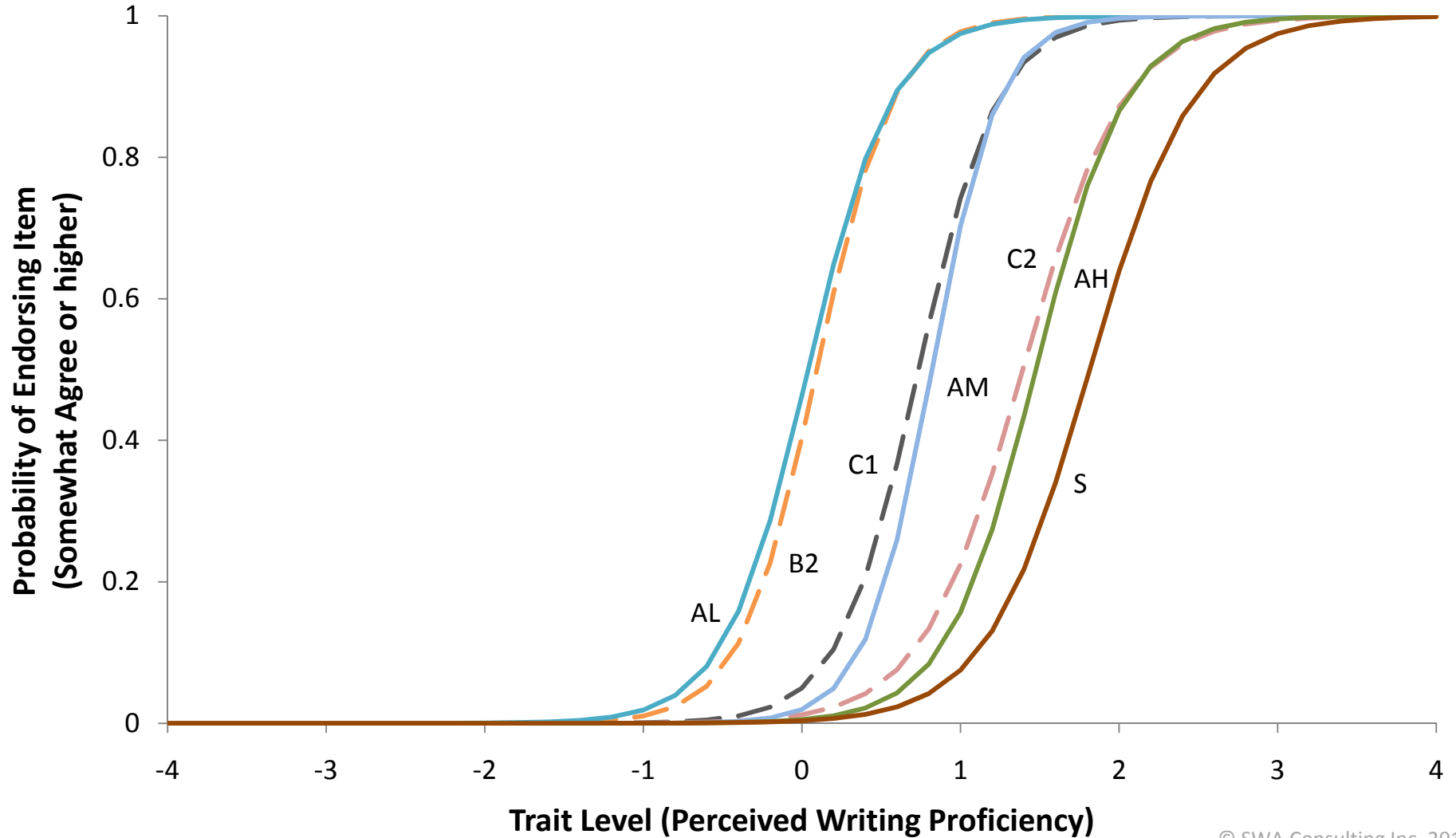
\* Endorse = "Agree" or higher

# Framework Comparisons



\* Endorse = "Agree" or higher

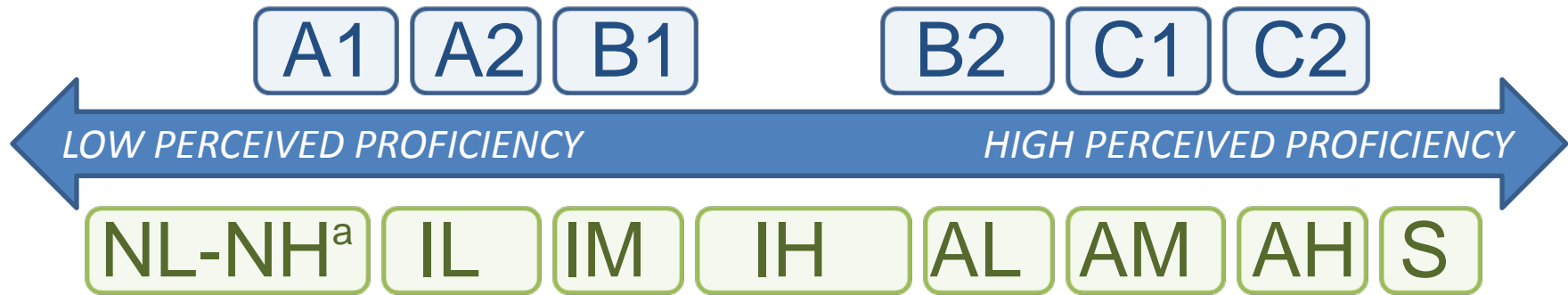
# Framework Comparisons



# Perceived Writing Proficiency Across Frameworks



## CEFR Writing



## ACTFL Writing

Note. Spacing is intentional.

<sup>a</sup> Novice items could not be modeled because all respondents endorsed them.



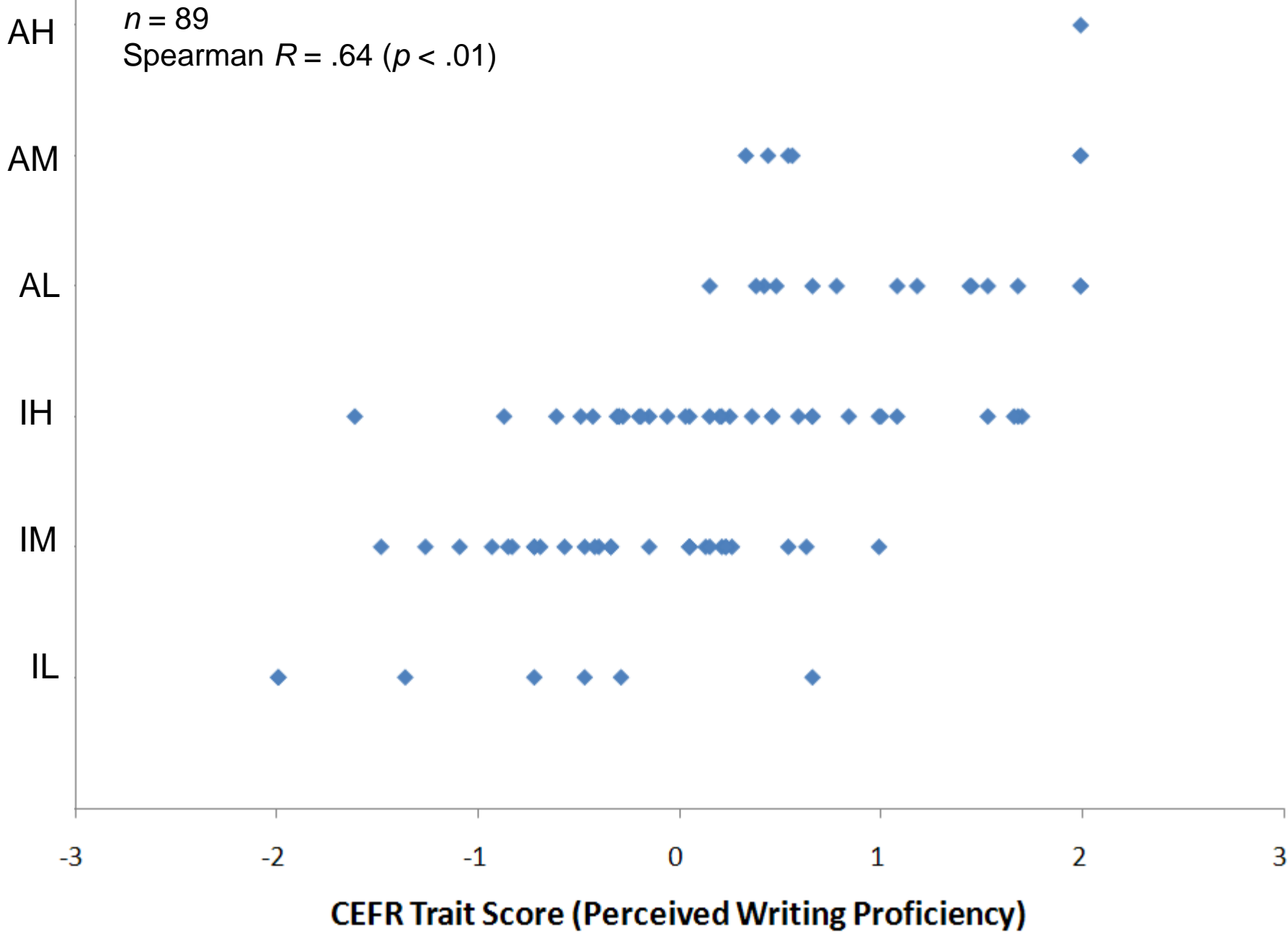
# Perceived CEFR Writing Proficiency vs. Actual ACTFL Writing

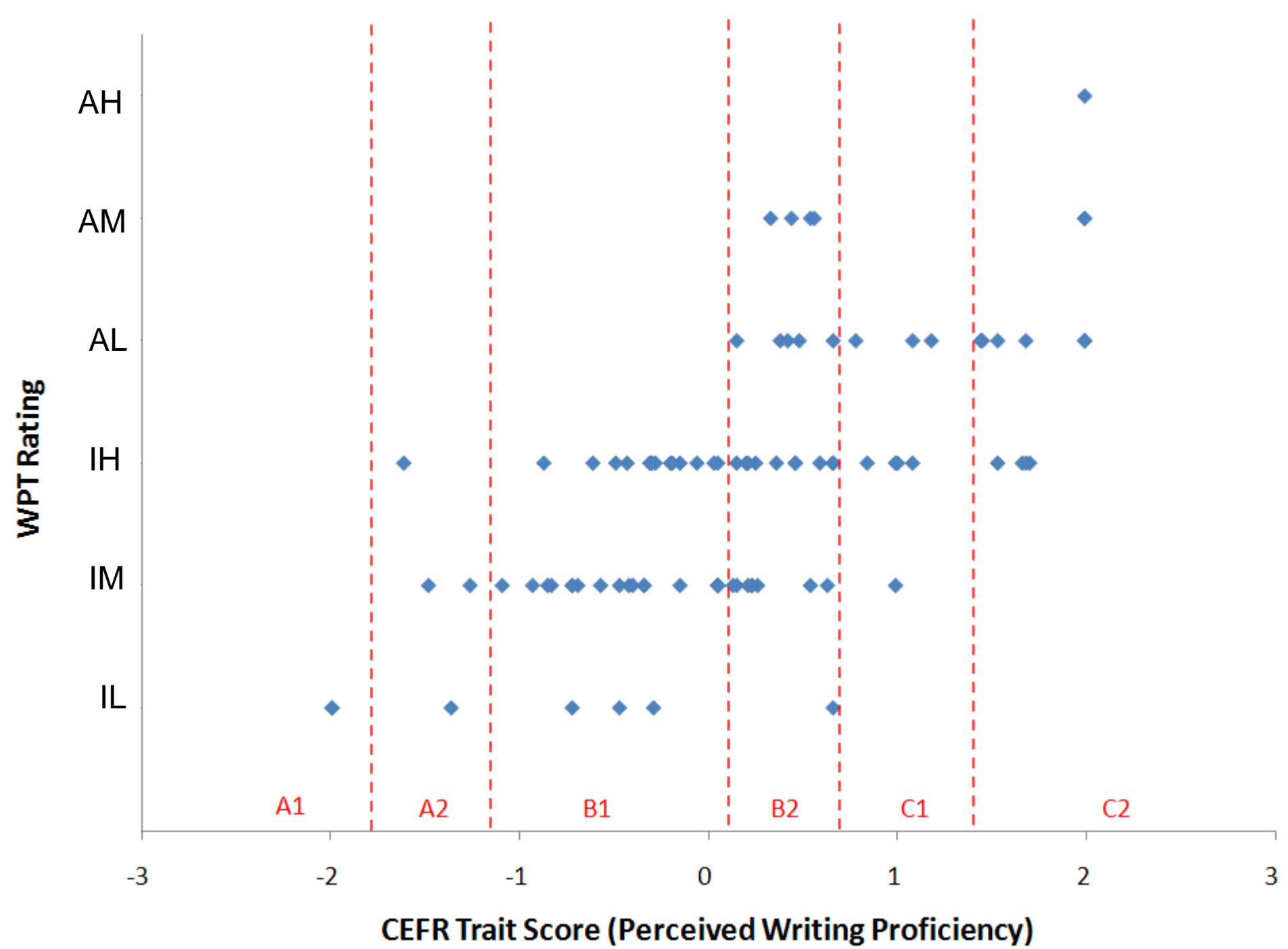


- What is the relationship between *perceived* CEFR writing proficiency (measured by ‘can-do’ ratings) and *actual* ACTFL writing proficiency (measured by WPT)?



- Placed CEFR items and respondents' perceived proficiency on same metric using item response modeling (IRT).
  - Results interpreted with caution given the small sample
- Looked at how actual writing proficiency on the WPT mapped onto the perceived capability construct.

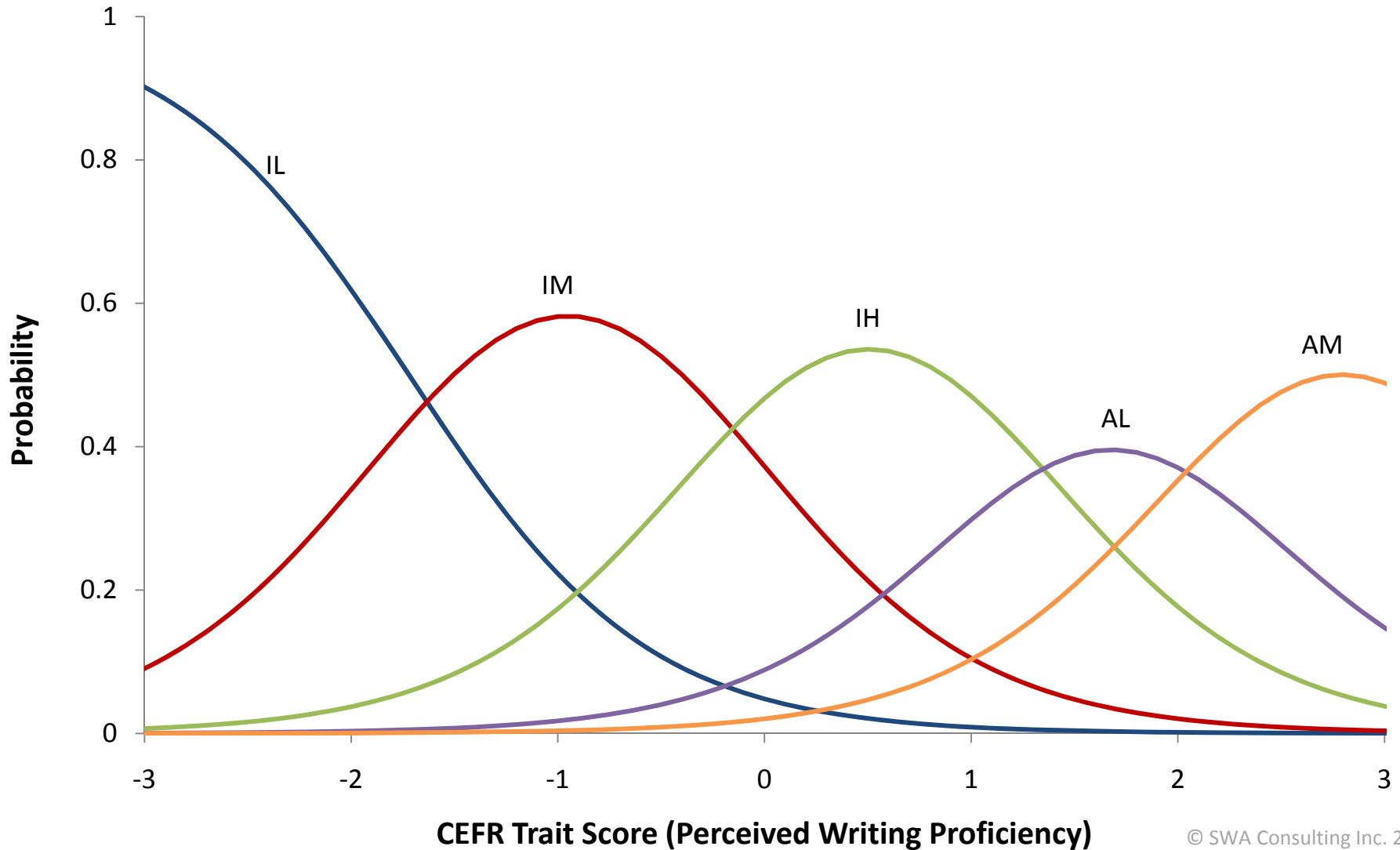


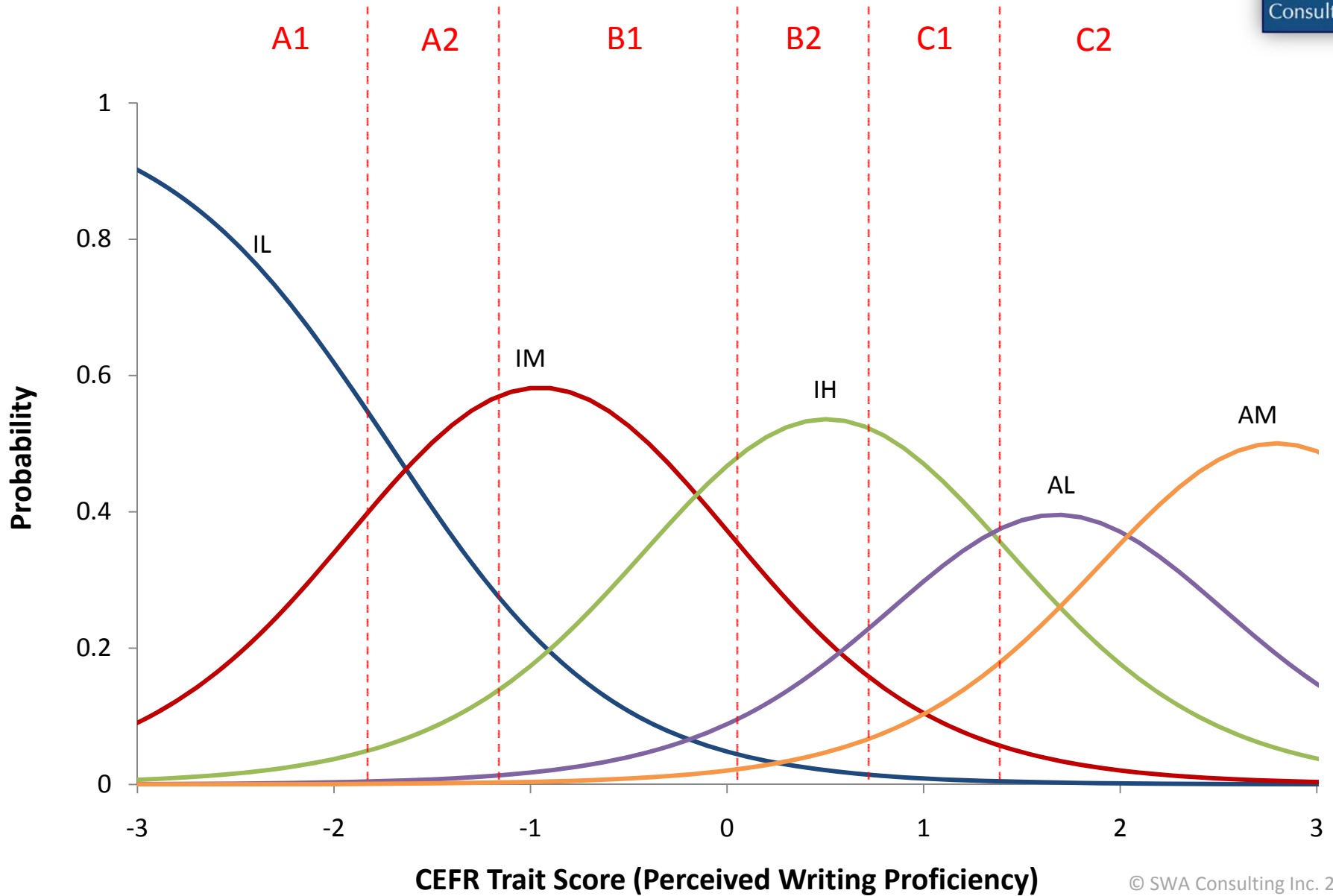


# Predicting WPT from CEFR Perceived Writing Proficiency



Universiteit Utrecht





# Predicting Actual Proficiency from Perceived Proficiency



## Estimated Probability of WPT Rating Based on Perceived CEFR Proficiency

		<u>Actual WPT Rating</u>				
		IL	IM	IH	AL	AM
Perceived Writing Proficiency (CEFR)	A1	.73	.24	.02	.00	.00
	A2	.36	.53	.10	.01	.00
	B1	.11	.53	.32	.04	.01
	B2	.02	.24	.53	.16	.04
	C1	.01	.09	.44	.32	.12
	C2	.00	.02	.15	.35	.38

*Shaded values are highest probability on the row.*

# Predicting Actual Proficiency from Perceived Proficiency



## Estimated Probability of WPT Rating Based on Perceived CEFR Proficiency

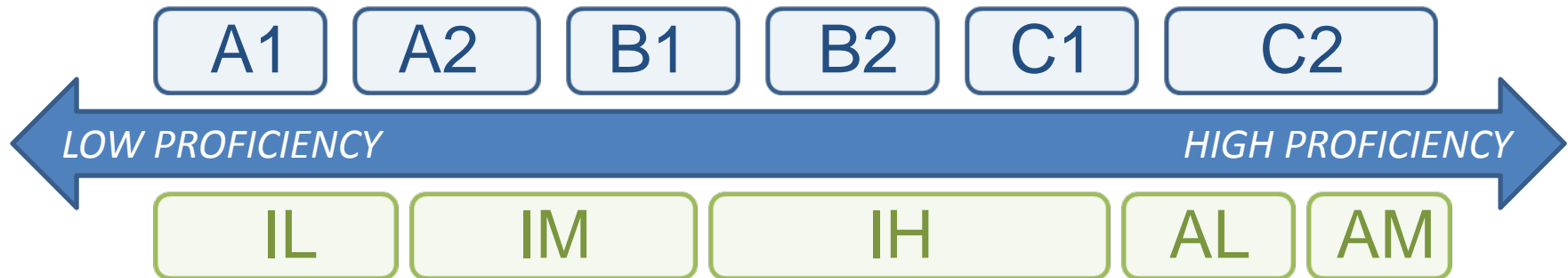
### Actual WPT Rating

		IL	IM	IH	AL	AM
Perceived Writing Proficiency (CEFR)	A1	0.73	0.24	0.02	0	0
	A2	0.36	0.53	0.1	0.01	0
	B1	0.11	0.53	0.32	0.04	0.01
	B2	0.02	0.24	0.53	0.16	0.04
	C1	0.01	0.09	0.44	0.32	0.12
	C2	0	0.02	0.15	0.35	0.38

# Perceived CEFR vs. Actual ACTFL Writing Proficiency



## Perceived CEFR Writing



## Actual WPT Rating



# **Perceived ACTFL Writing Proficiency vs. Actual ACTFL Writing Proficiency**

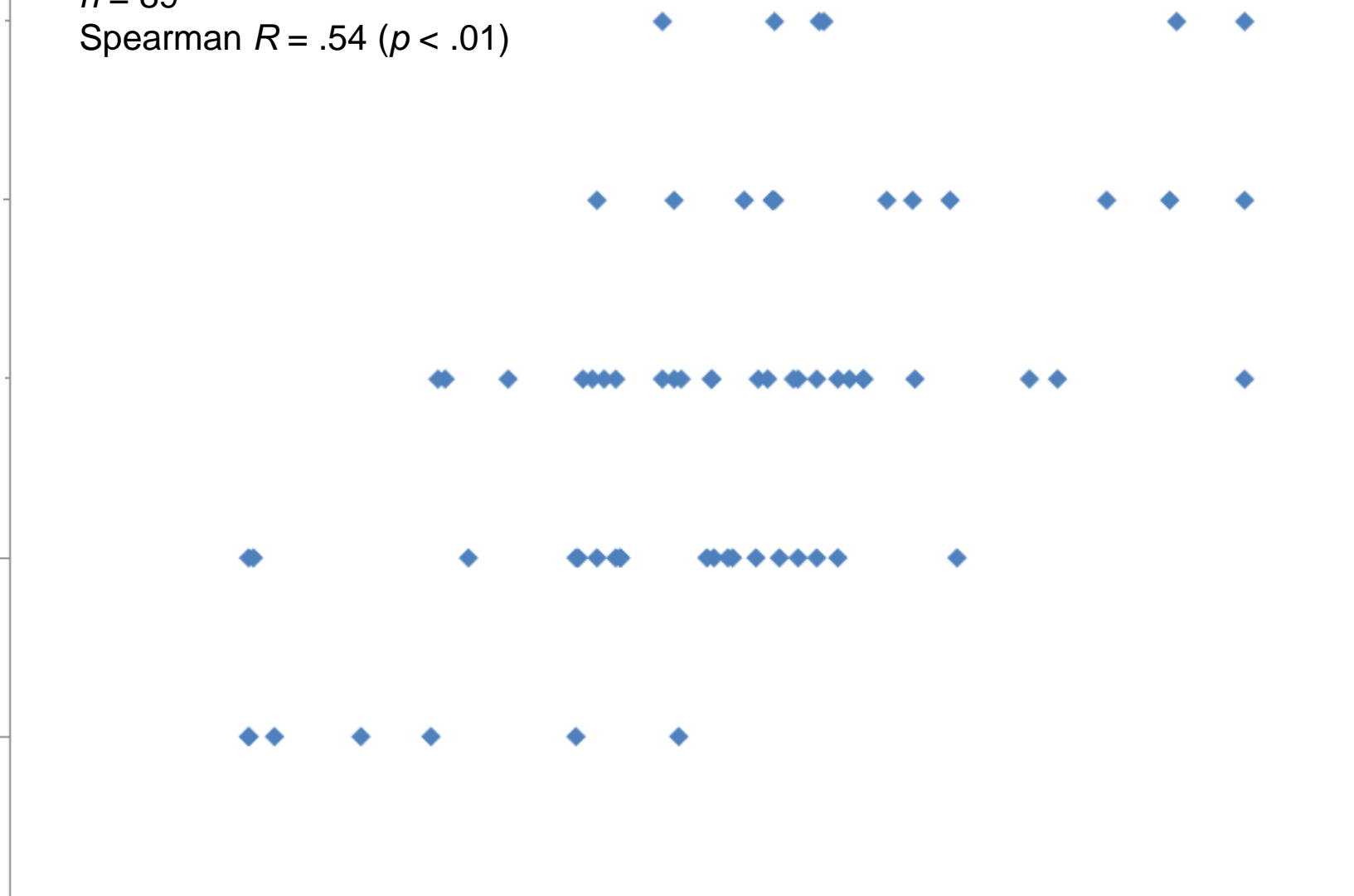
WPT Rating

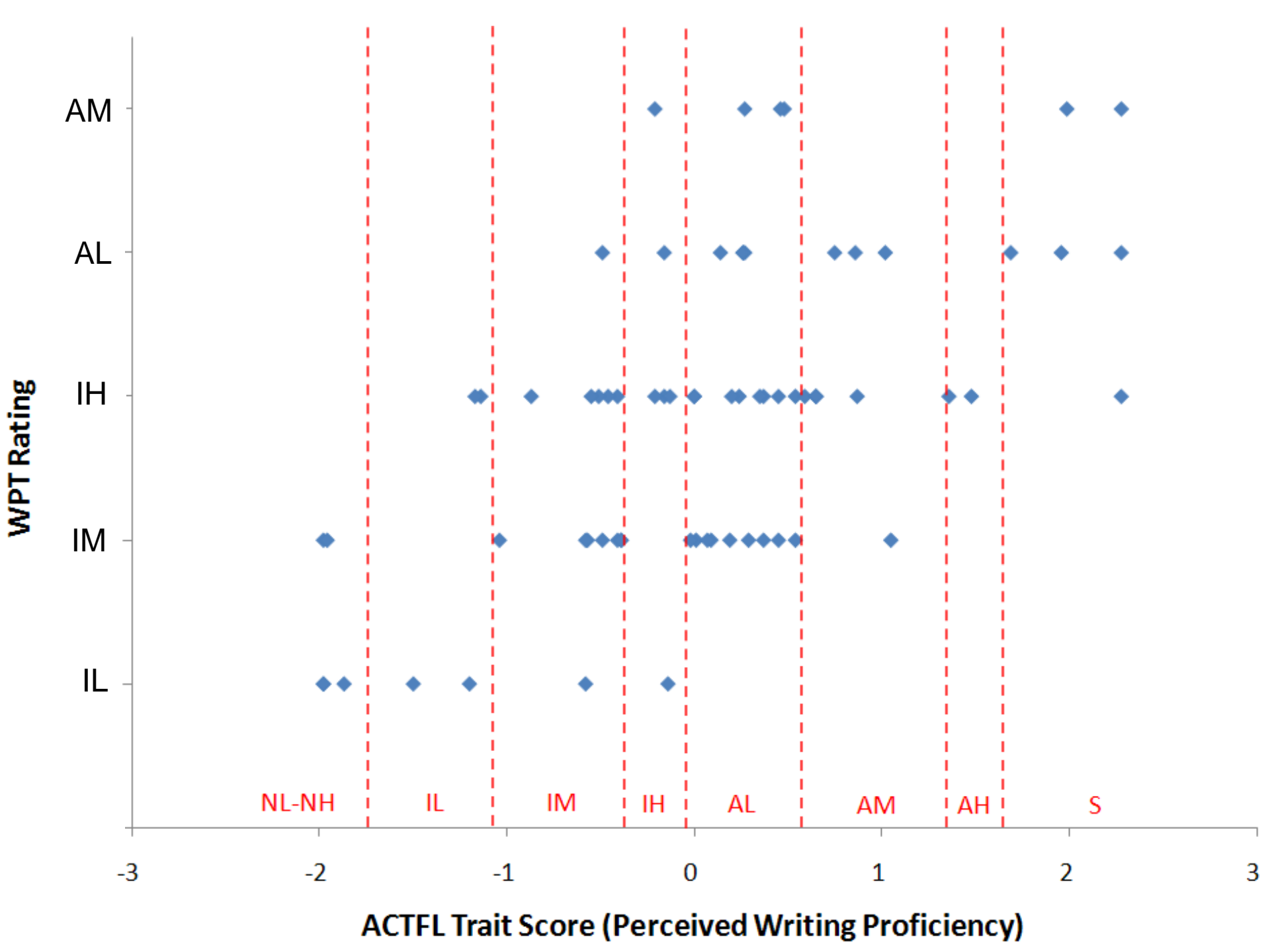
AM  
AL  
IH  
IM  
IL

$n = 89$   
Spearman  $R = .54$  ( $p < .01$ )

-3 -2 -1 0 1 2 3

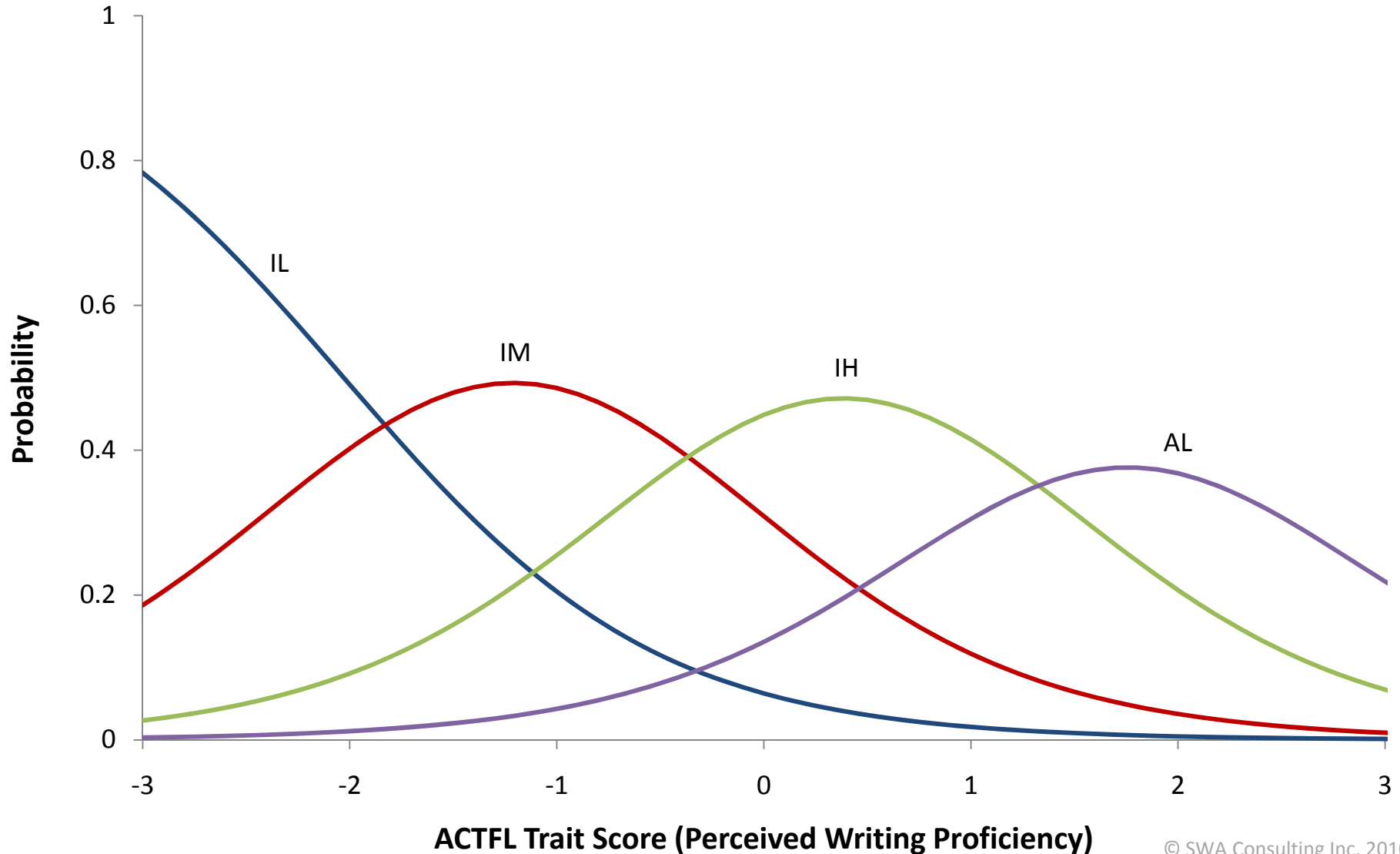
ACTFL Trait Score (Perceived Writing Proficiency)

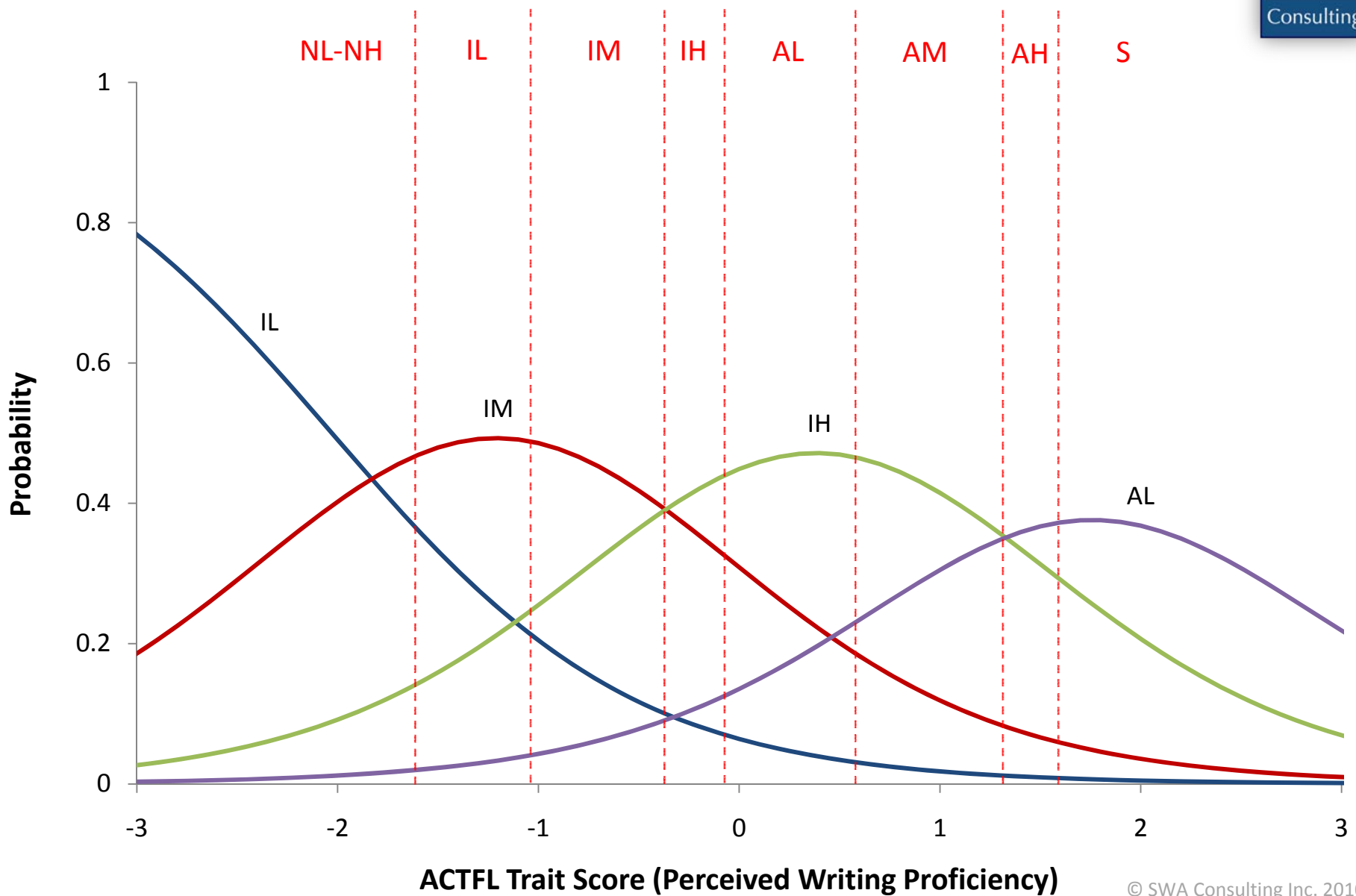




# Predicting of WPT from ACTFL

## Perceived Writing Proficiency





# Predicting Actual Proficiency from Perceived Proficiency



## Estimated Probability of WPT Rating Based on Perceived ACTFL Proficiency

### WPT Rating

		IL	IM	IH	AL
Perceived Writing Proficiency (ACTFL)	IL	<b>0.30</b>	0.49	0.18	0.03
	IM	0.15	<b>0.45</b>	0.32	0.06
	IH	0.08	0.35	<b>0.42</b>	0.11
	AL	0.04	0.24	0.47	<b>0.18</b>
	AM	0.02	0.12	0.42	0.31
	AH	0.01	0.07	0.31	0.37
	S	0	0.03	0.17	0.35

*Shaded values are highest probability on the row. Bold values are equivalent ACTFL rating levels.*

# Predicting Actual Proficiency from Perceived Proficiency



## Estimated Probability of WPT Rating Based on Perceived ACTFL Proficiency

		<u>WPT Rating</u>			
		IL	IM	IH	AL
Perceived Writing Proficiency (ACTFL)	IL	0.30	0.49	0.18	0.03
	IM	0.15	0.45	0.32	0.06
	IH	0.08	0.35	0.42	0.11
	AL	0.04	0.24	0.47	0.18
	AM	0.02	0.12	0.42	0.31
	AH	0.01	0.07	0.31	0.37
	S	0.00	0.03	0.17	0.35



- Collect more data
  - Larger sample size
  - More proficiency data
  - CEFR Proficiency test
  - Other languages
- Additions to data collection procedure
  - Add demographics
- We are looking for additional sites that wish to provide participants.



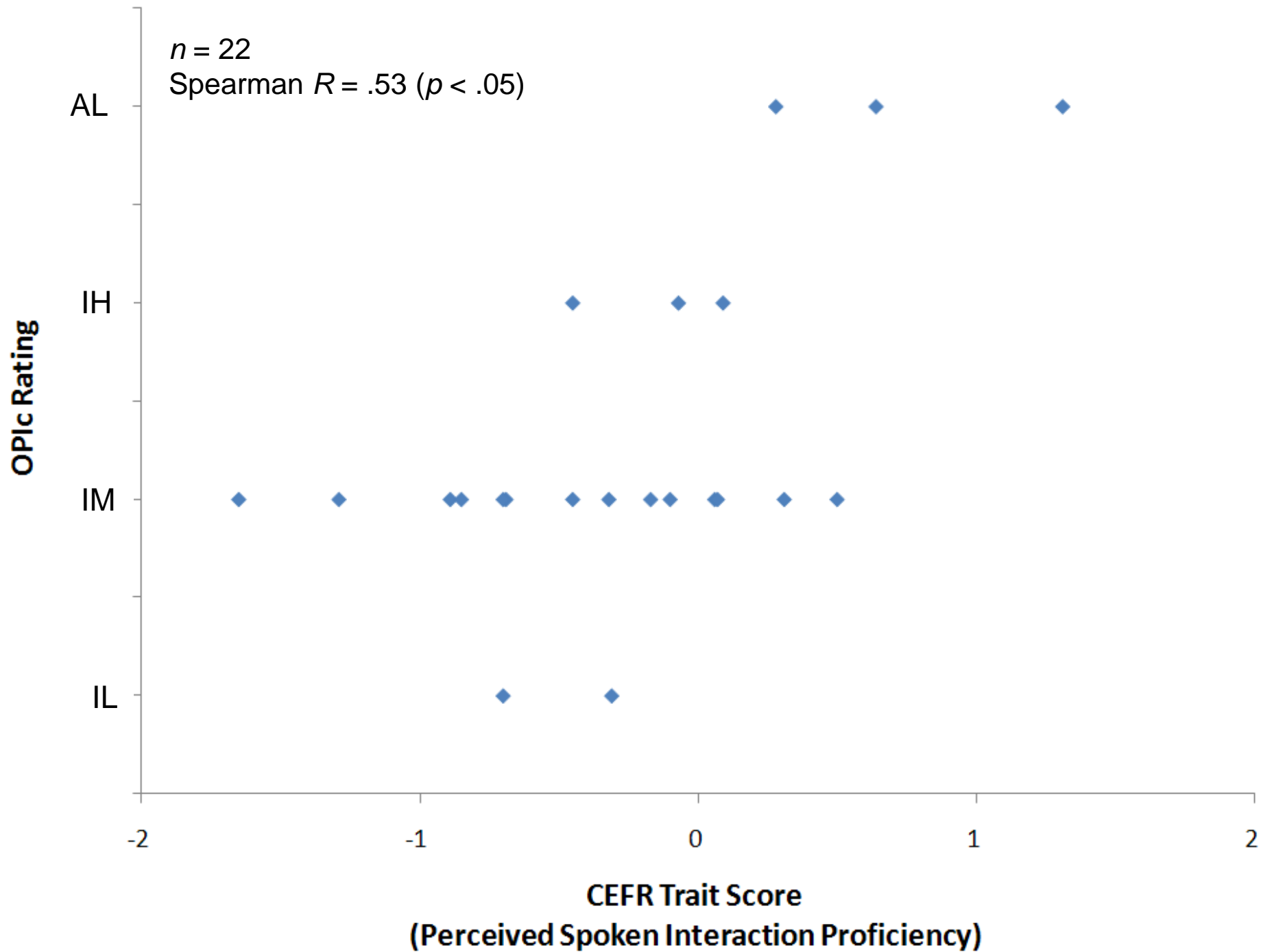
# Reserve Slides

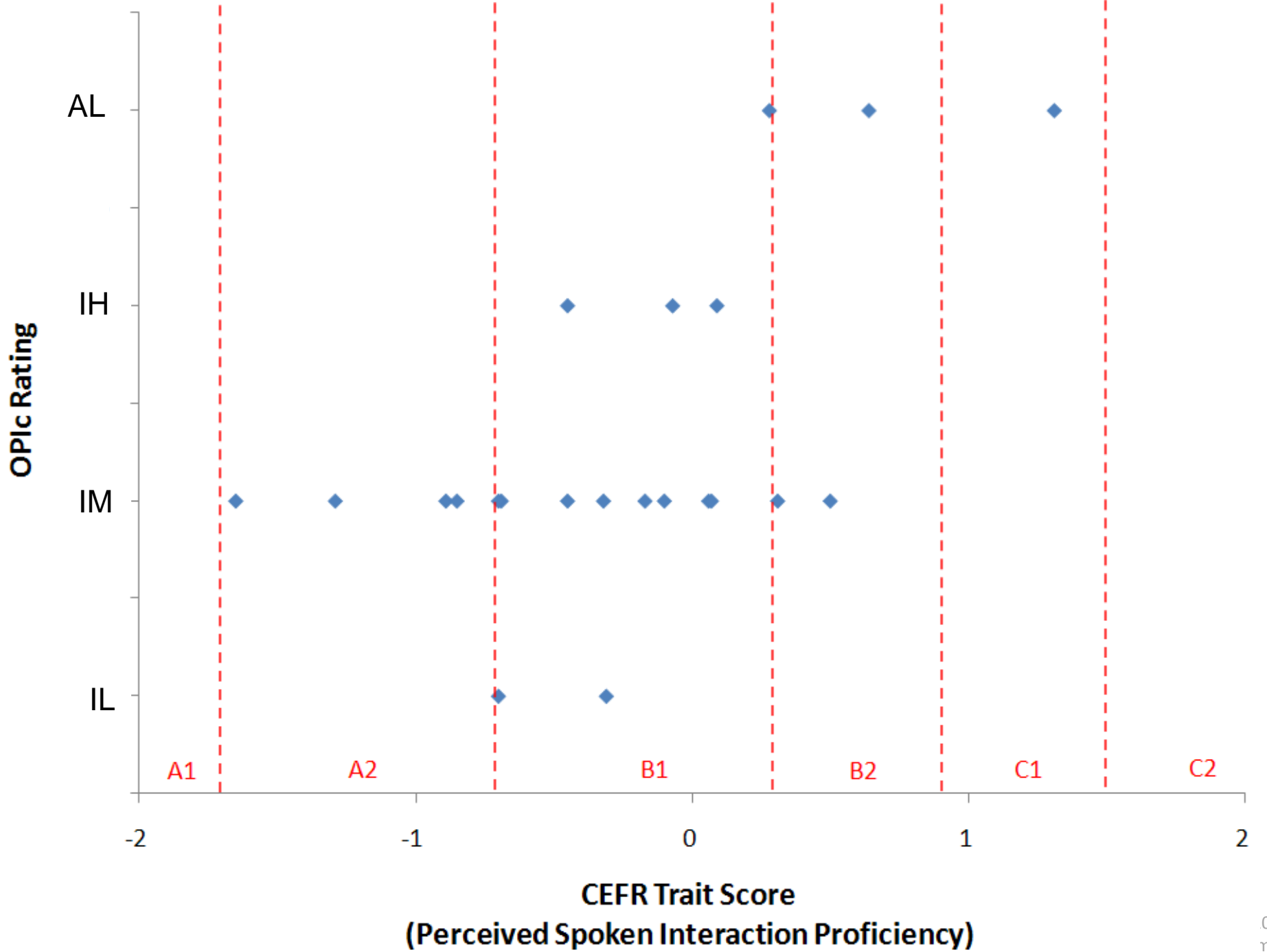


# RQ2: CEFR Spoken Interaction

## ↔ ACTFL OPIc

$n = 22$   
Spearman  $R = .53$  ( $p < .05$ )





# Perceived Spoken Interaction Proficiency and Actual OPIC Rating



## Actual OPIC Rating

		IL	IM	IH	A	Total
Perceived Spoken Interaction Proficiency (CEFR)	A1	-	-	-	-	0
		-	-	-	-	0%
	A2	-	4	-	-	4
		-	29%	-	-	18%
	B1	2	8	3	-	13
		100%	57%	100%	-	59%
	B2	-	2	-	2	4
		-	14%	-	67%	19%
	C1	-	-	-	1	1
		-	-	-	33%	5%
	C2	-	-	-	-	0
		-	-	-	-	0%
<i>Total</i>		2	14	3	3	22

% = percentages for column

## ABOUT SWA CONSULTING INC.

SWA Consulting Inc. (formerly Surface, Ward, and Associates) provides analytics and evidence-based solutions for clients using the principles and methods of industrial/organizational (I/O) psychology. Since 1997, SWA has advised and assisted corporate, non-profit and governmental clients on:

- Training and development
- Performance measurement and management
- Organizational effectiveness
- Test development and validation
- Program/training evaluation
- Work/job analysis
- Needs assessment
- Selection system design
- Study and analysis related to human capital issues
- Metric development and data collection
- Advanced data analysis

One specific practice area is analytics, research, and consulting on foreign language and culture in work contexts. In this area, SWA has conducted numerous projects, including language assessment validation and psychometric research; evaluations of language training, training tools, and job aids; language and culture focused needs assessments and job analysis; and advanced analysis of language research data.

Based in Raleigh, NC, and led by Drs. Eric A. Surface and Stephen J. Ward, SWA now employs close to twenty I/O professionals at the masters and PhD levels. SWA professionals are committed to providing clients the best data and analysis with which to make solid data-driven decisions. Taking a scientist-practitioner perspective, SWA professionals conduct model-based, evidence-driven research and consulting to provide the best answers and solutions to enhance our clients' mission and business objectives. SWA has competencies in measurement, data collection, analytics, data modeling, systematic reviews, validation, and evaluation.

For more information about SWA, our projects, and our capabilities, please visit our website ([www.swa-consulting.com](http://www.swa-consulting.com)) or contact Dr. Eric A. Surface ([esurface@swa-consulting.com](mailto:esurface@swa-consulting.com)) or Dr. Stephen J. Ward ([sward@swa-consulting.com](mailto:sward@swa-consulting.com)).