**UBC**

**The University of British Columbia**

# DIF and Various Item Response Formats: A study of the CES-D

Michaela N. Gelin, Bruno D. Zumbo, & Anita M. Hubley
The University of British Columbia

Presented at: National Council on Measurement in Education, Seattle, WA, April 12, 2001

A copy can be found on the web at:
www.educ.ubc.ca/faculty/zumbo/ins2001/index.html

## **Abstract**:

We report on a psychometric study of the Center for Epidemiologic Studies Depression (CES-D) scale with 600 community-dwelling adults between the ages of 17 and 87 years. The mean age for males is 46 years (N=310) and 42 years for females (N=290).  The 20-item CES-D was scored using two binary methods (presence and persistence) and one ordinal method. Gender differential item functioning (DIF) was explored using Zumbo's (1999) ordinal logistic regression method with corresponding effect size estimator with all three scoring methods. After statistically matching males and females on the underlying ability, gender DIF was found with the CES-D item *crying* for the ordinal and presence methods of scoring.  Two gender DIF items (*effort* and *hopeful*) were identified for the persistence scoring method; however, this scoring method was of limited use due to low response rates on some items.  Overall, the results indicate that the scoring method has an effect on DIF; thus DIF is a property of the item, scoring method, and purpose of the instrument.

## Research Problem:
• To investigate whether, for each scoring method, any CES-D items exhibit gender DIF

## Rationale:
• Few studies have explored gender DIF in depression measures
• Only one study has investigated gender DIF for the CES-D and this study (Cole et al., 2000) focused on one sample of seniors 65 years of age or older
• No study has compared DIF for ordinal and binary item formats on the same scale

## Purposes:
• To investigate gender DIF for the CES-D
• To investigate whether different scoring methods affect the DIF results.

## Measure:

- The CES-D (Radloff, 1977) is a widely used self-report measure developed for use in studies exploring the epidemiology of depressive symptomology in the general population.
- Used in numerous studies to compare the prevalence of depressive symptomology among men and women
- 20 items - each with four options
- 3 scoring methods: ordinal and two binary (presence & persistence)

## Participants:

- 600 community-dwelling adults living in Northern British Columbia who were part of a survey carried out by the Institute for Social Research and Evaluation (ISRE) at the University of Northern British Columbia, Canada, in the fall of 1998.

| Sex | Mean Age in Years | Std. Deviation Age | Minimum Age | Maximum Age | N |
|---|---|---|---|---|---|
| Female | 42.19 | 13.44 | 18 | 87 | 290 |
| Male | 46.05 | 12.07 | 17 | 82 | 310 |
| Total | 44.19 | 12.88 | 17 | 87 | 600 |

Figure 1.  Center for Epidemiologic Studies Depression (CES-D) Scale.


**Center for Epidemiologic Studies Depression (CES-D)
Scale: Format for Self-Administered Use**

---

INSTRUCTIONS:  Using the scale below, please circle the number for each statement that best describes how often you felt or behaved this way during the past week.

> 0 = Rarely or none of the time (less than 1 day)
> 1 = Some or a little of the time (1-2 days)
> 2 = Occasionally or a moderate amount of time (3-4 days)
> 3 = Most or all of the time (5-7 days)

---

| DURING THE PAST WEEK: | Less than 1 day | 1-2 days | 3-4 days | 5-7 days |
|---|---|---|---|---|
| 1. I was bothered by things that usually don't bother me. | 0 | 1 | 2 | 3 |
| 2. I did not feel like eating; my appetite was poor. | 0 | 1 | 2 | 3 |
| 3. I felt that I could not shake off the blues even with help from my family or friends. | 0 | 1 | 2 | 3 |
| 4. I felt that I was just as good as other people. | 0 | 1 | 2 | 3 |
| 5. I had trouble keeping my mind on what I was doing. | 0 | 1 | 2 | 3 |
| 6. I felt depressed. | 0 | 1 | 2 | 3 |
| 7. I felt that everything I did was an effort. | 0 | 1 | 2 | 3 |
| 8. I felt hopeful about the future. | 0 | 1 | 2 | 3 |
| 9. I thought my life had been a failure. | 0 | 1 | 2 | 3 |
| 10. I felt fearful. | 0 | 1 | 2 | 3 |
| 11. My sleep was restless. | 0 | 1 | 2 | 3 |
| 12. I was happy. | 0 | 1 | 2 | 3 |
| 13. I talked less than usual. | 0 | 1 | 2 | 3 |
| 14. I felt lonely. | 0 | 1 | 2 | 3 |
| 15. People were unfriendly. | 0 | 1 | 2 | 3 |
| 16. I enjoyed life. | 0 | 1 | 2 | 3 |
| 17. I had crying spells. | 0 | 1 | 2 | 3 |
| 18. I felt sad. | 0 | 1 | 2 | 3 |
| 19. I felt that people dislike me. | 0 | 1 | 2 | 3 |
| 20. I could not get "going". | 0 | 1 | 2 | 3 |

Note:  Items are summed after reverse scoring of items 4, 8, 12, and 16.  Total CES-D scores range from 0-60, with higher scores indicating higher levels of general depression.

Gelin, Zumbo, & Hubley     NCME 2001

<u>Figure 2</u>.  Scoring the CES-D.

# Scoring the CES-D

Each item has four options:

        Option 0, *rarely or none of the time / less than 1 day*

        Option 1, *some or a little of the time / 1-2 days*

        Option 2, *occasionally or a moderate amount of the time / 3-4 days*

        Option 3, *most of the time / 5-7 days*

**ORDINAL** scoring method

      &#10148; All four options are scored 0, 1, 2, or 3, respectively

      &#10148; The total score ranges from 0 - 60.

**PRESENCE** scoring method

  &#10148; The respondent's report of having experienced the symptom at least *some of the time* during the preceding week (i.e. for 1 to 7 days).

Option 0 $\xrightarrow{\text{is assigned}}$ 0 (indicating no depression)
Option 1 ⎫
Option 2 ⎬ assigned 1 (indicating depression)
Option 3 ⎭

  &#10148; The total score ranges from 0 – 20.

**PERSISTENCE** scoring method

  &#10148; The respondent's report of having experienced the symptom for *3-7 days* during the preceding week.

Option 0 ⎫
Option 1 ⎬ 0 (indicating no depression)
Option 2 ⎫
Option 3 ⎬ 1 (indicating depression)

  &#10148; The total score ranges from 0 – 20.

Gelin, Zumbo, & Hubley     NCME 2001

# Results

- As a preliminary to the DIF analyses, the results from a confirmatory factor analysis showed that a unidimensional model with method effects modelled for the four positively worded items fit well. (Zumbo, Gelin, & Hubley, in press; 2001).
- Differential item functioning (DIF) was explored using Zumbo's (1999) ordinal logistic regression & corresponding effect size estimator.

1. <u>Ordinal scoring method</u>
   - *Crying* item displayed large gender DIF (predominantly uniform)
2. <u>Presence scoring method</u>
   - *Crying* item displayed large gender DIF (predominantly uniform)
3. <u>Persistence scoring Method</u>
   - Items 7 (effort) and 8 (hopeful) showed moderate gender DIF (predominantly uniform)
   - A number of items (including *crying*) could not be computed because there was a low probability of endorsing the items

## DIF results from Zumbo's (1999) ordinal logistic regression

| DIF Computation | Ordinal Scoring | Presence Scoring | Persistence Scoring | |
|---|---|---|---|---|
| | Item 17 (*crying*) | Item 17 (*crying*) | Item 7 (*effort*) | Item 8 (*hopeful*) |
| $\chi^2(2)$ | 56.56 | 43.74 | 8.45 | 7.81 |
| $P$ | 0.000 | 0.000 | 0.007 | 0.010 |
| $R^2$ | 0.218 | 0.191 | 0.045 | 0.040 |

\* <u>Note:</u>  All other items showed no DIF with a $p<0.01$ and R-squared greater than or equal to 0.035

# <u>Discussion</u>

- Crying item displays gender DIF for ordinal <u>and</u> presence scoring methods
- Results found with persistence scoring method should be interpreted with caution
- Persistence scoring method is not appropriate with data from a general population
- DIF is a property of the scoring method

# References:

Cole, S.R., Kawachi, I., Maller, S.J., & Berkman, L.F. (2000). Test of item-response bias in the CES-D scale: Experience from the New Haven EPESE study. <u>Journal of Clinical Epidemiology, 53</u>, 285-289.

Radloff, L.S. (1977). The CES-D scale: A self-report depression scale for research in the general population. <u>Applied Psychological Measurement, 3</u>, 385-401.

Zumbo, B. D., Gelin, M. N., & Hubley, A. M. (in press). The construction and use of psychological tests and measures. <u>Encyclopedia of Life Support Systems</u>. United Nations Educational, Scientific and Cultural Organization Publishing (UNESCO-EOLSS Publishing), France.

Zumbo, B. D., Gelin, M. N., & Hubley, A. M. (February, 2001). <u>Psychometric Study of the CES-D: Factor Analysis and DIF</u>. Paper presented at the International Neuropsychological Society's 29th Annual Meeting, Chicago, IL.

Zumbo, B.D. (1999). <u>A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores</u>. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense. *Contact the author for a copy.