Item and Scale Differential Functioning in the Mini Mental Status Examination (MMSE)

Leo S. Morales, M.D., Ph.D.

Assistant Professor

UCLA School of Medicine

Assessing Measurement Equivalence Across Populations: Differential Item Functioning (DIF)

June 24, 2004

Support: RWJ Medical Minority Development Program, UCLA/DREW Project Export (P20-MD00148-01) and UCLA Resource Center for Minority Aging Research (AG-02-004)

Collaborators

Claudia Flowers, Ph.D., University of North Carolina

Peter Gutierrez, M.A., UCLA School of Medicine

Marjorie Kleinman, Ph.D., Columbia University

Jeanne Teresi, Ph.D., Columbia University

The Mini Mental Status Examination

Cognitive impairment affects 10% population over 65 and 50% of population over 85.

MMSE is the most familiar and widely used cognitive screening test worldwide.

Selected characteristics of MMSE:

- **Physician administered**
- **Seventeen (17) items**
- **Takes 5 to 10 minutes to administer.**
- **Multiple translations in existence including Spanish, Chinese, Danish, French German, Russian, …**

Purpose

To evaluate the Spanish version of the MMSE for *differential functioning* at the item and scale level (relative to the English version).

Data Source

Northern Manhattan Aging Project

- **13 adjacent census tracks in Northern Manhattan**
- **9,349 individuals 65 years of age and older**

Dementia Registry

- **Survey of representative sub-samples**
- **Reporting network**

Analytical file of 1,578 completed surveys

- **913 English-language respondents**
- **665 Spanish-language respondents**

MMSE Instrument

17 items assessing five areas of cognitive functioning including:

- **Orientation**
- **Attention**
- **Memory**
- **Recall**
- **Language**

MMSE Spanish Translation

Blind independent forward translations

Reconciliation of new translations and with previous translations

Pilot test and cognitive interviews (n=50)

MMSE Items 1-11

Item	Content	Categories
MMSE1	State year	2
MMSE2	State season	2
MMSE3	State day of month	2
MMSE4	State day of week	2
MMSE5	Name month	2
MMSE6	Name state	2
MMSE7	Name city	2
MMSE8	Name 2 nearby streets	2
MMSE9	Name floor of building	2
MMSE10	Name of type of place	2
MMSE11	Name 3 objects	4

MMSE Items 12-21

Item	Content	Categories
MMSE12	Serial 7s	6
MMSE13	Spell WORLD	6
MMSE14	Recall 3 objects	4
MMSE15	Name pencil	2
MMSE16	Name wristwatch	2
MMSE17	Repeat phrase	2
MMSE18	Close eyes	2
MMSE19	Instructions w/paper	4
MMSE20	Write sentence	2
MMSE21	Copy design	2

Sample Characteristics (n=1,578)

Gender	**English (n=913)**	**Spanish (n=665)**	**P-Value**
Female	75	76	0.82
Age			
<75 years	28	44	<0.001
75 – 84 years	42	39	
85+ years	31	17	
Race/Ethnicity			<0.001
Latino	5	98	
African American	62	<1	
White	32	1	
Asian	6	1	
Education			<0.001
0 – 8 years	46	82	
9+ years	54	18	

Statistical Analysis

Test IRT Model Assumptions

- **Dimensionality**
- **Local Independence**

Evaluate Item and Scale Differential Functioning

- **Estimate item parameters**
- **Evaluate item and scale differential functioning**

Testing Model Assumptions

Evaluate dimensionality and local independence separately for English and Spanish samples:

- **Item-scale correlation**
- **Internal reliability consistency (Alpha)**
- **Confirmatory factor analysis**

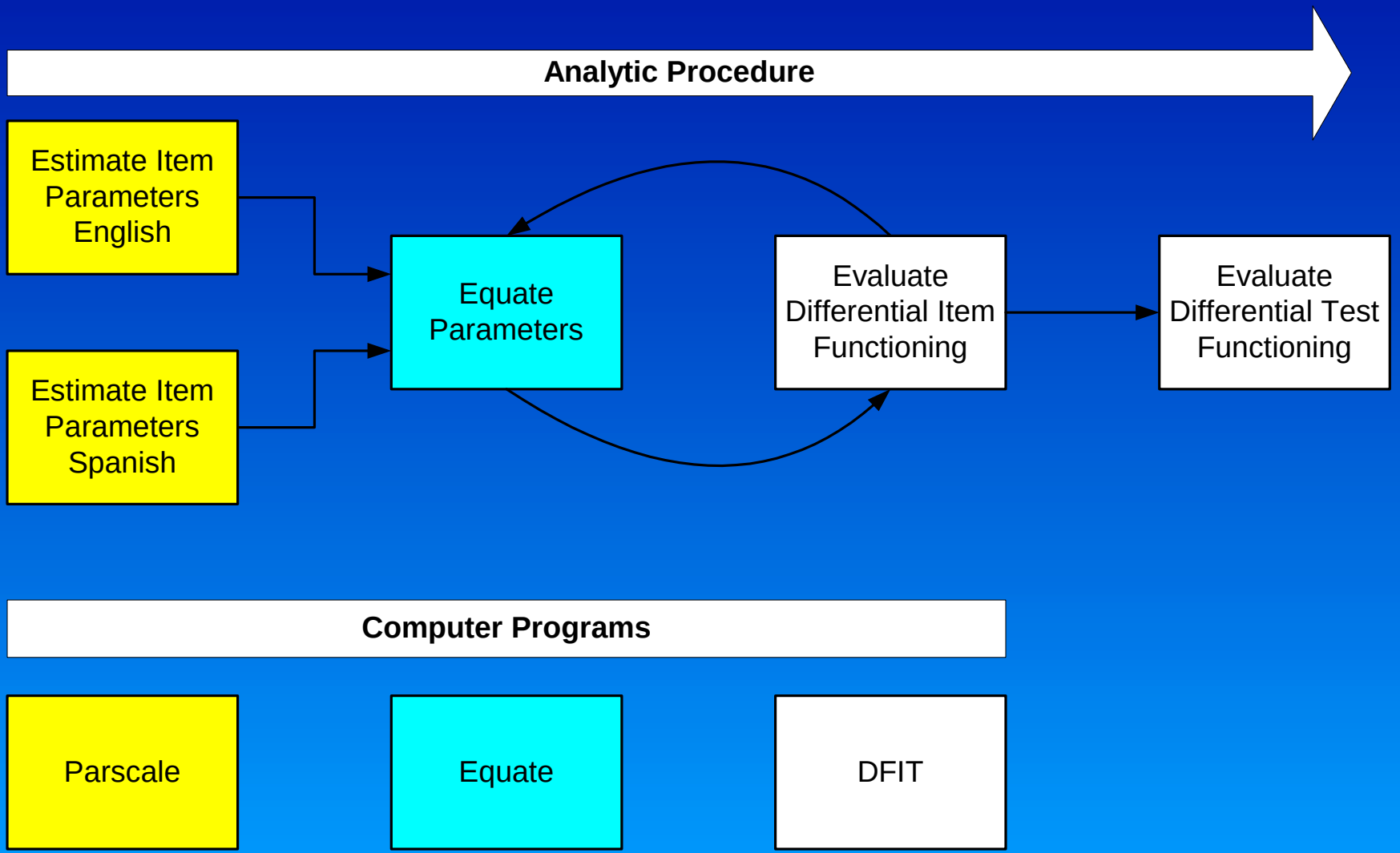
Item-Scale Correlation and Alpha

MMSE	**English**	**Spanish**
Item-Test Correlations		
Mean	0.63	0.58
Median	0.67	0.62
Range	0.48-0.72	0.43-0.68
Alpha	0.89	0.87

Results from Confirmatory Factor Analysis 1-Factor Solution

MMSE	Comparative Fit Index Rec >0.90	Tucker-Lewis Index Rec >0.90	Root Mean Square Error of Approximation Rec: <0.05
English	0.97	0.99	0.06
Spanish	0.97	0.99	0.06

Evaluating Differential Item and Test Functioning



Testing for Differential Functioning

Item-level differential functioning:

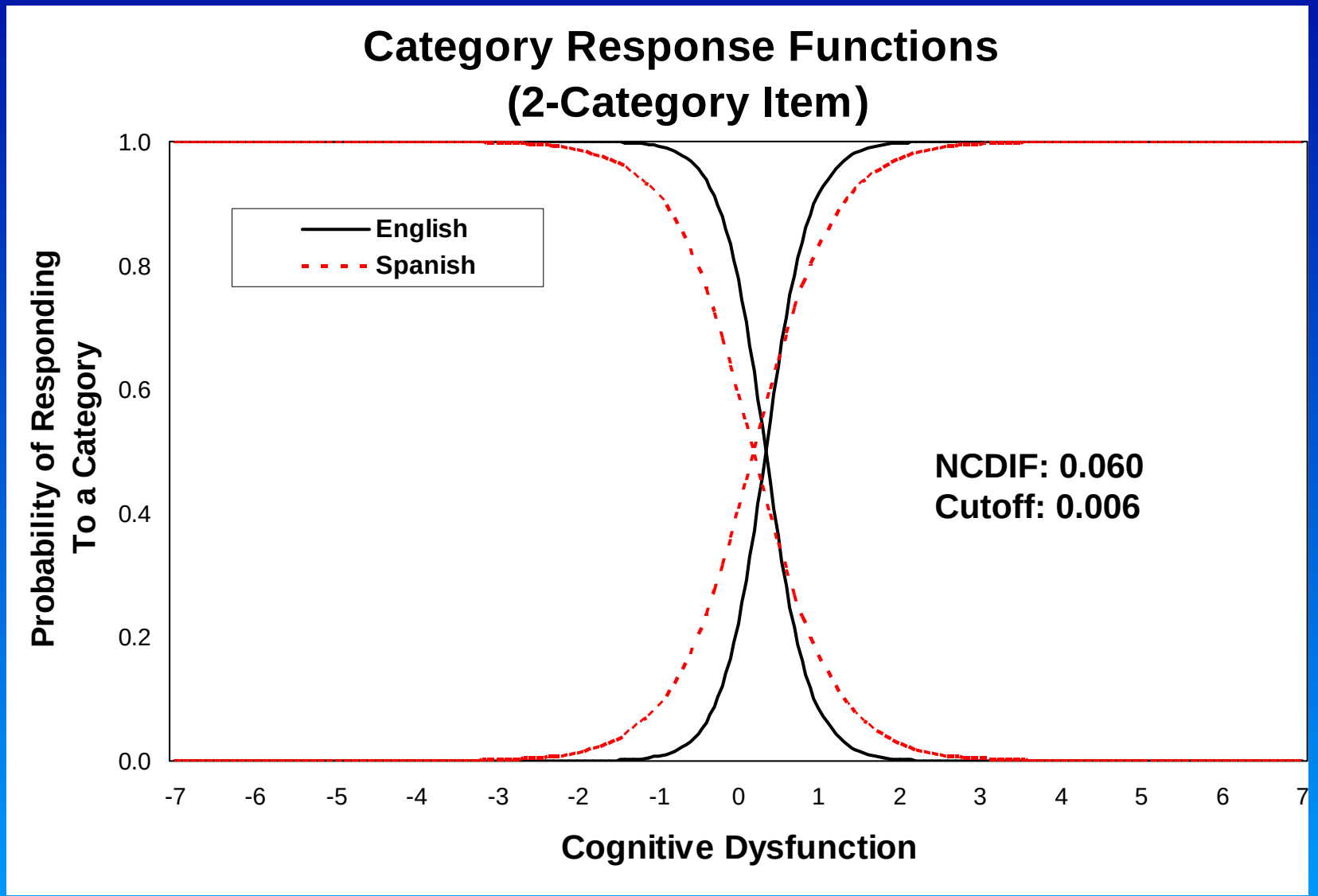
- **Chi-square statistics**
- **Non-compensatory DIF (NC-DIF)**

Scale-level differential functioning:

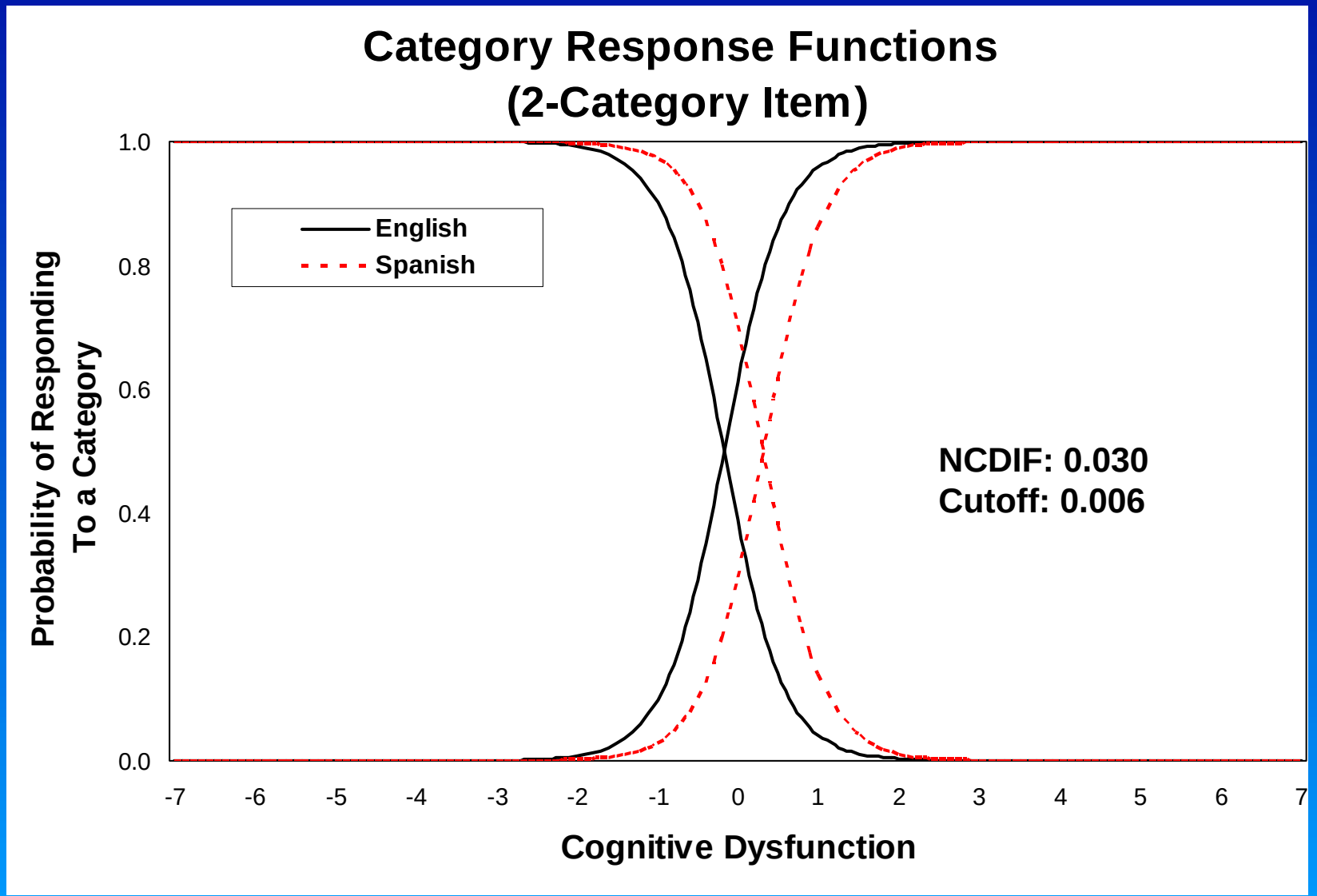
- **Chi-square statistics**
- **Compensatory DIF (C-DIF)**
- **Differential test functioning (DTF)**

Raju, van der Linden, Fleer (1995) IRT-Based Internal Measures of Differential Functioning of Item and Tests. Applied Psych Measurement 19(4): 353-368.

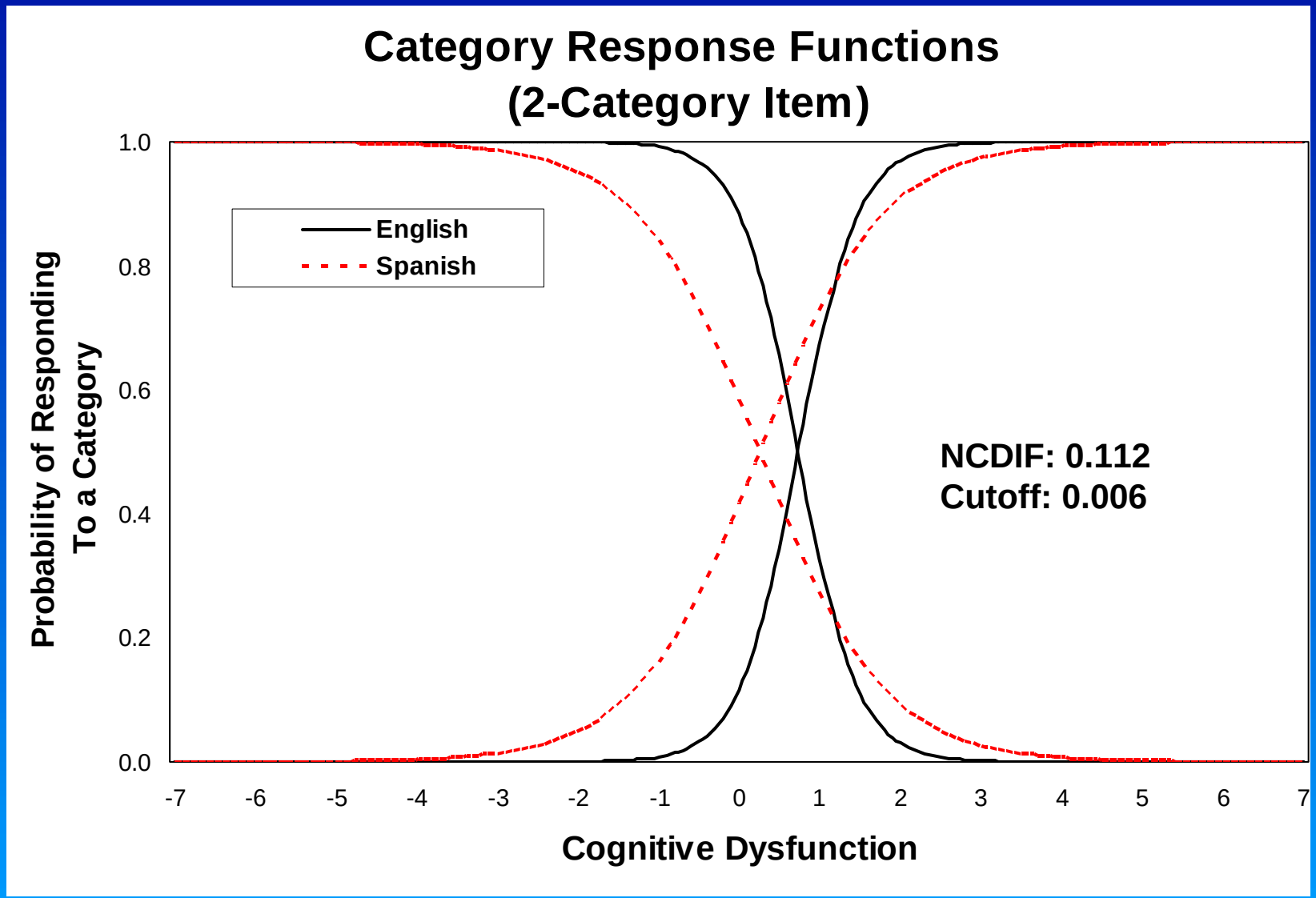
MMSE 2: Name Season



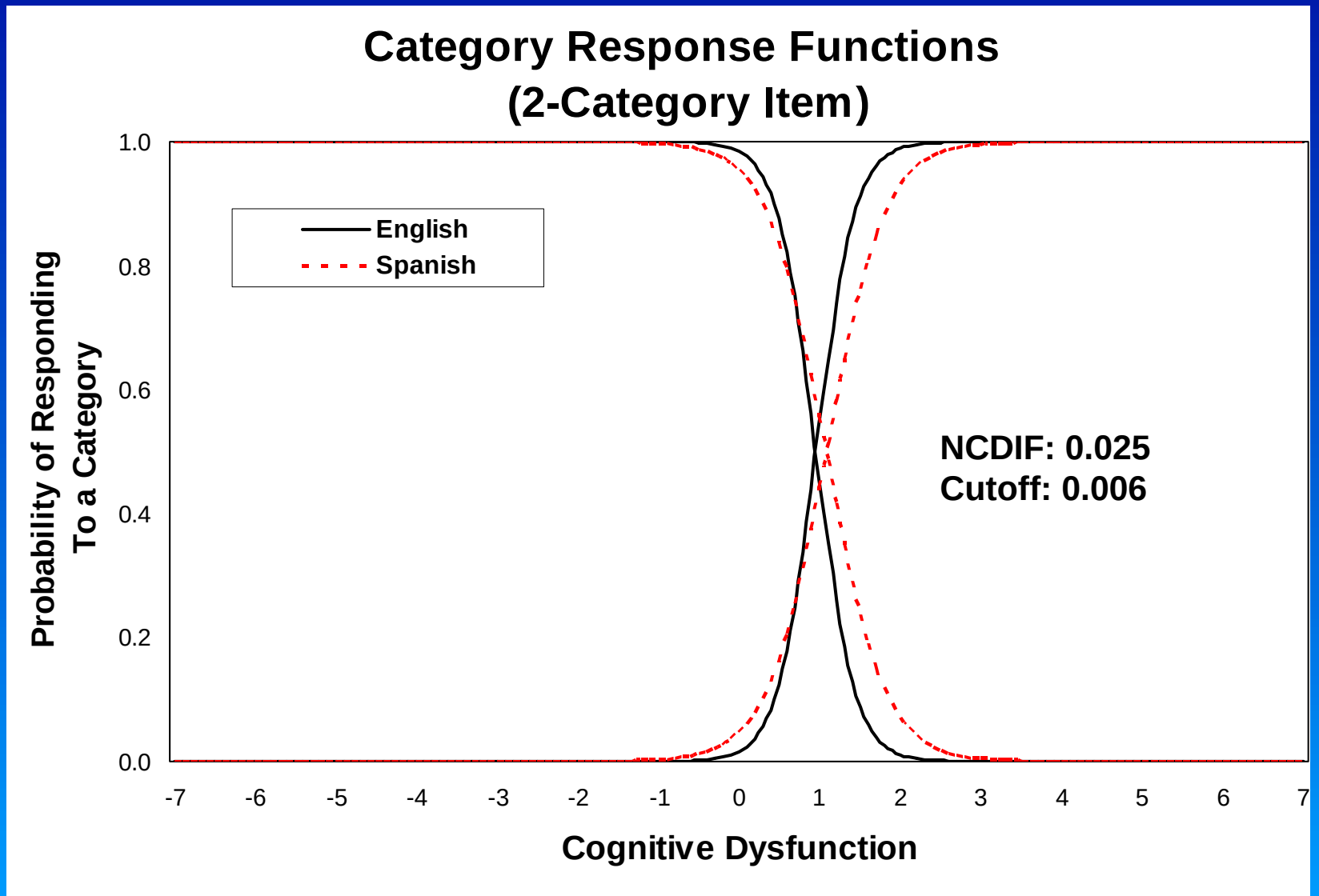
MMSE 3: Name Day of Month



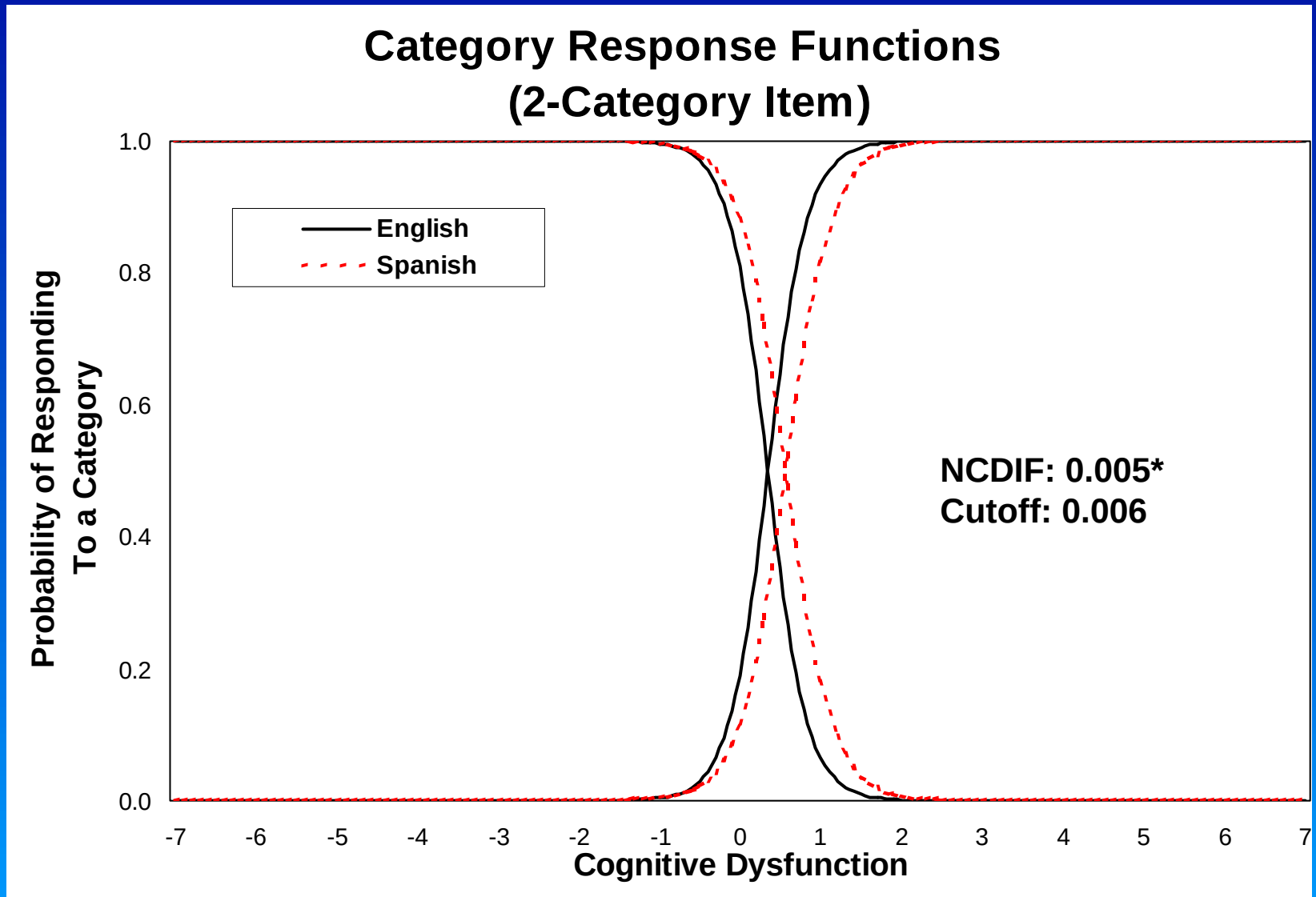
MMSE 6: Name State



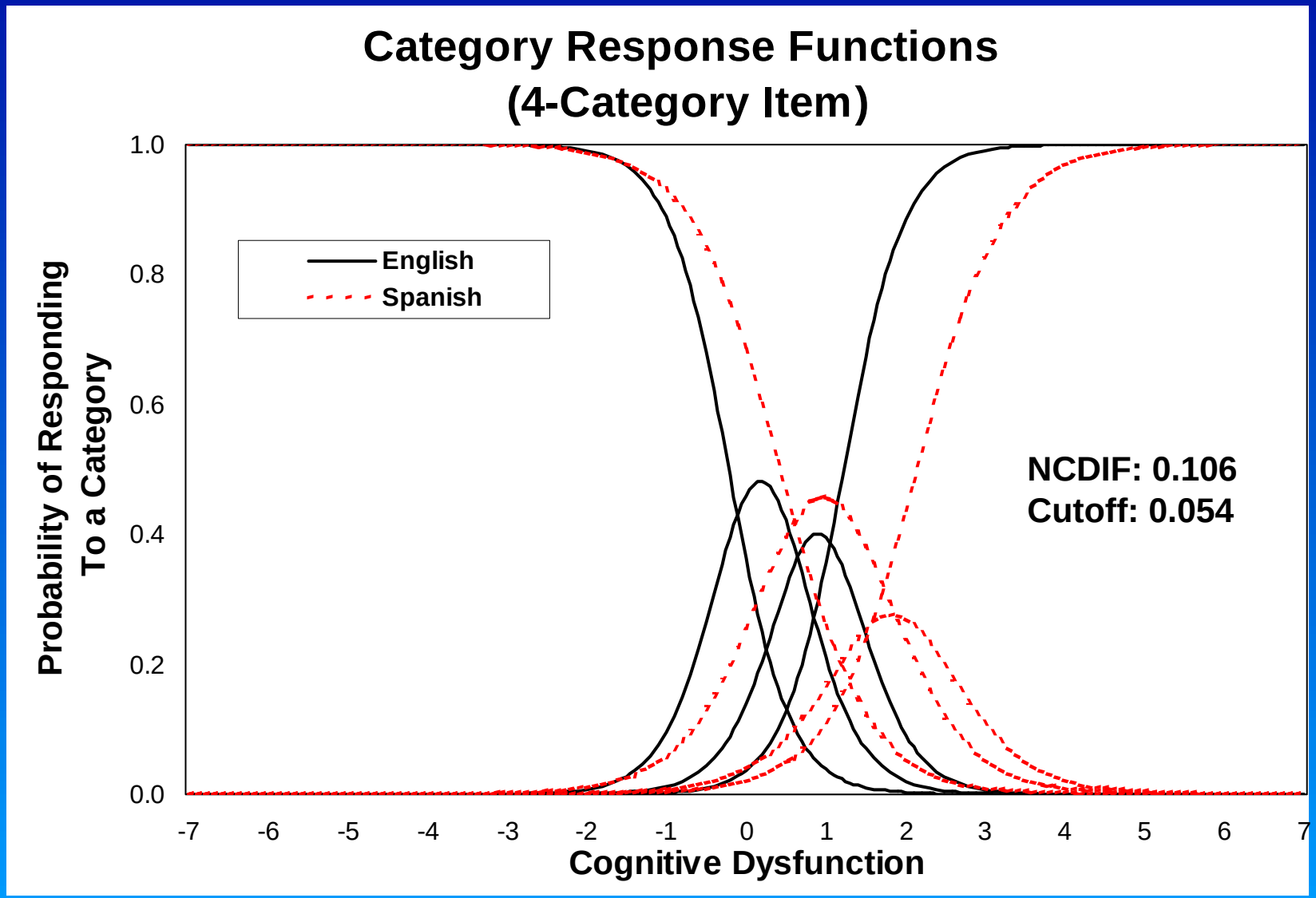
MMSE 7: Name City



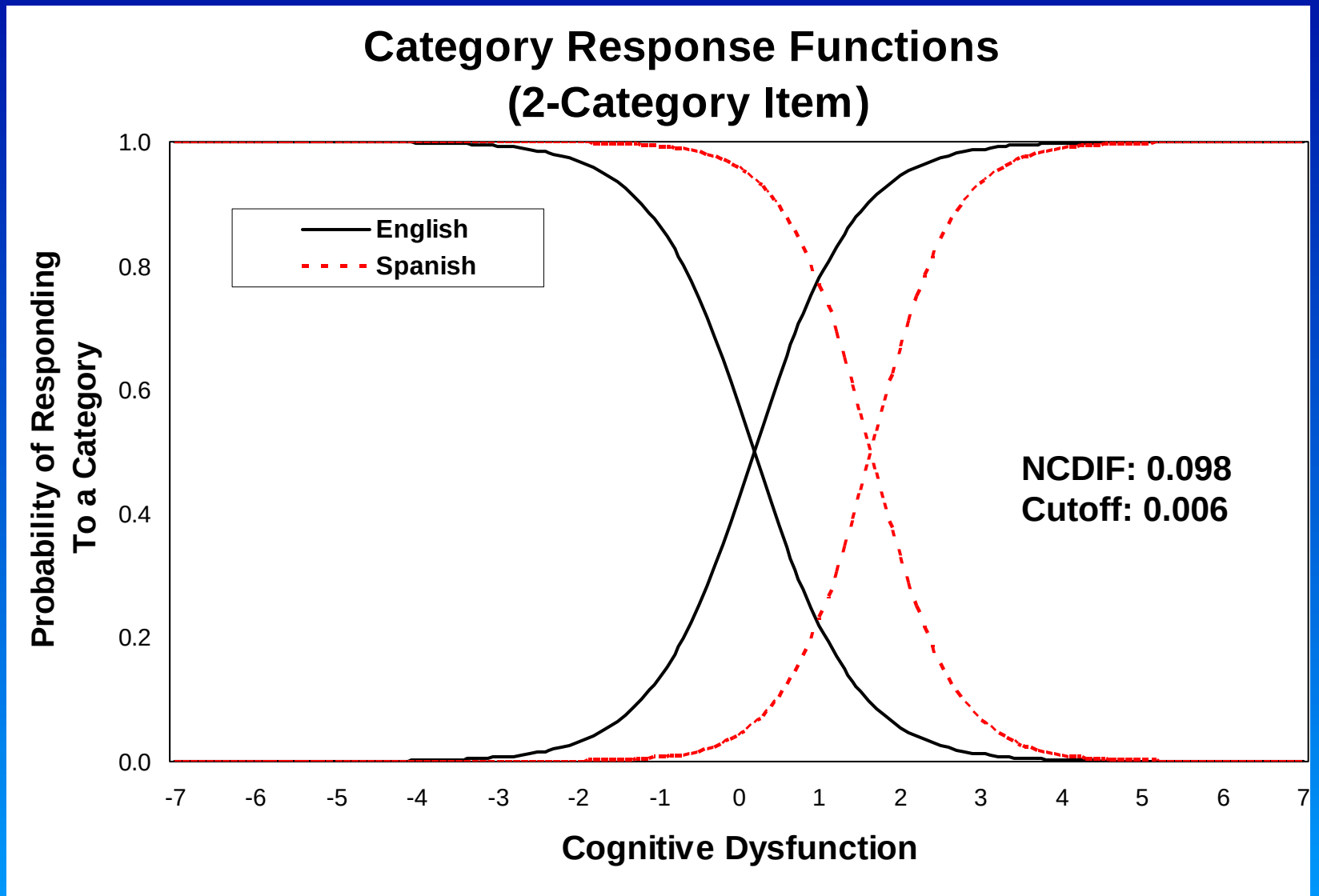
MMSE 8: Name 2 Nearby Streets



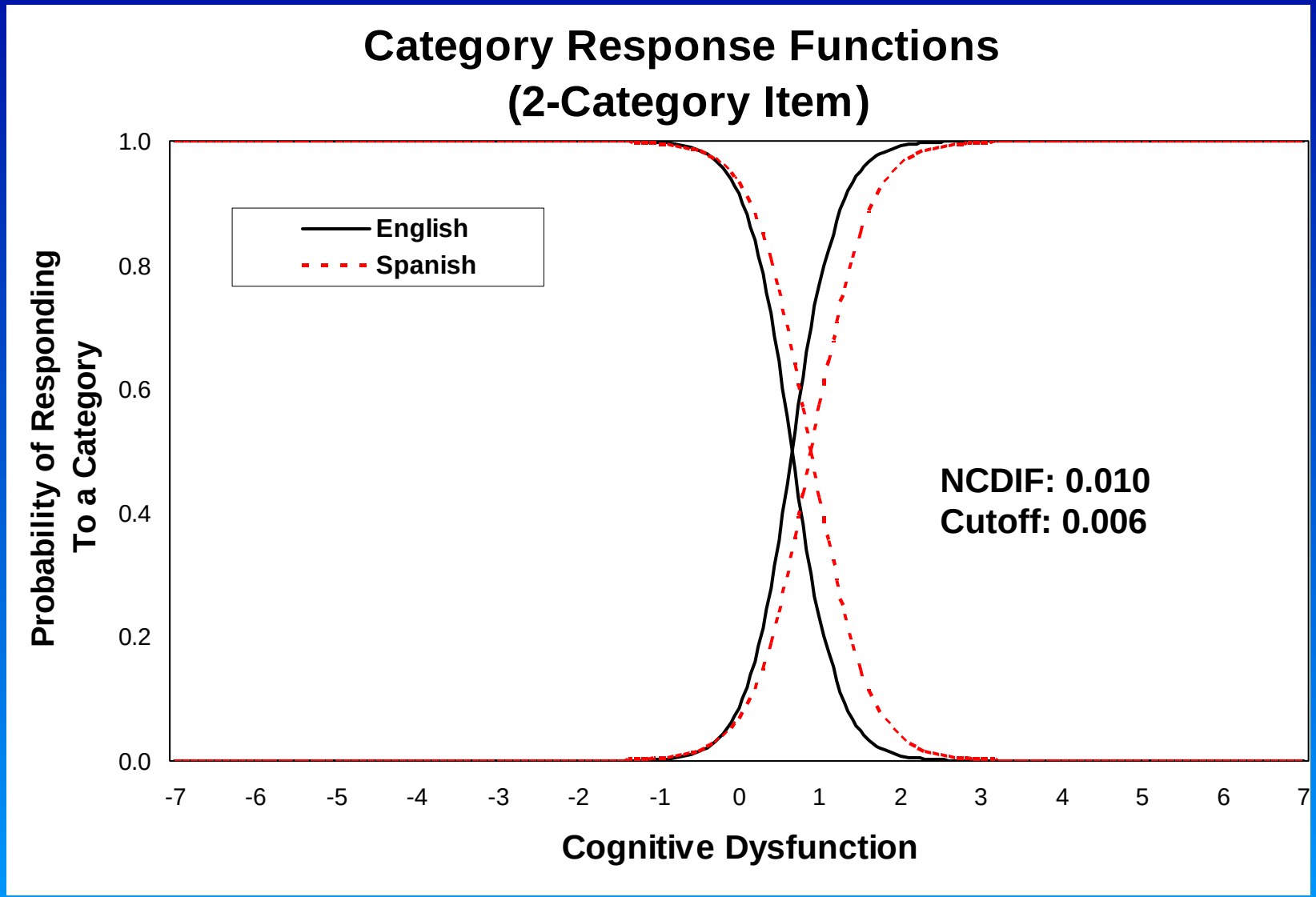
MMSE 14: Recall 3 Objects



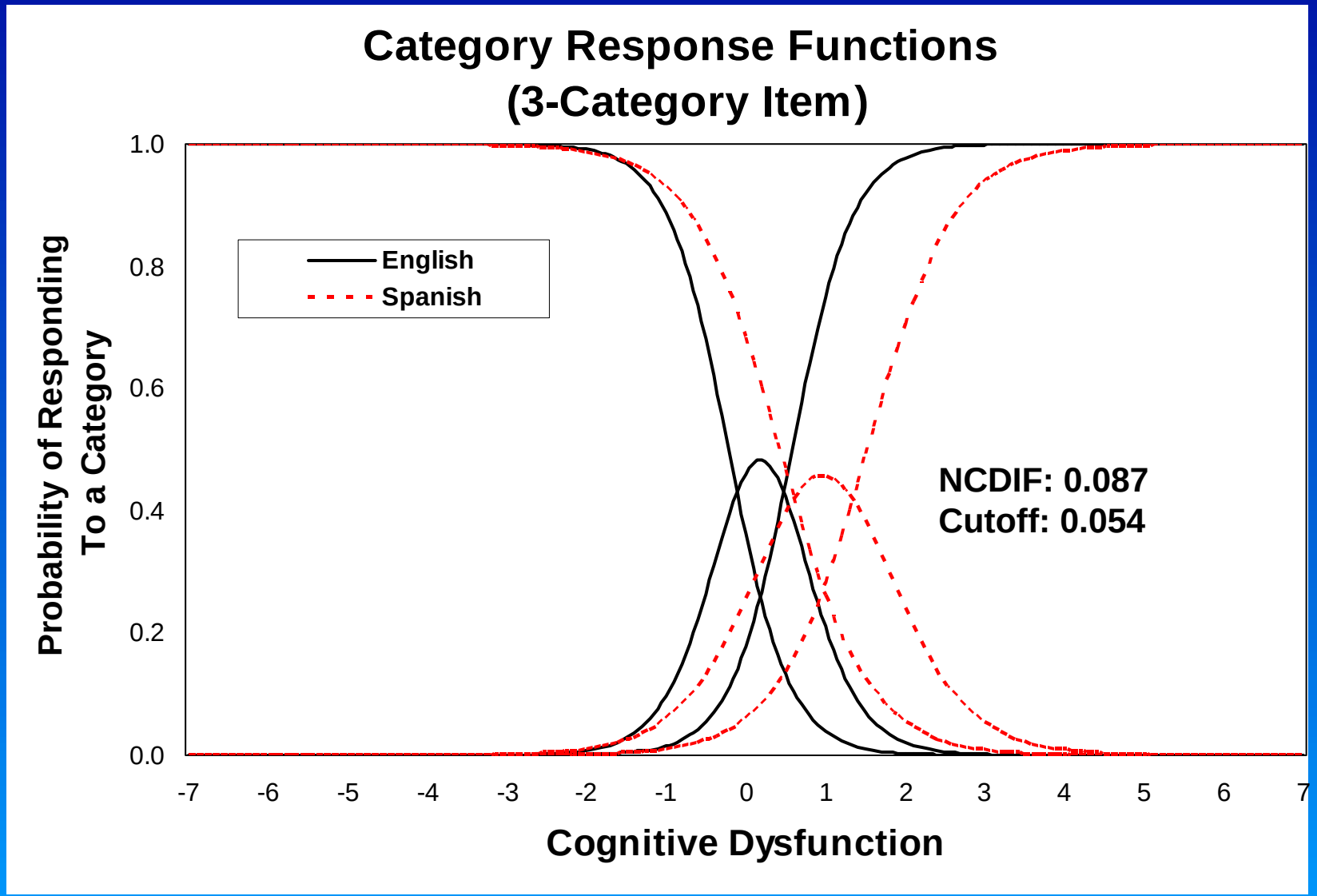
MMSE 17: Repeat Phrase



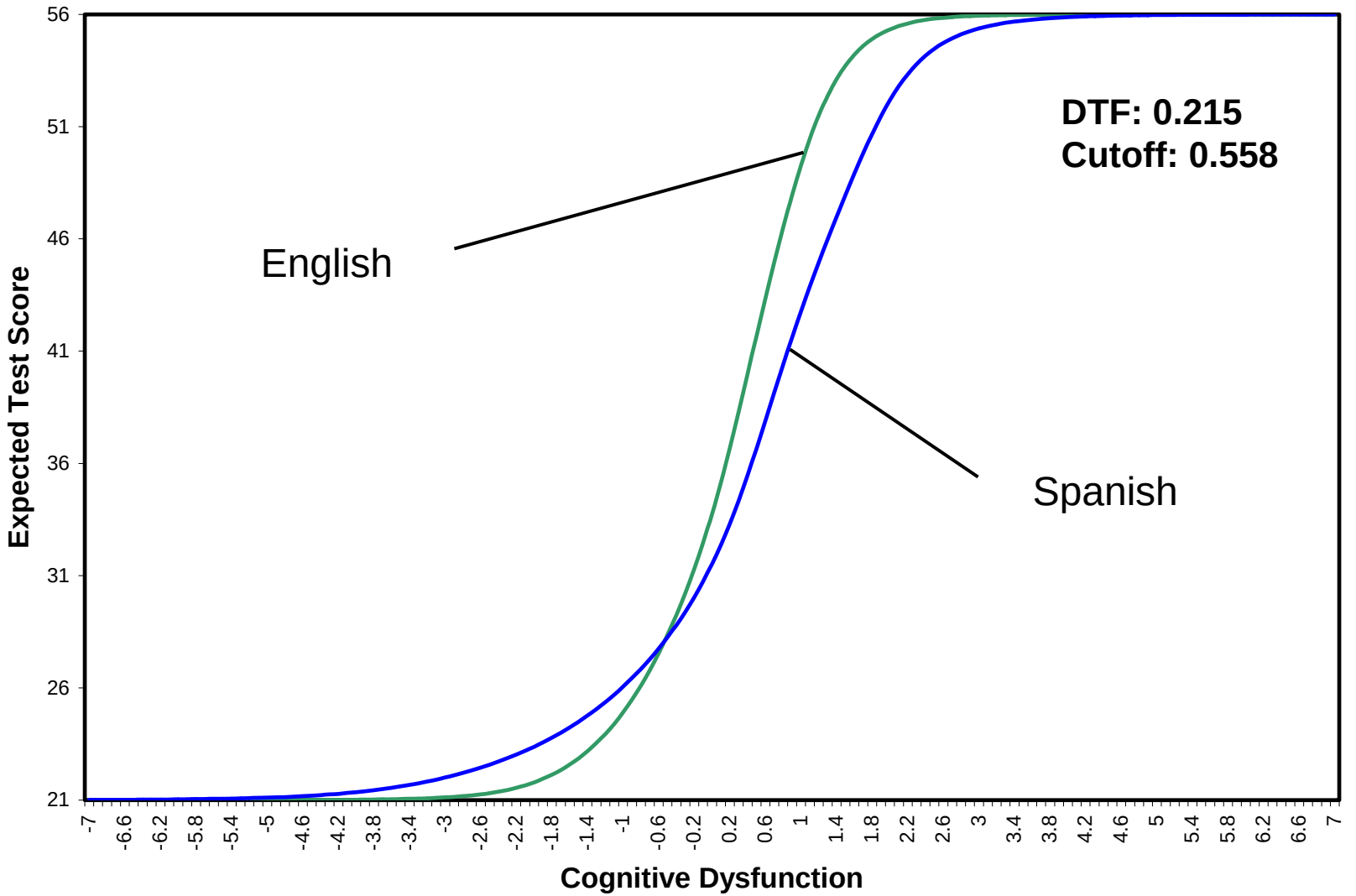
MMSE 18: Close Eyes



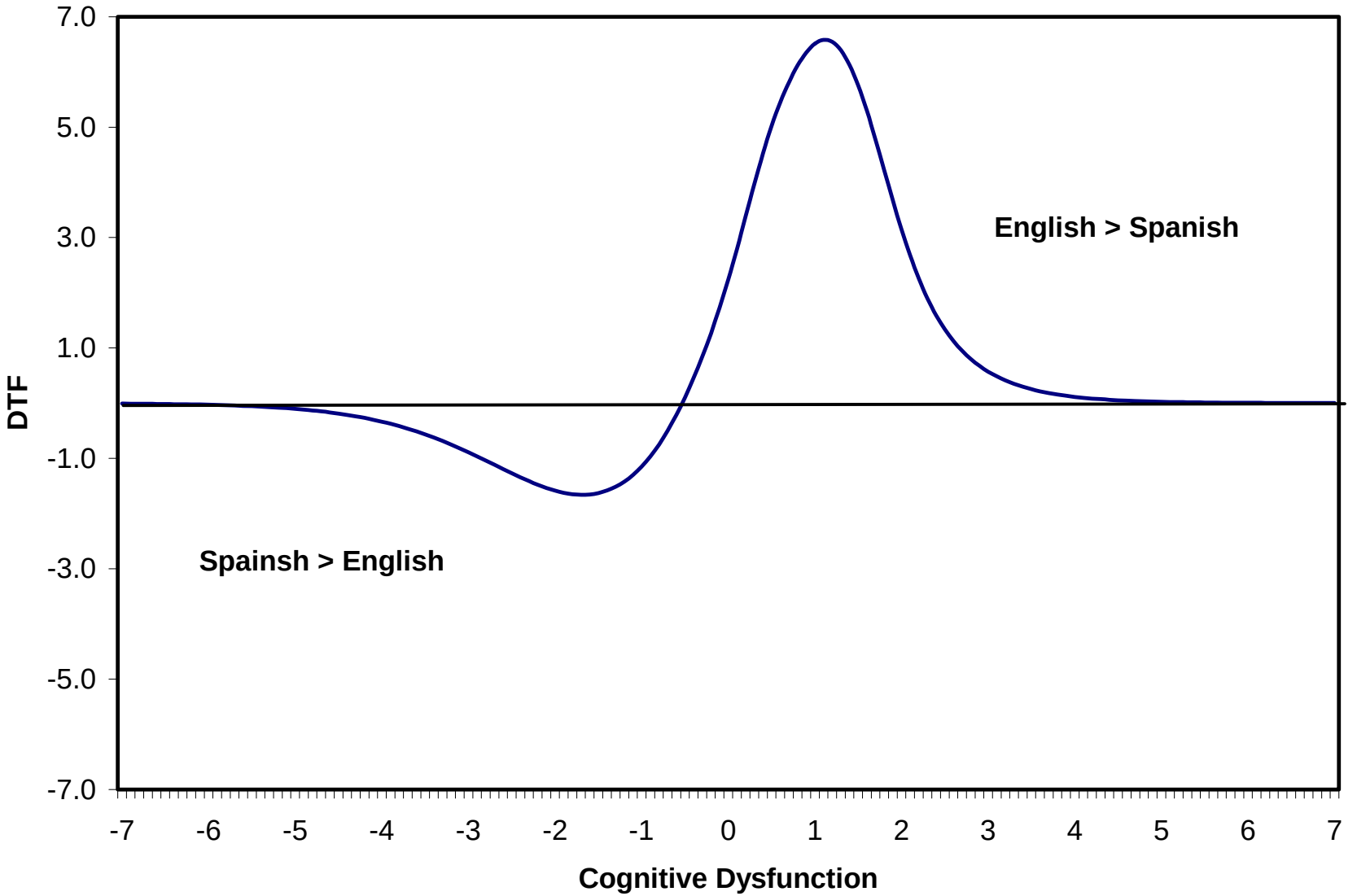
MMSE 19: Instructions with Paper



MMSE Scale Response Functions



Differential Test Functioning



Item Parameters for Non-DIF Items

Item	a	b_1	b_2	b_3	b_4	b_5
MMSE1	4.27	0.49				
MMSE4	3.09	0.61				
MMSE5	3.52	0.60				
MMSE9	3.31	0.99				
MMSE10	3.01	0.90				
MMSE11	3.18	1.44				
MMSE12	3.74	1.31				
MMSE15	2.64	0.58				
MMSE16	1.51	-0.25				
MMSE17	2.58	1.33	1.53	1.73		
MMSE20	1.50	-1.55	-0.79	-0.33	0.02	0.77
MMSE21	2.18	-0.67	-0.28	0.17	0.62	0.98

Dichotomous DIF Items

	English		Spanish	
Item	**a**	**b**	**a**	**b**
MMSE2	2.96	0.62	2.10	2.10
MMSE3	2.18	-0.02	0.15	0.15
MMSE6	2.26	2.26	1.40	1.40
MMSE7	1.09	1.09	0.21	0.21
MMSE8	3.59	3.59	2.93	2.93
MMSE17	1.27	1.27	1.97	1.97
MMSE18	0.42	0.42	1.54	1.54

Polytomous DIF Items

Item	Lang	a	b_1	b_2	b_3
MMSE14	English	1.53	-1.53	-0.84	-0.20
MMSE14	Spanish	1.53	-1.53	-0.57	0.17
MMSE19	English	2.14	-0.08	0.90	1.69
MMSE19	Spanish	1.91	0.38	1.42	2.02

Impact of DIF on MMSE Scale Scoring

Model Assumptions	Group Mean Score Difference: English-Spanish
Standardized Observed Scores	0.12 (t=2.37; p=0.02)
DIF-Adjusted IRT Scores Constrained: 1,4-5,8-16,20-21 Unconstrained: 2-3,6-7,17-19	0.09 (t=1.78; p=0.08)

Summary of Findings

Item-Level Findings:

- **Based on chi-square statistics, all items show DIF.**
- **Based on NC-DIF cutoffs, 9 items have DIF (MMSE2, MMSE3, MMSE6, MMSE7, MMSE8, MMSE14, MMSE17, MMSE18, MMSE19).**

Scale-Level Findings:

- **Based on chi-square statistics, MMSE shows DF**
- **Based on DTF cutoff, the MMSE does not shows differential test functioning**
- **The group mean score difference is probably not “clinically significant”**

Conclusions

Some Spanish and English items in the MMSE show different measurement characteristics.

Comparisons between respondents to the English and Spanish respondents are probably unbiased, whether using IRT to account for items with DIF or not.

END