

# **State of the Science of Health Literacy Measures: Validity Implications for Minority Populations**

**6<sup>th</sup> Health Literacy Annual Research Conference  
November 3-4, 2014**

Tam H Nguyen, PhD, MSN/MPH, RN  
Hae-Ra Han, PhD, RN  
Hyun Jeong Park, PhD, FNP, RN  
Michael Paasche-Orlow, MD, MA, MPH  
Kim, Miyong, PhD, RN

---

- Limited HL is a problem for many minority populations
- Existing reviews have not systematically assessed the validity of current measures for minority populations (Mancusco, 2009; Luk & Aslani, 2011)
- Using tools that are not well-validated for a given population can lead to measurement error (Nunnally & Bernstein, 1994)

# Background

---

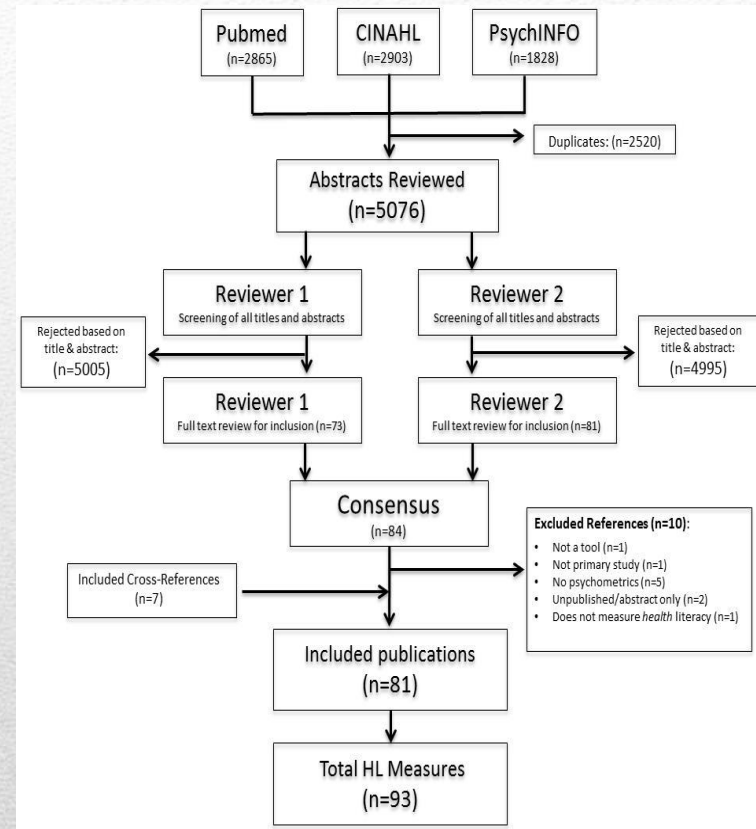


- Systematically review the literature and examine the validity of existing HL measures for a wide range of minority populations

# Study Aim

---

- 3 databases searched
- January 1966 and September 2013
- Terms/synonyms related to “health literacy” AND “measures”
- Included: Articles that reported original validation data on a HL measure, shorted measure, or translated measure
- Excluded: If psychometrics were not reported



# Methods



(health literacy [mh] OR "health literacy" OR (health [tiab] AND (literate [tiab] OR literacy [tiab])) OR numeracy [tiab])

**AND**

(measur\* [tw] OR language tests [mh] OR psychometrics [mh] OR tool\* [tiab] OR survey\* [tiab] OR instrument\* [tiab] OR questionnaire\* [tiab] OR screen\* [tw] OR assessment [tiab])

# **Methods:**

## **Example of PubMed search**

---

Author, Year	Instrument Name	Description (Items)	Scoring	Time, min	Reliability	Validity	Sample	Domain Measured
Davis et al, 1991	Rapid Estimate of Adult Literacy in Medicine (REALM)	General list of medical words in descending levels of difficulty (125-items)	Possible range: 0-125 Low HL: 0-78: $\leq 3^{rd}$ grade 79-103: $4^{th}$ - $6^{th}$ grade 104-114: $7^{th}$ - $8^{th}$ grade 115-125: $\geq 9^{th}$ grade	~3-5	Test-retest: $r=.98$ Inter-rater: $r=.99$	Convergent: SORT-R <sup>1</sup> , $r=.95$ PIAT-R <sup>1</sup> , $r=.94^*, .81^*$	(n=207) 54% Black 46% White	Print
Davis et al, 1993	REALM (Shortened)	General list of medical words in descending levels of difficulty, shortened (66-items)	Possible range: 0-66 0-18: $\leq 3^{rd}$ grade 19-44: $4^{th}$ - $6^{th}$ grade 45-60: $7^{th}$ - $8^{th}$ grade 61-66: $\geq 9^{th}$ grade	~1-3	Test-retest: $r=.99$	Convergent: SORT-R <sup>1</sup> , $r=.96$ WRAT-R <sup>1</sup> , $r=.88$ PIAT-R <sup>1</sup> , $r=.97^*$	(n=203) 76% Black 24% White	Print
Bass et al, 2003	REALM-R (Revised)	General list of medical words in descending levels of difficulty, shortened (8-items)	Possible range: 0-8 ↑scores = ↑HL	<2	Internal Consistency: $\alpha=.91$	Convergent: WRAT-R <sup>1</sup> , $r=.64$	(n=157) Unspecified	Print
Aronzallah et al, 2007	REALM-SF (Shorten Form)	General list of medical words in descending levels of difficulty, shortened (7-items)	Possible range: 0-7 0: $\leq 3^{rd}$ grade 1-3: $4^{th}$ - $6^{th}$ grade 4-6: $7^{th}$ - $8^{th}$ grade 7: $\geq 9^{th}$ grade	<1	N/A	Convergent: REALM, $r=.96$ WRAT-R <sup>1</sup> , $r=.83$ CFA <sup>1</sup> : $R^2=.92, 1\%$ $\chi^2=14.54(p=.41)$	(n=1500) 65% Black 30% White 5% 'Other'	Print
Hanson-Divers, 1997	Medical Achievement Reading Test (MART)	General medical word recognition test modeled after the WRAT <sup>1</sup> , with small font size and glossy covering to resemble prescription bottles (42-items)	Possible range: 0-42 *Converted into grade level equivalences similar to the WRAT <sup>1</sup>	<5	Internal Consistency: $\alpha=.98$	Convergent: WRAT <sup>1</sup> , $r=.98$	(n=405) 38% Black 65% White 6% 'Other'	Print
Lee et al, 2006	Rapid Estimate of Health Literacy for Spanish-speaking Adults (SAHLSA-50)	Spanish general medical word recognition and matching test (50-items)	Possible range: 0-50 $\leq 37$ = inadequate HL	~3-6	Test-retest: $r=.86$ Internal Consistency: $\alpha=.92$	Convergent: TOFHLA, $r=.65$ REALM, $r=.76$ Concurrent: Physical health status, $\beta=.17$	(n=403) 50% Spanish-speaking 50% English-speaking	Print

# Methods:

## Example of extraction tables



- 93 HL measures identified
  - 47 general, 46 content/context specific
  - 20 REALM or REALM-*like*
  - 14 TOFHLA or TOFHLA-*like*
- 65 were English language measures
- 28 were non-English measures
- Reliability was generally strong across all measures
- Weaker evidence of validity
  - 82 used Classical Test Theory
  - 11 used Item Response Theory/Rasch alone or with CTT

# Results

---

- Of the 65 English language measures:
  - 15 measures did not specify the racial/ethnic characteristic of sample
  - Of the remaining 50 measures,
    - 7 did not include any blacks (14%)
    - 29 did not include any Hispanics (58%)
    - 33 did not include any Asians (66%)
  - When Hispanic and Asian Americans were included, they usually accounted for a small % of overall sample

# Results

---



- Of the 28 “other” language measures
  - Translations for REALM (n=3) and TOFHLA (n=9)
    - Issues with phonetic structure of language
    - Issues with cultural equivalence
  - New HL measures in languages other than English (n=16)
    - Poor description of subgroups sampled
    - For some cultural groups this is more/less important

# Results

---

- Many HL measures exist
- Most have not been properly validated for minority populations.
  - Challenges exist with translations & cultural equivalence
  - Problematic b/c most measures are validated using CTT vs IRT
- To address this issue researchers/clinicians have started to translate and developing new ones.

# Conclusions

---



- 93 HL measures -> public repository
- Pilot testing should be done if not yet validated
- Use Item Response Theory/Rasch Modeling with Classical Test Theory to guide instrument development and refinement
- Think about how to collaboratively “re-engineer” HL measurement

# Practice Implications

---



# Questions?

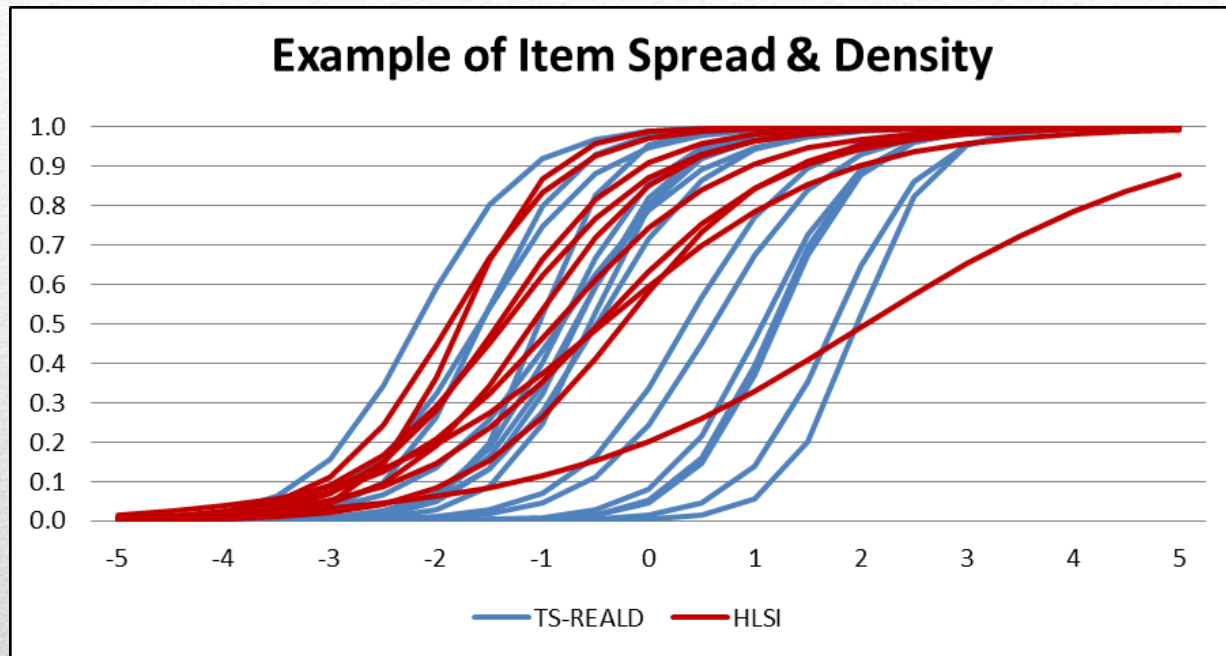
[tam.nguyen@bc.edu](mailto:tam.nguyen@bc.edu)

# Thank you

---



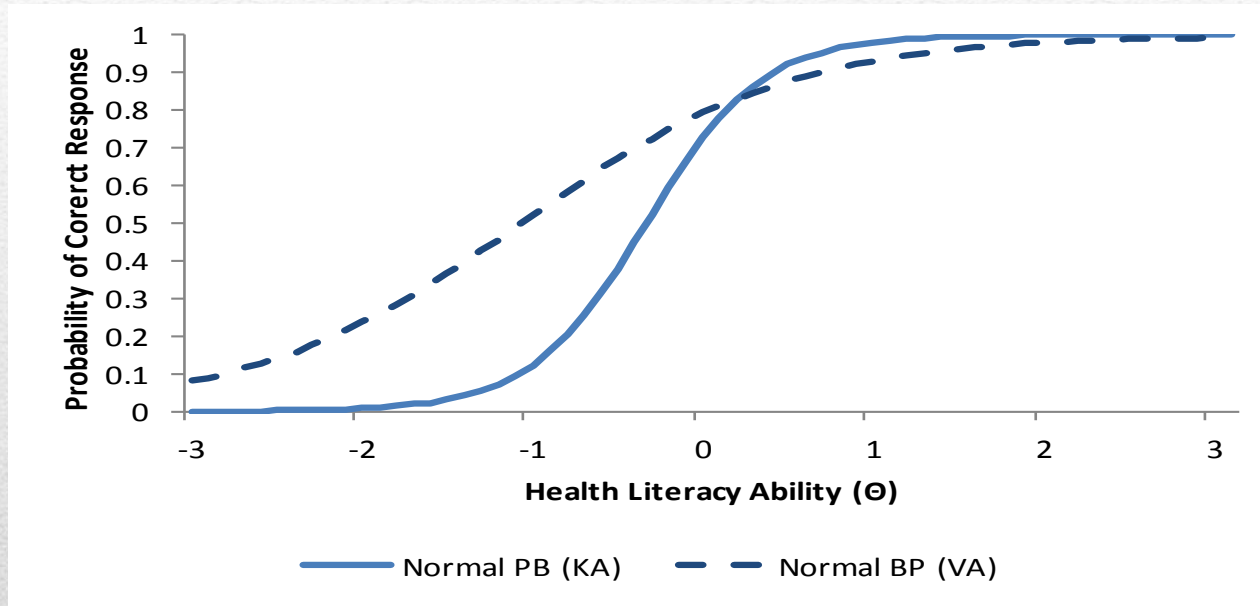
- Item **Difficulty**, **Discrimination**, and **Spread**



# Extra slide (IRT/RASCH)

---

## ○ Examining “Differential Item Functioning” (DIF)



# Extra slide (IRT/RASCH)

---