

Empirical Validation of Reading Proficiency Scales

Results 2012

Ray Clifford & Dana Bourgerie

Prior Research Results: Reading

- Students were able to capture some meaning from all levels of texts, so there is no sequence of text difficulty as implied by the ACTFL [and by ILR] scales. [Allen et al, 1988]
- Text types do not form a hierarchy of increasing difficulty. [Lee and Musimeci, 1998:]

Prior Results: Reading

- There is no empirical evidence to validate the *a priori* [ILR and ACTFL] definitions of levels. [Alderson, 2000.]
- The ACTFL and ILR scales are frequently described as criterion-referenced, however there is a circularity in the descriptors between texts and readers ability. [Brown, 2002]

A Real-World Example

- The following post card was delivered to me a few years ago.
- Are all of the text types used equally difficult?

A Postcard Offer

PHONE-A-GRAM



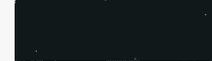
URGENT-U-REPLY

CONGRATULATIONS:

You will definitely receive one of the free prizes listed below:

- | | |
|------------------------------|-----------------------------|
| (1) PORTABLE CASSETTE PLAYER | (6) \$500 CASH |
| (2) COLOR TV | (7) 6 PIECE STEAK KNIFE SET |
| (3) 35MM CAMERA | (8) BLENDER |
| (4) GIFT CERTIFICATE | (9) INSTANT LOAD CAMERA |
| (5) FISHING ROD AND REEL | (10) 10 SPEED BICYCLE |

Coded Gift Number



Void if Tampered

★ **VALID THIS WEEK ONLY!!** ★

To Claim
Your Gift

OFFER EXPIRES _____

Call:

ASK FOR UNITED EXT. _____



(408)

4435

11 AM-8 PM

Call Immediately!

Urgent That You Call At Once!

This special offer is made available to you by your local Independent Distributor for United Laboratories of America, Inc. — Now celebrating over two decades of reliable service.

★ **GIFT AT NO COST OR OBLIGATION** ★

Although the purpose of this offer is to acquaint you with our film processing service, you are not required to make a purchase to receive your gift. If married, both husband and wife must be present. We ask only that you present your Phone-A-Gram in person so that we can erase the magic ink and determine your prize.

SEE REVERSE SIDE FOR ADDITIONAL INFORMATION

D.L.R. NO. _____

A Fine Print Definition of "NO COST"

United Laboratories Of America, Inc.
P. O. Box 4499 • Station A
Dallas, Texas 75208

PHONE-A-GRAM



URGENT-U-REPLY

Note: The address was printed here.

FIRST CLASS MAIL

(See Reverse Side for Details)

This program is not valid for United Distributors, relatives, friends or acquaintances of United Distributors, previous recipients, or in any state where prohibited by law. This phone-a-gram is not transferable and identification will be required. Your coded gift number will be uncovered at the time of your visit and you will be awarded the appropriate prize determined by your number. Styles and brands are subject to availability. We reserve the right to substitute merchandise of equal or greater value. This is a gift program to promote the sale of our photo processing services. Any tampering with the winning numbers will void this offer. A minimum of three million seven hundred thirty-six thousand dollars in prizes are included in every 467,000 mailers. All prizes will be distributed except those that are unclaimed. Void if used in an unauthorized manner.

Gift value and odds are as follows: (1) 60.00 — 1/4,670 (2) 300.00 — 1/46,700 (3) 79.00 — 1/934 (4) 8.00 — 464,378/467,000 (5) 69.96 — 1/934 (6) 500.00 — 1/233,500 (7) 8.00 — 1/934 (8) 49.00 — 1/934 (9) 19.00 — 1/934 (10) 100.00 — 1/46,700.

Should a legal adjudication be made that valid consideration was transferred or expended by offeree pursuant to the specific terms of this written offer then in such event restitutory indemnification shall be accomplished hereunder upon written request. Such request must be executed at time of offeree's visit. However, any such indemnification shall not exceed reasonable and ordinary amounts expended hereunder by the offeree.

PRINTED IN USA 8

A Fine Print Definition of “NO COST”

United Laboratories
P. O. Box 4499 • S
Dallas, Texas 7520

PHONE-

FIRST CLASS M

Should a legal adjudication be made that valid consideration was transferred or expended by offeree pursuant to the specific terms of this written offer, then in such event restitutory indemnification shall be accomplished hereunder upon written request.

This program is not valid for United States residents, relatives, employees, agents, or previous recipients, or in any state where prohibited by law. This phone-a-gram is not transferable and identification will be required at the time of your visit and you will be awarded the appropriate prize determined by your number. Styles and brands are subject to availability. Prizes will be awarded in cash or merchandise of equal value. This is a gift program to promote the sale of our photo processing services. Any tampering with the winning number will void the prize. A minimum of three million seven hundred thirty-six thousand dollars in prizes are included in every 467,000 mailers. All prizes will be distributed except those which are voided. Void if used in an unauthorized manner.

Gift value and odds are as follows: (1) 60.00 — 1/4,670 (2) 300.00 — 1/46,700 (3) 600.00 — 1/934 (4) 8.00 — 464,378/467,000 (5) 69.96 — 1/934 (6) 500.00 — 1/233,500 (7) 8.00 — 1/934 (8) 49.00 — 1/934 (9) 19.00 — 1/934 (10) 100.00 — 1/46,700.

Should a legal adjudication be made that valid consideration was transferred or expended by offeree pursuant to the specific terms of this written offer then in such event restitutory indemnification shall be accomplished hereunder upon written request. Such request must be executed at time of offeree's visit. However, any such indemnification shall not exceed reasonable and ordinary amounts expended hereunder by the offeree.

What Researchers Haven't Understood

- Text types are not defined by what a reader can do, but by the task or purpose the author is trying to accomplish.
- Language proficiency (and what the proficiency scales describe) is a multi-dimensional trait.
- The best way to measure the development of multidimensional traits is to define progress using steps or a hierarchy of levels within the multidimensional continuum.

C-R Tests of Reading Proficiency

1. Define the construct to be tested.
2. Use the proficiency scales to describe Task, Conditions, and Accuracy (TCA) criteria for each level to be tested.
3. Train the items writers to develop C-R, TCA-aligned items that target a specific level.
4. Test whether the C-R, TCA-aligned item sets cluster in a difficulty hierarchy.

1. The Construct to be Tested

Proficient reading: The active, automatic, far-transfer process of using one's internalized language and culture expectancy system to efficiently comprehend an authentic text for the purpose for which it was written.

Author purpose

Orient

Inform

Evaluate

–

–

–

Reading purpose

Get necessary information

Learn

Evaluate and synthesize

2. Level-specific, TCA-aligned Criteria

- When an educated writer is writing for an educated audience...
 - The author's purpose normally determines the genre or text mode that is used.
 - The text mode normally determines the text type that is used for the intended audience.
- When testing for proficient reading, ...
 - The reader task matches the writer's purpose.
 - The question asked or test item used must match the reader's task.

The Effect of Test Item Choices



3. Trained Item Writers

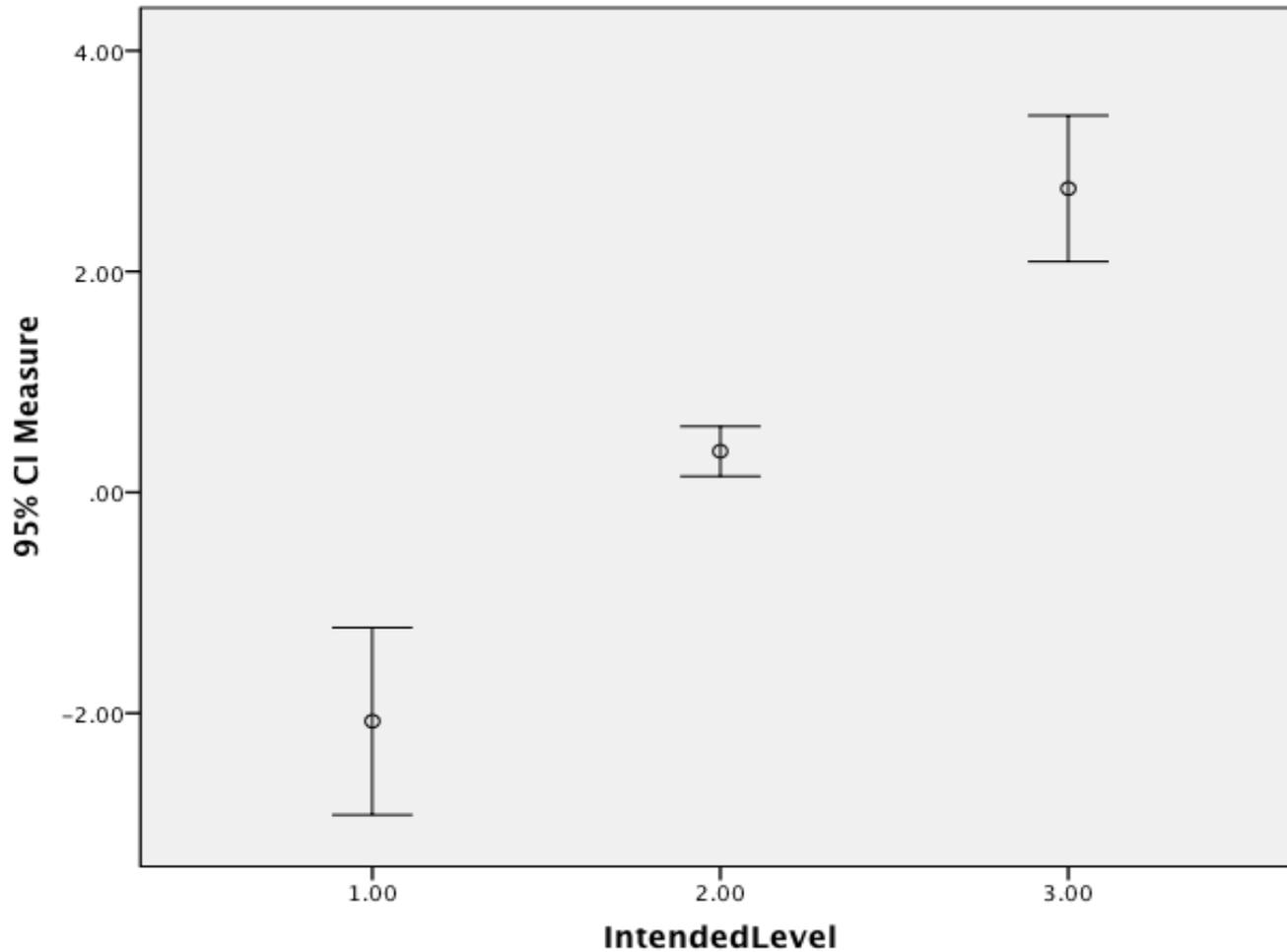
- Trained the items writers on the construct.
- Trained item writers on the proficiency scale.
- Trained item writers on external text modes, internal text types, and their relationship to author purpose.
- Trained item writers to align the reader's task with the author's purpose.
- Retrained the item writers.
- Mentored the item writers.

4. Confirming that the C-R, TCA-aligned items do cluster in a difficulty hierarchy.

- Checked for alignment before piloting items using a Modified-Modified-Angoff procedure.
 - Estimated at-level difficulty.
 - Estimated difficulty for learners at 1 level lower.
 - Estimated difficulty for learners at 1 level higher.
 - Estimated item discrimination across levels.
- Confirmed alignment and discrimination with WINSTEPS IRT analyses.

Reading Simplified Chinese

Initial Means and Confidence Intervals



With this documented alignment,

- We can produce Reading tests that:
 - Are criterion-referenced.
 - Can be computer adaptive.
 - Include the advantages of testlet-based design.
 - Use “floor and ceiling” scoring.
- Generate reports that show:
 - The reader’s sustained level of proficiency.
 - Progress the reader has made toward the next higher proficiency level.

Thank you.

Questions?