

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188		
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b>					
<b>1. REPORT DATE</b> (DD-MM-YYYY) March 2013		<b>2. REPORT TYPE</b> Briefing		<b>3. DATES COVERED</b> (From - To) Sep 2012 – Mar 2013	
<b>4. TITLE AND SUBTITLE</b> Preliminary Study on Using ILR Can-Do Statements for Placement in SOFTS Virtual Training Courses			<b>5a. CONTRACT NUMBER</b> H92222-10-D-0017 / 0007 J7 SOFLO		
			<b>5b. GRANT NUMBER</b>		
			<b>5c. PROGRAM ELEMENT NUMBER</b>		
<b>6. AUTHOR(S)</b>  SWA Consulting Inc.			<b>5d. PROJECT NUMBER</b>		
			<b>5e. TASK NUMBER</b>		
			<b>5f. WORK UNIT NUMBER</b>		
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>  SWA Consulting Inc. 311 S Harrington Street Suite 200 Raleigh, NC 27603 SWA was a subcontractor to CACI, Inc. on Contract # H92222-10-D-0017 / 0007 J7 SOFLO			<b>8. PERFORMING ORGANIZATION REPORT</b>  2013011202		
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>  Special Operations Forces Language Office HQ USSOCOM Attn: SOKL-J7—SOFLO 7701 Tampa Point Blvd MacDill AFB, FL 33621-5323			<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b> SOFLO		
			<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>		
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b>  A. Approved for public release; distribution is unlimited					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b> This presentation was delivered to the Interagency Language Roundtable (ILR) Testing Committee on 20 MAR 2013. SWA Consulting Inc. investigated the effectiveness of the ILR Can Do Statements as a Special Operations Forces Tele-training System (SOFTS) course placement tool. Researchers asked the following questions: 1. Are the Can Do Statements measuring perceived language speaking proficiency consistently and accurately for all students? and 2. Are the Can Do Statements related to similar constructs such as students' confidence in their ability to perform language tasks? The Can Do Statements measure perceived speaking proficiency (i.e., an individual's perception of his or her language ability). The sample consisted of 1,710 student responses to the Can Do Statements. Course feedback open ended comments were also analyzed. Although this study provides some support for the use of the ILR Can Do Statements as a placement tool for SOFTS courses, some limitations may restrict the usefulness of the findings. Most notably, this was an archival study. A follow-up study could involve measuring actual proficiency with an Oral Proficiency Interview (OPI) at the beginning of language training. This would allow for comparison between students' OPI ratings and their responses to the Can Do Statements.					
<b>15. SUBJECT TERMS</b> Special Operations Forces Tele-training System virtual training interagency language roundtable ILR can-do statements placement					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION</b> UU (SAR)	<b>18. NUMBER</b>  23	<b>19a. NAME OF RESPONSIBLE PERSON</b>
<b>a. REPORT</b> U	<b>b. ABSTRACT</b> U	<b>c. THIS PAGE</b>			<b>19b. TELEPHONE NUMBER</b> (include area code) 919-480-2751



Bienkowski, S. C. (2013, March). *Preliminary study on using ILR can-do statements for placement in SOFTS virtual training courses*. Presented to the Interagency Language Roundtable Testing Committee, National Foreign Language Center, College Park, MD.

## **Preliminary Study on using ILR Can-Do Statements for Placement in SOFTS Virtual Training Courses**



**MARCH 2013**

APPROVED FOR PUBLIC RELEASE;

DISTRIBUTION UNLIMITED

SPONSORED BY: SOFLO, USSOCOM

RESEARCH CONDUCTED BY: SWA CONSULTING INC.



SWA

Consulting

# Preliminary Study on Using ILR Can-Do Statements for Placement in SOFTS Virtual Training Courses

Presented by:

**Ms. Sarah C. Bienkowski**

Presented to:

**ILR Testing Committee**

Sponsored by:

**SOFLO, USSOCOM**

03.22.2013

Slide 1

SWA Briefing #2013011202

Prepared by:

**SWA Consulting Inc.**

© SWA Consulting Inc. 2013  
esurface@swa-consulting.com



- Special Operations Forces Teletraining System (SOFTS)
  - Synchronous online language training platform
  - Courses available in variety of languages and a range of proficiency levels
  - Course placement is determined in part by self-assessment measure



- RQ1: Are the ILR *Can Do Statements* measuring perceived language proficiency consistently and accurately for all SOFTS students?
  - Preliminary look at the ILR *Can Do Statements* functioning
- RQ2: Are the *Can Do Statements* related to similar constructs such as students' confidence in their ability to perform language tasks?
  - Preliminary look based on convergence with limited available data



- Sample
  - 709 student responses to the [Can Do Statements](#)
    - Students categorized as having no language proficiency were removed
  - Majority of participants enrolled in Spanish ( $n = 168$ ), French ( $n = 115$ ) or Modern Standard Arabic ( $n = 89$ ) at the ILR 0+ ( $n = 284$ ), 1 ( $n = 114$ ), or 2 ( $n = 84$ ) course level
- Analyses
  - Classical Test Theory (CTT) reliability analyses
  - Item response theory (IRT) analyses
  - Correlation with other perceived capability



- Internal consistency reliability (Cronbach's alpha)
  - ILR Level 1 = .88
  - ILR Level 2 = .90
  - ILR Level 3 = .87
  - ILR Level 4 = .82
- Item difficulty [Results Table](#)
  - *Can do* items at higher levels of the ILR scale were more difficult (endorsed less frequently) than at lower levels
    - However, a few items seemed to be too easy or too difficult for their specified ILR Level
- Item-total correlations
  - Most items had moderate to large item-total correlations
  - One Level 1 item had a small item-total correlation



- IRT analyses [Results Table](#)
  - Item difficulty estimate provides similar information as item difficulty estimate in CTT
  - Item discrimination estimate in IRT is similar to item-total correlation in CTT
  - **Results of the IRT analyses are consistent with the CTT findings**
- In general, findings suggest that the *Can Do Statements* subscales are consistently measuring the same construct; however, more evidence is needed to be confident that what the *Can Do Statements* are measuring is perceived language speaking ability

# RQ1 Results: Potential Item Issues



SWA

Consulting

## Can Do Statements that may be too easy for assigned ILR Level:

Level 2: *Can you take and give simple messages over the telephone or leave a message on voicemail?*

Level 4: *Can you take a discussion in different directions (friendly, controversial, collaborative)?*

## Can Do Statements that may be too difficult for their assigned ILR level:

Level 2: *Can you interview an employee, taking care of details such as salary, qualifications, hours and specific duties?*

Level 3: *Can you use the language to speculate at length about abstract topics such as how some change in history or the course of human events would have affected your life or civilization?*

Level 3: *Can you carry out any job assignment as effectively as you could in your native language?*

## Can Do Statements with low item-to-total correlations:

Level 1: *Are you often unable to finish a sentence because of grammatical or vocabulary limitations?*

# RQ2 Results: Convergent Validity



SWA

Consulting

- Strong correlation ( $r = .79$ ) between assigned course level with perceived speaking proficiency ratings on the pre-training survey (but after placement)
  - As course level increased, student perceived speaking capability increased
- Strong correlation between students' average *Can Do Statements* ratings and their average [Confidence ratings](#) on the pre-training survey ( $r = .77$ ,  $n = 147$ )
- **Summary:** *Can Do Statements* converged with other indicators of perceived capability on a different survey after class assignment—limited evidence of effective placement

# Results for RQ2



Table 5. Correlations for the Four *Can Do Statements* Subscales and Three *Confidence Factors*

	Can Do Level 1	Can Do Level 2	Can Do Level 3	Can Do Level 4	Basic Language Tasks	Daily Language Tasks	Military-Specific Language Tasks
Can Do Level 1	1	–	–	–	–	–	–
Can Do Level 2	.744 (709)	1	–	–	–	–	–
Can Do Level 3	.521 (709)	.753 (709)	1	–	–	–	–
Can Do Level 4	.389 (709)	.597 (709)	.808 (709)	1	–	–	–
Basic Language Tasks	.731 (147)	.655 (147)	.517 (147)	.406 (147)	1	–	–
Daily Language Tasks	.672 (147)	.709 (147)	.567 (147)	.512 (147)	.945 (316)	1	–
Military-Specific Language Tasks	.559 (147)	.705 (147)	.603 (147)	.595 (147)	.842 (316)	.934 (316)	1

All correlations were statistically significant.

# Future Directions



- Continue to monitor item and scale functioning and determine whether or not changes are needed
- Use actual measure of proficiency to investigate effectiveness of ILR *Can Do Statements*
- Develop and use other measures of placement effectiveness to validate ILR *Can Do Statements* use as a placement tool



SWA

Consulting

# SWA Consulting Inc.

311 S. Harrington Street  
Suite 200  
Raleigh, NC 27603

919.480.2751

<http://www.swa-consulting.com>



SWA

Consulting

# RESERVE SLIDES

# Can Do Statements



Question	Level
Can you explain or understand directions to a nearby hotel, restaurant, post office, or other establishment?	1
Can you order a meal?	1
Can you buy a needed item, such as bus or train ticket, groceries, or clothing?	1
Are you often unable to finish a sentence because of grammatical or vocabulary limitations?	1
Can you make social introductions and use greeting and leave-taking expressions?	1
Can you ask and answer simple questions about date and place of birth, nationality, marital status, and occupation?	1
Can you arrange for a hotel room or taxi ride?	1
Can you tell a story?	2
Can you take and give simple messages over the telephone or leave a message on voicemail?	2
Can you give detailed information about your job, your family, your house, and your community?	2
Can you report on news that you have seen recently on television or read?	2
Can you describe in detail a person or place that is very familiar to you?	2
Can you talk about an everyday event that happened in the recent past or that will happen soon?	2
Can you interview an employee, taking care of details such as salary, qualifications, hours, and specific duties?	2

# Can Do Statements (cont.)



## Question

## Level

Can you adjust your language to suit your audience, whether you're talking to diplomats, an O7, an E2, close friends, employees, or others?	3
Can you defend personal opinions about social and cultural topics?	3
Can you follow and contribute to a conversation among native speakers?	3
Can you cope with unexpected, difficult situations such as broken-down plumbing, an undeserved traffic ticket, or a serious social blunder?	3
Can you use the language to speculate at length about abstract topics such as how some change in history or the course of human events would have affected your life or civilization?	3
Can you discuss a hypothetical situation?	3
Can you carry out any job assignment as effectively as you could in your native language?	3
Can you prepare and give a lecture at a professional meeting about your area of specialization and debate complex aspects of it with others?	4
Can you take a discussion in different directions (friendly, controversial, collaborative)?	4
Can you persuade someone effectively to take a course of action in a sensitive situation, such as to improve their health, reverse a decision, or establish a policy?	4
Can you naturally integrate appropriate cultural and historical references into your speech?	4
Do you practically never make a grammatical mistake?	4
In professional discussions, is your vocabulary extensive and precise enough to enable you to convey your exact meaning?	4

[Return](#)

# Results Table



Table 1. CTT Item Difficulties and Item-Total Correlations for Subscales/ILR Levels 1 and 2

Level 1			Level 2		
Item	Diff	Item-Total	Item	Diff	Item-Total
Can you order a meal?	0.69	0.76	Can you take and give simple messages over the telephone or leave a message on voicemail?*	0.53	0.67
Can you buy a needed item, such as bus or train ticket, groceries, or clothing?	0.67	0.80	Can you give detailed information about your job, your family, your house, and your community?	0.40	0.77
Can you make social introductions and use greeting and leave-taking expressions?	0.66	0.69	Can you talk about an everyday event that happened in the recent past or that will happen soon?	0.38	0.77
Can you ask and answer simple questions about date and place of birth, nationality, marital status, and occupation?	0.62	0.72	Can you tell a story?	0.38	0.74
Can you explain or understand directions to a nearby hotel, restaurant, post office, or other establishment?	0.61	0.71	Can you describe in detail a person or place that is very familiar to you?	0.37	0.78
Can you arrange for a hotel room or taxi ride?	0.50	0.64	Can you report on news that you have seen recently on television or read?	0.30	0.76
Are you often unable to finish a sentence because of grammatical or vocabulary limitations?***	0.49	0.39	Can you interview an employee, taking care of details such as salary, qualifications, hours, and specific duties?*	0.14	0.53
Average	0.61	—	Average	0.36	—

n = 709

Diff = Item Difficulty: the percentage of students who endorsed the item. The larger the item difficulty, the easier the item.

Item Total = Item-Total Correlation: A measure of how related a given item is to the measure as a whole. Items with low item-total correlations do not discriminate well between individuals with different proficiency levels.

A single asterisk (\*) indicates items that may be too easy or too difficult for some individuals.

Two asterisks (\*\*) indicate an item does not discriminate well between individuals with different levels of language proficiency.

# Results Table (cont.)



Table 2. CTT Item Difficulties and Item-Total Correlations for Subscales/ILR Levels 3 and 4

Level 3			Level 4		
Item	Diff	Item-Total	Item	Diff	Item-Total
Can you follow and contribute to a conversation among native speakers?	0.25	0.66	Can you take a discussion in different directions (friendly, controversial, collaborative)?*	0.18	0.63
Can you adjust your language to suit your audience, whether you're talking to diplomats, an O7, an E2, close friends, employees, or others?	0.22	0.63	Can you persuade someone effectively to take a course of action in a sensitive situation, such as to improve their health, reverse a decision, or establish a policy?	0.11	0.63
Can you discuss a hypothetical situation?	0.20	0.67	Can you naturally integrate appropriate cultural and historical references into your speech?	0.10	0.68
Can you defend personal opinions about social and cultural topics?	0.19	0.77	Can you prepare and give a lecture at a professional meeting about your area of specialization and debate complex aspects of it with others?	0.06	0.57
Can you cope with unexpected, difficult situations such as broken-down plumbing, an undeserved traffic ticket, or a serious social blunder?	0.15	0.68	In professional discussions, is your vocabulary extensive and precise enough to enable you to convey your exact meaning?	0.05	0.62
Can you use the language to speculate at length about abstract topics such as how some change in history or the course of human events would have affected your life or civilization?*	0.09	0.63	Do you practically never make a grammatical mistake?	0.03	0.47
Can you carry out any job assignment as effectively as you could in your native language?*	0.06	0.48			
Average	0.17	—	Average	0.09	—

n = 709

Diff = Item Difficulty: the percentage of students who endorsed the item. The larger the item difficulty, the easier the item.

Item Total = Item-Total Correlation: A measure of how related a given item is to the measure as a whole. Items with low item-total correlations do not discriminate well between individuals with different proficiency levels.

A single asterisk (\*) indicates items that may be too easy or too difficult for their assigned ILR Level.

Two asterisks (\*\*) indicate an item does not discriminate well between individuals with different levels of language proficiency.

[Return](#)

# RQ1 Result: IRT



Table 3. IRT Item Difficulties and Item Discriminations for Subscales/ILR Levels 1 and 2

Level 1			Level 2		
Item	Diff	Disc	Item	Diff	Disc
Can you order a meal?	-0.48	2.15	Can you take and give simple messages over the telephone or leave a message on voicemail?*	0.00	2.46
Can you make social introductions and use greeting and leave-taking expressions?	-0.46	1.37	Can you give detailed information about your job, your family, your house, and your community?	0.33	2.86
Can you buy a needed item, such as bus or train ticket, groceries, or clothing?	-0.37	2.98	Can you talk about an everyday event that happened in the recent past or that will happen soon?	0.38	3.13
Can you ask and answer simple questions about date and place of birth, nationality, marital status, and occupation?	-0.30	1.67	Can you tell a story?	0.39	2.66
Can you explain or understand directions to a nearby hotel, restaurant, post office, or other establishment?	-0.21	2.57	Can you describe in detail a person or place that is very familiar to you?	0.40	3.17
Can you arrange for a hotel room or taxi ride?	0.05	1.26	Can you report on news that you have seen recently on television or read?	0.56	3.12
Are you often unable to finish a sentence because of grammatical or vocabulary limitations? **	0.07	0.30	Can you interview an employee, taking care of details such as salary, qualifications, hours, and specific duties?*	1.17	1.99
Average	-0.25	—	Average	0.46	—

Two asterisks (\*\*) indicate an item does not discriminate well between individuals with different levels of language proficiency.

n = 709

*Diff* = Item Difficulty: Measured in standard deviations around the mean. Items of average difficulty are equal to zero, positive values (+) are more difficult than average, and negative values (-) are easier than average.

*Disc* = Discrimination: High discrimination values indicate that the item discriminates well between individuals with different speaking proficiency levels. Negative values or values close to zero indicate that the item does not differentiate well between individuals with different proficiency levels. A single asterisk (\*) indicates items that may be too easy or too difficult for their assigned ILR Level.

# RQ1 Result: IRT



Table 4. IRT Item Difficulties and Item Discriminations for Subscales/ILR Levels 3 and 4

Level 3			Level 4		
Item	Diff	Disc	Item	Diff	Disc
Can you follow and contribute to a conversation among native speakers?	0.74	2.05	Can you take a discussion in different directions (friendly, controversial, collaborative)?*	0.96	2.20
Can you discuss a hypothetical situation?	0.85	2.68	Can you persuade someone effectively to take a course of action in a sensitive situation, such as to improve their health, reverse a decision, or establish a policy?	1.21	3.13
Can you defend personal opinions about social and cultural topics?	0.89	2.93	Can you naturally integrate appropriate cultural and historical references into your speech?	1.32	2.13
Can you adjust your language to suit your audience, whether you're talking to diplomats, an O7, an E2, close friends, employees, or others?	0.90	1.72	Can you prepare and give a lecture at a professional meeting about your area of specialization and debate complex aspects of it with others?	1.71	1.95
Can you cope with unexpected, difficult situations such as broken-down plumbing, an undeserved traffic ticket, or a serious social blunder?	1.06	2.41	In professional discussions, is your vocabulary extensive and precise enough to enable you to convey your exact meaning?	1.74	2.33
Can you use the language to speculate at length about abstract topics such as how some change in history or the course of human events would have affected your life or civilization?*	1.35	2.68	Do you practically never make a grammatical mistake?	2.26	1.44
Can you carry out any job assignment as effectively as you could in your native language?*	1.78	1.78			
Average	1.08	—	Average	1.53	—

n = 709

**Diff** = Item Difficulty: Measured in standard deviations around the mean. Items of average difficulty are equal to zero, positive values (+) are more difficult than average, and negative values (-) are easier than average.

**Disc** = Discrimination: High discrimination values indicate that the item discriminates well between individuals with different proficiency levels. Negative values or values close to zero indicate that the item does not differentiate well between individuals with different proficiency levels.

A single asterisk (\*) indicates items that may be too easy or too difficult for their assigned ILR Level.

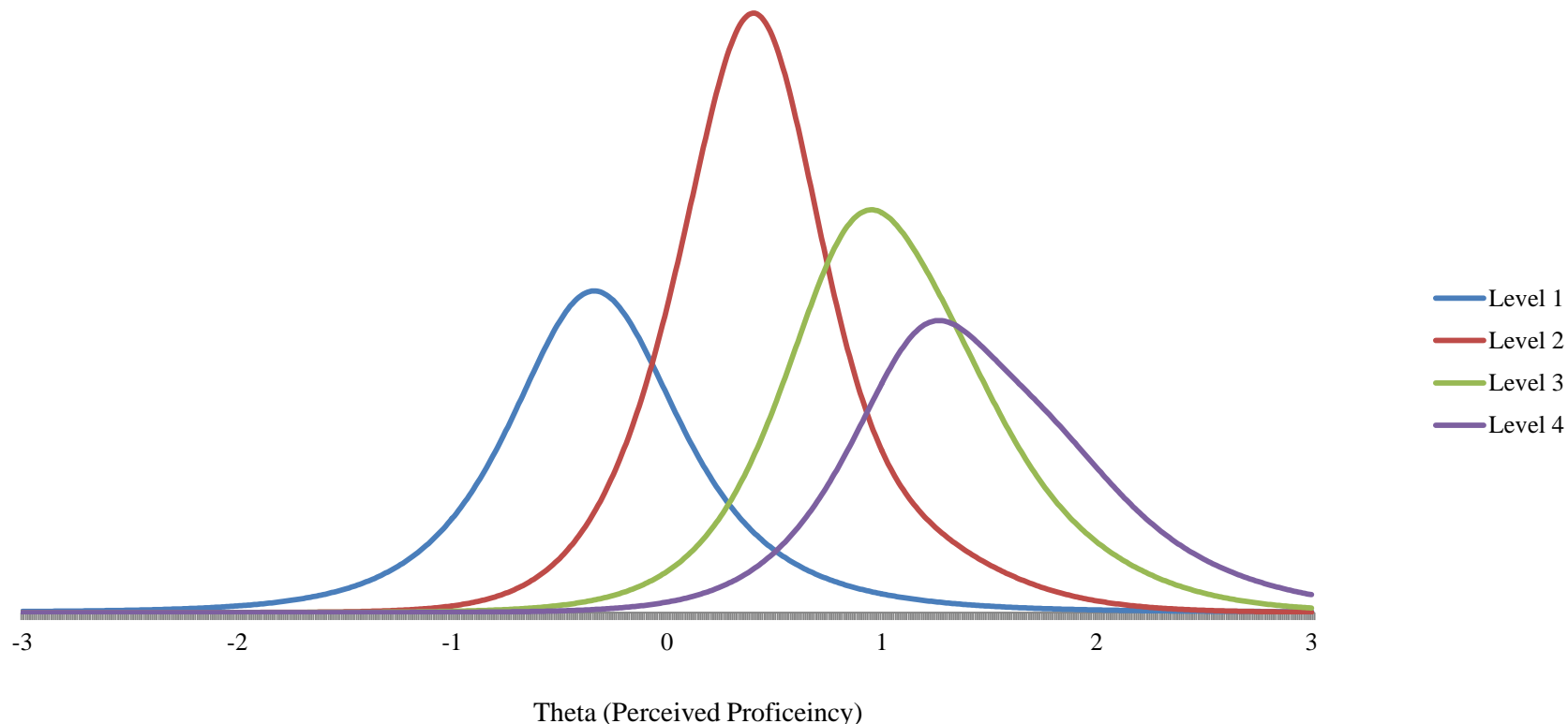
Two asterisks (\*\*) indicate an item does not discriminate well between individuals with different levels of language proficiency.

[Return](#)

# RQ1 Result: IRT



Figure 1. Test Characteristic Curves for *Can Do Statements* Subscales/ILR Levels 1 through 4

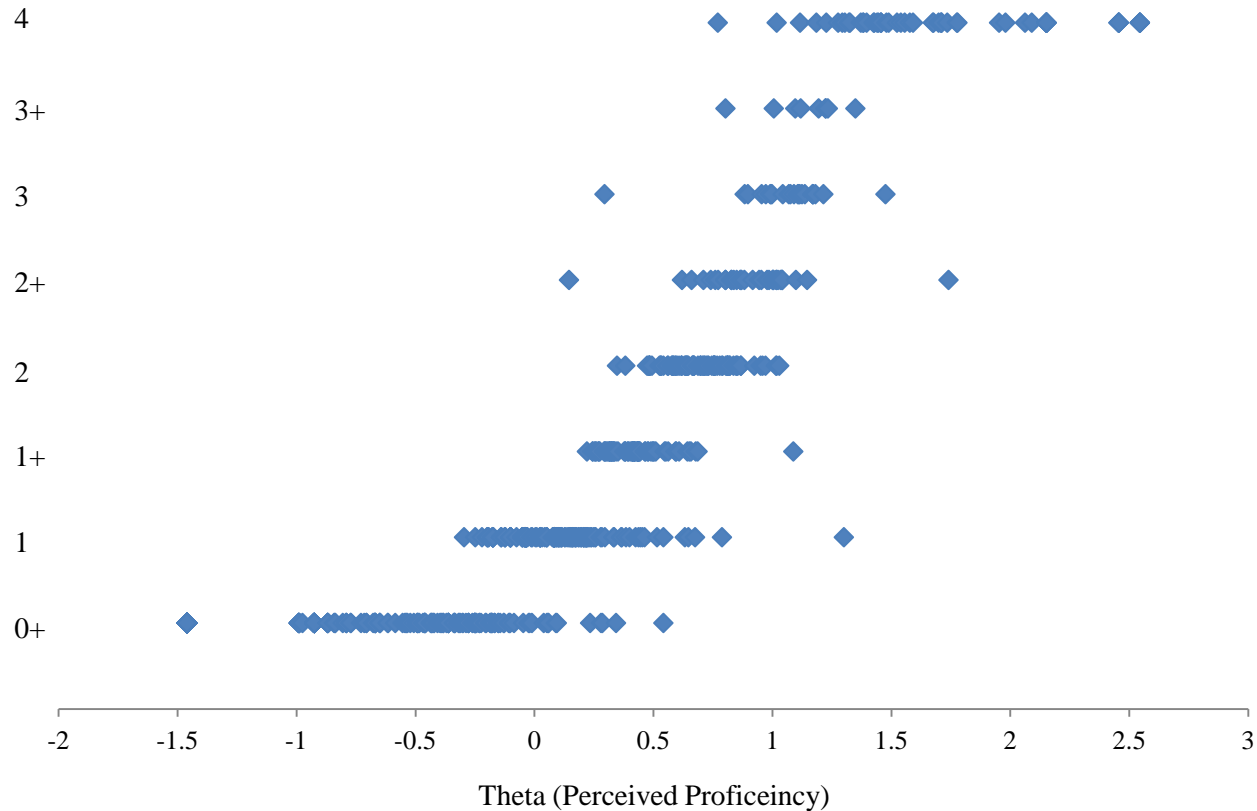


*Theta* (Perceived Proficiency) is measured in standard deviation units around the mean. Average perceived proficiency is equal to zero, positive (+) theta values indicate perceived proficiency levels that are higher than the mean, and negative (-) theta values indicate perceived proficiency levels that are below the mean.

# RQ1 Result: IRT



Figure 2. Comparison of Students' Perceived Speaking Proficiency Ratings and Assigned Course Level



*Theta* (Perceived Proficiency) is measured in standard deviation units around the mean. Average perceived proficiency is equal to zero, positive (+) theta values indicate perceived proficiency levels that are higher than the mean, and negative (-) theta values indicate perceived proficiency levels that are below the mean.

[Return](#)



- Basic Tasks
  - Exchange personal information (e.g., age, marital status, etc.).
  - Use various levels of formality when interacting with individuals.
  - Exchange social courtesies (e.g., extend or accept an invitation).
  - Express numerical ideas (e.g., count to 20).
  - Communicate information about time.
  - Communicate with others about my daily routine.
- Daily Tasks
  - Communicate with others about military and civilian occupations in the target region.
  - Engage in a telephone conversation.
  - Arrange for accommodations in the host country.
  - Make purchases in the host country.
  - Conduct daily business transactions (e.g., exchange currency).

# Confidence Items (cont.)



- Military Tasks
  - Acquire supplies for my mission.
  - Use military-technical vocabulary.
  - Exchange mission-related information with a counterpart in the host country.
  - Describe features of the environment.
  - Conduct negotiations in the training language.
  - Use this language to train or teach others.

[Return](#)