

Understanding the Washington Group Measures:

Distinguishing disability status from
response error

Goals of This Analysis

1. Assess how well the WG questions work to broadly identify disabled people for prevalence estimates
2. Assess how each question captures functional difficulties within its specific domain

Goal 1: Assess how well the WG questions work to broadly identify disabled people

- Taken as a set can they be used to construct meaningful prevalence estimates
- Distinguishing “disabled” from “non-disabled”
- Not distinguishing types of disability, e.g., blind people from deaf people

Goal 2: Assess how well each question captures the specific domain

- Determine extent to which each question falsely identifies people as having a disability
- Determine reason for misidentifications
 - e.g. question design flaw, translation problem
- Determine which (if any) population may be more likely to be misidentified
 - e.g. less educated, particular country, elderly

Outline of Presentation

- Background information
 - Purpose of cognitive test
 - Instrument and method for collecting data
 - Brief review of initial pattern analysis
- Goal 1 Analysis -- “prevalence”
- Goal 2 Analysis – “false positives”

Purpose of Cognitive Testing

- Interviews: Designed to examine the stages of the question response process
 - Comprehension, retrieval, judgment, response
- Analysis:
 - Identify potential response errors
 - Identify patterns of interpretation

Cognitive Interview Protocol

- I. Demographic Section: Country, language, gender, age, SES
- II. Question Testing Sections
 - A. Core Question
 - B. Interviewer Coding
 - C. Open-ended Follow-up Probe
 - D. Cognitive Follow-up Probe
 - E. Functioning Follow-up Probe
- III. Health Questions: subjective health, chronic condition list
- IV. Interviewer Debriefing

Data Collection for Cognitive Test

- 15 Countries: South & Central American, Asia, Africa
- Country roles:
 - Trained interviewers (with prepared materials)
 - Conducted interviews
 - Entered data
 - Prepared Excel spreadsheet
 - Performed quality check
 - Performed initial analysis
 - Sent data to WG for larger, combined analysis
- Total Sample: N=1290

Initial Analysis of Test Data

- Examined the consistency between Washington Group question responses and follow-up questions responses
- Goal: explain the discrepancies
 - Misunderstood word?
 - Cultural difference?
 - Translation problem?
 - Interviewer error?
 - Error in the follow-up questions?
 - WG Question captures more dimensions of the disability

Initial Analysis: Problematic Responses

Inconsistencies between the WG question and follow-up questions

Do you have difficulty seeing, even if wearing glasses? No

Do you have difficulty seeing and recognizing a person you know from 7 meters (20 feet) away? A lot of difficulty

Do you have difficulty seeing the print in a map, newspaper or book? Some difficulty

Vision

	Washington Group	
	No Difficulty	Yes Difficulty
Not Problematic Responses	617 53.1%	352 30.3%
Problematic Responses	45 3.9%	149 12.8%

Vision Response Patterns

Pattern	WG Disability	Wears Glasses	Follow-up Disability
A	No	No	No
B	No	Yes	No (corrected)
C	Yes, disability	No	Yes
D	Yes, disability	Yes	Yes (not corrected)
E	Yes, disability	Yes	No (corrected)
F	Yes, disability	No	No
G	No	Yes	Yes (not corrected)
H	No	No	Yes

Reasons for inconsistency

1. True response error (in WG or follow-ups)—**potential for Bias**
2. Characteristic of respondent's disability not captured in follow-up questions
3. Data entry/Interviewer error

Hearing Response Patterns

Pattern	WG Disability	Aid	Missed Words	Functioning Problem
A	No	No	No	No
B	Yes	No	Yes	Yes
C	No	No	Yes	No
D	No	No	No	Yes
E	Yes	No	No	Yes
F	Yes	No	Yes	No
G	No	No	Yes	Yes
H	Yes	No	No	No
I	No	Yes	Yes	Yes
J	Yes	Yes	Yes	Yes
K	No	Yes	No	No
L	Yes	Yes	No	Yes
M	Yes	Yes	Yes	No
N	No	Yes	No	Yes
O	No	Yes	Yes	No
P	Yes	Yes	No	No

Generating a Prevalence Rate

- Non-problematic Response Patterns (clear cells)
 - Consistent responses to Washington Group and Follow-up Questions
 - Easy to classify as either disabled or not disabled
- Problematic Response Patterns (light and dark shaded cells)
 - Inconsistent responses make them difficult to categorize
 - Unclear if respondent is disabled or not disabled
 - Unclear if response contains error

How to work with problematic cases:

- Goal 1: Assess whether the problematic cases matter when combining 6 questions to create 1 prevalence rate
- Goal 2: Conduct more analysis to identify the cases of “true response error”

Analysis Goal 1

Assess how well the WG questions work to broadly identify disabled people for prevalence estimates

Disabilities Across Domains (“some” threshold)

	Inconsistent responses	Pct. With Difficulties in Another Domain
Vision	149	75.2
Hearing	59	76.3
Mobility	104	82.7
Cognition	306	80.1
Self-Care	115	96.5

Disabilities Across Domains (“a lot” threshold)

	Inconsistent responses	Pct. With Difficulties in Another Domain
Vision	17	52.9
Hearing	14	57.1
Mobility	20	90.0
Cognition	45	88.9
Self-Care	31	96.8

Correlation across domains stronger among those aged 65+ (“some”)

	Inconsistent	Pct. With Other Difficulties
Vision	22	72.7
Hearing	11	81.8
Mobility	16	87.5
Cognition	40	85.0
Self-Care	8	100.0

Vision: Breakdown by Pattern and Difficulty Threshold

	“Some”	“A Lot”
Inconsistent	29.7	9.2
Somewhat inconsistent	0.0	0.0
Not Problematic	70.3	90.8

Hearing: Breakdown by Pattern and Difficulty Threshold

	“Some”	“A Lot”
Inconsistent	7.3	1.9
Somewhat inconsistent	15.4	11.1
Not Problematic	77.2	87.0

Mobility: Breakdown by Pattern and Difficulty Threshold

	“Some”	“A Lot”
Inconsistent	4.4	1.1
Somewhat inconsistent	14.8	6.2
Not Problematic	80.8	92.8

Cognition: Breakdown by Pattern and Difficulty Threshold

	“Some”	“A Lot”
Inconsistent	27.3	8.3
Somewhat inconsistent	34.6	33.3
Not Problematic	38.2	58.3

Self-Care: Breakdown by Pattern and Difficulty Threshold

	“Some”	“A Lot”
Inconsistent	24.5	15.8
Somewhat inconsistent	29.7	25.0
Not Problematic	45.8	59.2

Prevalence and Equalization of Opportunities

- Using the WG questions together gives us a good methodology for dividing the population up into those with and without disabilities
- But can we delve deeper into the reasons behind the response errors?

Analysis Goal 2

Assess how each question captures functional difficulties within its specific domain

Goal 2: Assess how well each question captures the specific domain

- Determine extent to which each question falsely identifies people as having a disability
- Determine reason for misidentifications
- Determine which (if any) population may be more likely to be misidentified

- Analytically, this means we must examine the problematic cases
 - Which cases are true response error?
(*Should not include in measure*)
 - Which cases are characteristics of disability?
(*Should include in measure*)

Potential False Negatives/Positives

	False Negative	False Positive
Vision	3.9%	12.8%
Hearing	9.9%	5.4%
Mobility	19.7%	8.6%
Cognition	10.3%	25.4%
Self-Care	17.1%	9.5%

Potential False Negatives/Positives

	False Negative		False Positive	
	Less Problematic	More Problematic	Less Problematic	More Problematic
Vision	0%	3.9%	0%	12.8%
Hearing	6.7%	2.3%	3.7%	1.8%
Mobility	14.9%	4.8%	6.6%	2.0%
Cognition	8.7%	1.6%	14.2%	11.2%
Self-Care	12.9%	4.1%	5.2%	4.3%

Why it's important to identify cases of "true error"

- Bias if there is pattern in the error
 - Gender
 - Country
 - Age
 - Disability or Health Status
- Conduct demographic analysis of error cases to identify bias

Vision: Problematic Responses

Inconsistencies between the WG question and follow-up questions

Do you have difficulty seeing, even if wearing glasses? No

Do you have difficulty seeing and recognizing a person you know from 7 meters (20 feet) away? A lot of difficulty

Do you have difficulty seeing the print in a map, newspaper or book? Some difficulty

Vision

	Washington Group	
	No Difficulty	Yes Difficulty
Not Problematic Responses	617 53.1%	352 30.3%
Problematic Responses	45 3.9%	149 12.8%

Identifying True False Positives

12.8% (149 cases) of Potential False Positives

Next analysis to identify true error:

Additional follow-ups:

1. With your glasses, how often do you have difficulty seeing well?
2. With your glasses, how much effort do you have to put into seeing?

Vision False-positive Errors

- **53.7% No Effort and No Frequency**
 - 80 out of the 149 potential false positives
 - 71 were Pattern E, 9 were Pattern F
- **15.4% No Effort or No Frequency**
 - 23 out of the 149 potential false positives
 - 22 were Pattern E, 1 were Pattern F

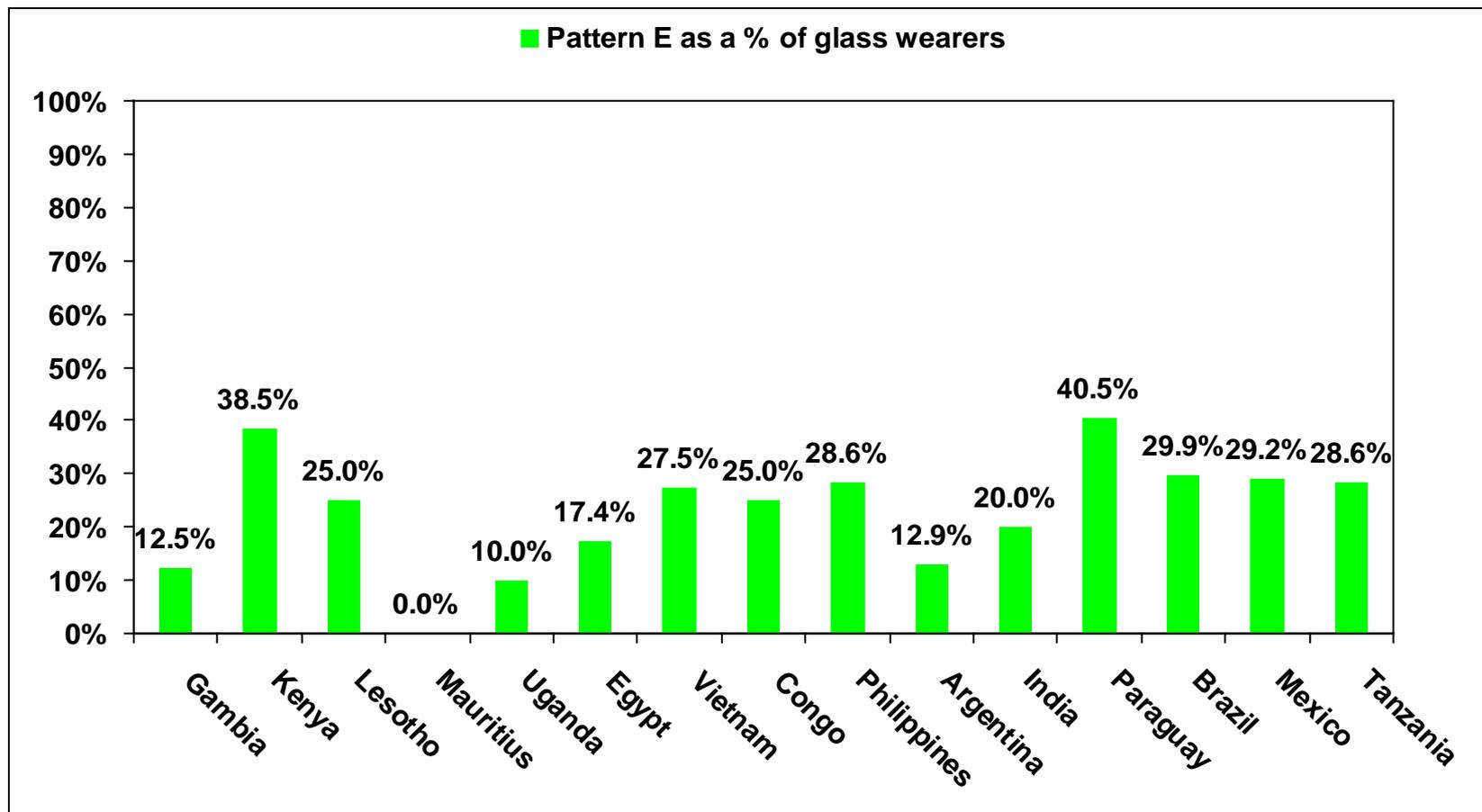
Conclusions: True Error for Vision

- Pattern E:
 - Cases in Pattern E are likely true error
 - Related to the glasses clause
- Pattern F:
 - More problems with effort and frequency
 - Cannot assume is error
 - Not to be included in demographic/bias analysis

Vision: Demographic/Bias Analysis

- Does Pattern E occur more often among specific subgroups?
 - Country
 - Age
 - Gender
 - Disability Status

Glasses Clause Problems by Country



Glasses Clause Problems by Age

	10-44	45-64	65+	p-value
Pattern E	6.1%	17.3%	14.7%	p<.001
Wear Glasses	24.6%	62.0%	67.9%	p<.0001
Percent of glass wearers who are in Pattern E	25.0%	28.7%	21.5%	p=.42

Glasses Clause Problems by Gender

	Female	Male	p-value
Pattern E	8.1%	11.9%	$p < .05$
Wear glasses	38.3%	39.7%	$p = .6175$
Percent of glass wearers in pattern E	21.4%	29.6%	$p < .05$

Glasses Clause Problems by Disability in Other Domains

	No	Yes	P-value
Pattern E	9.7%	10.4%	$p=.72$
Wear glasses	34.9%	40.8%	$p=.06$
Percent of glass wearers in pattern E	27.2%	25.7%	$<.75$

Cognition: Do you have difficulty remembering or concentrating?

Do you have difficulty remembering the names of people or places?

Do you have difficulty remembering appointments?

Do you have difficulty remembering how to get to familiar places?

Do you have difficulty remembering important tasks, like taking medications or paying bills?

Do you have difficulty concentrating on doing something for ten minutes?

‡Do you have difficulty learning a new task, for example, learning how to get to a new place?

‡Do you have difficulty finding solutions to problems in day to day life?

Response patterns: Cognition

Pattern	Disability	Functioning Problem					
		Yes/No Questions		Scale Questions TEN/NEW/SOLUT		Combined Questions Yes/No and Scale	
		NAMES/APPT	TASK/PLACE	One	2 or 3	2 or 3 (excluding E & L)	4 +
A	No	No	No	No	No	-----	-----
B	No	No	No	Yes	-----	-----	-----
C	No	Yes	No	No	No	-----	-----
D	No	No	No	-----	Yes	-----	-----
E	No	Yes	No	Yes	-----	-----	-----
F	No	Yes or No	Yes	No	No	-----	-----
G	No	-----	-----	-----	-----	Yes	-----
H	No	-----	-----	-----	-----	-----	Yes
I	Yes	-----	-----	-----	-----	-----	Yes
J	Yes	-----	-----	-----	-----	Yes	-----
K	Yes	Yes or No	Yes	No	No	-----	-----
L	Yes	Yes	No	Yes	-----	-----	-----
M	Yes	No	No	-----	Yes	-----	-----
N	Yes	Yes	No	No	No	-----	-----
O	Yes	No	No	Yes	-----	-----	-----
P	Yes	No	No	No	No	-----	-----

Cognition

	Washington Group	
	No Difficulty	Yes Difficulty
Not Problematic Responses	586 48.6%	189 15.7%
Problematic Responses	124 10.3%	306 25.4%

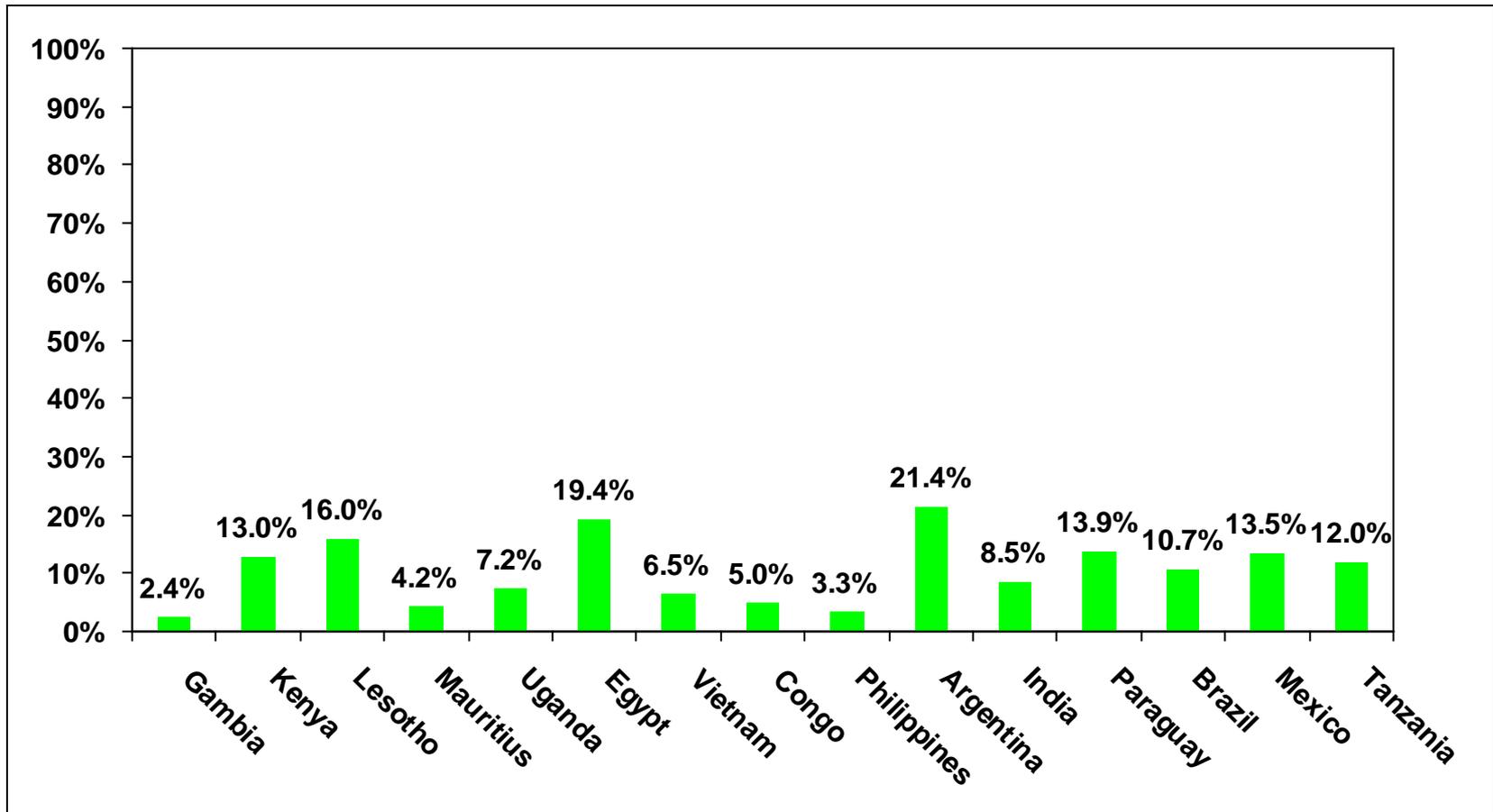
Cognitive Potential False Positives

- 25.4% (306 out of 1205) are potential false positives.
- Only two of these cases answered no to both the frequency and effort questions.
- Different from Vision; Likely to be more of an interpretation issue, not blatant misunderstanding

Follow-up Questions by Positive Response Patterns

	More Problematic Patterns	Less Problematic Patterns	All others	p-value
Too busy	42.8%	36.8%	23.4%	<.001
Effort	21.0%	37.5%	58.1%	<.0001
Frequency	10.8%	30.5%	49.7%	<.0001
Worry	13.4%	24.4%	40.9%	<.0001

Percentage of Respondents in More Problematic Patterns by Country



Percentage of Respondents in More Problematic Patterns by Age

Pattern	10-44	45-64	65+
More Problematic Patterns	10.4%	11.2%	15.6%
All others	89.6%	88.9%	84.4%

Chi-square = 3.01, DF=2, p = .2215

Percentage of Respondents in More Problematic Patterns by Gender

Pattern	Female	Male
More Problematic Patterns	12.0%	10.5%
All others	88.0%	89.5%

Chi-square = .6160, DF = 1, p = .4325

Conclusions

- WG questions taken as a group are good at generating general prevalence estimate
- Confirm that glasses clause is significant issue, but needs to be addressed at country level – language, custom
- Country differences in response error are significant -- suggests need for country specific cognitive testing in question development
- Preliminary results suggest no real sign of demographic bias