

Original Research Report

Positive Aspects of Family Caregiving for Dementia: Differential Item Functioning by Race

David L. Roth,¹ Peggye Dilworth-Anderson,² Jin Huang,¹ Alden L. Gross,³ and Laura N. Gitlin⁴

¹Division of Geriatric Medicine and Gerontology, School of Medicine, Johns Hopkins University, Baltimore, MD. ²Department of Health Policy and Management, University of North Carolina – Chapel Hill. ³Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD. ⁴Department of Community-Public Health, School of Nursing, Johns Hopkins University, Baltimore, MD.

David L. Roth, PhD, Division of Geriatric Medicine and Gerontology, 2024 East Monument Street, Johns Hopkins University, Baltimore, MD 21205. E-mail: droth@jhu.edu.

Received October 9, 2014; Accepted April 14, 2015

Decision Editor: Bob G. Knight, PhD

Abstract

Objectives. Due to increasing interest in the positive experiences associated with family caregiving, potential demographic group differences were examined on the Positive Aspects of Caregiving (PAC) scale at both the item and scale levels.

Method. Family caregivers ($N = 642$) completed the PAC as part of their participation in the Resources for Enhancing Alzheimer's Caregiver Health (REACH II) clinical trial. Multiple indicators, multiple causes models were used to examine potential differential item functioning (DIF) across demographic subgroups.

Results. Overall PAC scale scores indicated that both Hispanics and African Americans experienced more PAC than Whites. Two items with statistically significant ($p < .004$) and practically meaningful (odds ratio > 2.0) DIF were found for African American caregivers. After controlling for the underlying unidimensional construct, African Americans reported that caregiving gave them “a more positive attitude toward life” and enabled them to “appreciate life more” than either Whites or Hispanics. No instances of meaningful DIF were found between Hispanics and Whites, women and men, or spouses and nonspouses.

Discussion. PAC scores differ significantly by race. In addition, 2 items with meaningful race DIF identify content areas that are particularly relevant to the cultural experiences of African American caregivers.

Key Words: Caregiving—Dementia—Differential item functioning (DIF)—Measurement—Minority and diverse populations

Caring for, and being cared for by, members of one's family are universal human experiences that are desired by virtually everyone across the life span. Providing care for an older family member with dementia is one important and increasingly common type of caring within the family (Centers for Disease Control and Prevention [CDC], 2010; Pruchno & Gitlin, 2012). Most persons with dementia live at home, and most of the assistance they receive comes

from nonprofessional, family caregivers (Schulz & Martire, 2004). Although providing care to a family member with dementia is often described as being stressful and burdensome (Ory, Hoffman, Yee, Tennstedt, & Schulz, 1999; Pinguart & Sörensen, 2003), there is an increasing emphasis on balancing this point of view with a better understanding of the full range of caregiving experiences, including an increased appreciation of the potential positive aspects and

benefits of caregiving (Brown & Brown, 2014; Roth, Fredman, & Haley, 2015; Zarit, 2012).

Recent population-based studies have shown that informal caregivers, as a general group and including, but not limited to, dementia, have significantly reduced all-cause mortality rates in comparison with specific noncaregiving comparison samples (Brown et al., 2009; Fredman et al., 2010; O'Reilly et al., 2008; Roth et al., 2013). Caregivers who report little or no strain from caregiving activities have also been found to report better health-related quality of life than noncaregivers (Roth et al., 2009). Both the stressors and the potential benefits associated with dementia caregiving depend, to some extent, on many contextual factors, including cultural and gender-based considerations that can affect caregiving expectations both within families and throughout their surrounding communities.

In addition to comparing the well-being of caregiving and non-caregiving comparison groups, another approach to better understanding the potential positive experiences associated with family caregiving is to directly ask caregivers about such experiences. The Positive Aspects of Caregiving (PAC) scale was designed specifically for this purpose (Tarlow et al., 2004). Each item on the PAC asks about potential benefits from caregiving for the caregivers.

The potential benefits that might result from providing care to a family member with dementia or another disabling illness might be perceived or experienced quite differently by caregivers from different backgrounds or distinct demographic subgroups. Men, for example, might experience and report the positive aspects of dementia caregiving differently from women, and a substantial literature has identified both gender (e.g., Pinquart & Sörensen, 2006; Yee & Schulz, 2000) and race (e.g., Dilworth-Anderson et al., 2002; Pinquart & Sörensen, 2005) differences in family caregiving experiences, stressors, and outcomes. Culturally, African Americans may sometimes report more positive caregiving outcomes due to experiences that foster active coping styles and create attitudes of greater resilience, which can then be used to buffer negative life experiences such as caregiving to older relatives with dementia (Dilworth-Anderson et al., 2005; McCallum, Sorocco, & Fritsch, 2006; Merritt, McCallum, & Fritsch 2011). This is consistent with findings that African American caregivers often report less depression, stress, and strain due to caregiving than Whites, although both groups report negative physical health effects of caregiving over time (Calderón & Tennstedt, 1998; Haley et al., 1996; Haley, et al., 2004; Roth, et al., 2001). Findings also show that cultural values and beliefs serve as the lens through which caregiving is experienced, shaping both its perceived threats and its potential positive effects (Dilworth-Anderson et al., 2005; Dilworth-Anderson, Pierre, & Hilliard, 2012; Goins et al., 2011; Knight & Sayegh, 2010).

In addition to differences by race and gender shaping the caregiving experiences, caregiving relationship also shapes these experiences. Spouse caregivers provide more hours of care and report more depression symptoms, greater financial and physical burden, and lower levels of psychological well-being than adult-child caregivers (Pinquart & Sörensen, 2011). Lin, Fee, and Wu (2012) found that wife caregivers, in particular, were least likely to report positive experiences and female and adult-child caregivers reported having had more negative experiences than male and spouse caregivers, respectively. Their findings also show that different risk factors for negative (i.e., care recipients' problem behavior and dependency) and positive (e.g., reciprocal help from care recipients for wife, daughter, and son caregivers, but not for husband caregivers) caregiving experiences are related to caregivers' gender and relationship to the care recipient. Furthermore, differences sometimes attributed

to race might be partially confounded with differences in the caregiving relationship. Spousal caregiving, for example, is more prevalent among older Whites than among African Americans and other minority groups, where adult children and other family members are often the primary caregivers (Pinquart & Sörensen, 2005).

This study provides additional detailed information on understanding how different subgroups of caregivers experience positive aspects of caregiving. When examining group differences on summary scores from self-report instruments, the validity of those comparisons depends on an assumption that the items on the scale have the same meaning and measurement properties across the groups being compared (Stewart & Nápoles-Springer, 2003). This is often an untested assumption, but psychometric analysis methods are available for detecting whether individual items on a scale are differentially sensitive to certain population subgroups. Differential item functioning (DIF) is present when an item from a scale measures a construct differently across two or more subgroups (Holland & Wainer, 1993). Although previous studies have reported race group differences on the overall PAC summary score and on other summary score measures of positive caregiving experiences (Picot, Debanne, Namazi, & Wykle, 1997; Roff et al., 2004), there are no previous examinations of whether individual items with significant DIF by race might have biased or contributed to these overall observed race group differences.

The REACH II data set provides a unique opportunity to evaluate potential DIF on the PAC by race, gender, and relationship (i.e., spouse vs. nonspouse) factors. These analyses were conducted in order to evaluate further the psychometric properties of the PAC items and to determine whether this instrument is suitable for measuring differences in positive caregiving experiences across three race/ethnic groups as well as gender- and relationship-based demographic subgroups.

Method

Participants

The REACH II investigation enrolled 642 dyads that consisted of persons with dementia and their primary family caregivers. Participants were enrolled at five sites: Birmingham, AL; Memphis, TN; Miami, FL; Palo Alto, CA; and Philadelphia, PA. Eligibility criteria for the caregivers included being persons who were providing care to a family member with dementia of the Alzheimer type or a related disorder for at least 4 hr per day over at least the previous six months. Caregivers were at least 21 years of age; of African American ($N = 211$), Hispanic ($N = 212$), or White ($N = 219$) race or ethnicity; and reported at least two signs of caregiving distress at enrollment (e.g., feeling overwhelmed, cutoff from family or friends). The sample of caregivers included 271 persons who were providing care for a spouse with dementia and 371 nonspouse caregivers. There were 110 men and 532 women in the sample. The recruitment procedures and eligibility criteria for REACH II have been described in more detail elsewhere (Belle et al., 2006).

The REACH II data are publically available through the Inter-University Consortium for Political and Social Research, and all analyses reported in this paper are based on data downloaded in October 2013 from the ICPRS website (<http://www.icprs.umich.edu/icprsweb/landing.jsp>).

Measures

The PAC scale was administered as part of a comprehensive baseline battery prior to the delivery of any interventions. The 11-item scale used in REACH II was modified from an earlier 9-item version that

was used in the initial REACH I project (Roff et al., 2004; Tarlow et al., 2004). Caregivers were instructed that “in spite of all the difficulties involved in giving care to a family member with memory or health problems, good things can come out of caregiving experiences, too.” They were then asked to indicate their agreement with 11 statements about possible positive experiences associated providing this type of care. A listing of the 11 PAC items is provided in Table 1. A response card was used that instructed caregivers to reply on a 0 to 4-point scale (0 = disagree a lot, 1 = disagree a little, 2 = neither agree nor disagree, 3 = agree a little, 4 = agree a lot). There were no reverse-scored items. The measure is typically used by obtaining an unweighted sum of the item responses, yielding a score that could range from 0 to 44.

Statistical Analyses

Overall scale differences on the PAC were examined with a multiple regression analysis that included caregiver age, caregiver gender, race (African American, Hispanic, White), and caregiving relationship (spouse vs. nonspouse) as the predictors.

Possible DIF by race, caregiver gender, or caregiving relationship on the individual PAC items was assessed using multiple indicators, multiple causes (MIMIC) models. All MIMIC analyses were conducted using version 7.0 of the Mplus software package (Muthén & Muthén, 2013). A robust version of the maximum likelihood method (MLR estimation method in Mplus) was used to model the ordinal item responses (CATEGORICAL option in Mplus). The general factor analytic model that was used is illustrated in Figure 1. First, a single factor was extracted from all 11 items. Factor loadings were examined, and normalized residuals were calculated based on differences between sample and model-estimated polychoric correlations. These residuals were used to evaluate the fit of an underlying 1-factor model to the observed item data (Bollen, 1989). Next, a binary grouping variable (e.g., African American vs. White) was included to the model that predicted differences in the responses to each individual item (path a) and differences in the mean on the latent variable (path b). The estimate for path a assesses DIF by capturing the effect of the grouping variable on the variability of a specific individual item after accounting for the impact of the underlying latent construct.

Each PAC item was tested individually in separate models, with the first item model illustrated in Figure 1. Similar MIMIC models

have been used in other item response theory analyses of DIF across demographic groups (Jones, 2006; Yang & Jones, 2007). In these models, the nonfocal items were assumed to have fixed (invariant) effects across groups. Alternative MIMIC models using a free baseline-designated anchor approach (Woods, 2009) were also examined. This approach yielded similar and largely consistent results with the DIF reported in this paper when using the fixed effects approach illustrated in Figure 1. Consequently, only the results of the fixed effects approach are reported in this paper, but the comparable findings from the free baseline-designated anchor analyses are available from the authors upon written request.

Because DIF in each of the 11 items was examined in separate models, a Bonferroni correction was applied to the type I error rate, and DIF was considered to be statistically significant only if the estimate for path a was different from 0 at the $\alpha < .004$ (.05/11) level. In addition, in analyses with large samples, items with rather small or trivial amounts of DIF can, nonetheless, sometimes show statistically significant DIF. Consequently, the methods and criteria of Cole, Kawachi, Maller, & Berkman (2000) were used to identify items with practically meaningful DIF. Specifically, the MLR estimate for path a was exponentiated to yield a proportional odds ratio (OR), and if the OR exceeded 2.0 (or was <0.50), then the corresponding DIF was considered to be practically meaningful.

After DIF was examined for grouping factors individually, subsequent models examined DIF for one grouping contrast after controlling for another. Race, for example, overlapped substantially with caregiving relationship in REACH II, and race-based DIF could be overlapping or redundant with relationship-based DIF detected in separate models. In order to examine this possibility, subsequent MIMIC models were conducted that included two group contrasts simultaneously, as illustrated in Figure 2. The covariate-adjusted a paths from this model tested whether DIF on one group contrast (e.g., African American vs. White) was statistically independent and distinct from DIF observed for a correlated grouping variable (e.g., spouse vs. nonspouse caregiving relationship).

Results

In the REACH II sample, the mean unweighted sum of the 11 PAC items was 30.95 (SD = 10.92). This general scale score indicates

Table 1. Odds Ratios for Positive Aspects of Caregiving Items With Bonferroni-Adjusted, Statistically Significant Differential Item Functioning

Item	Comparison				
	African American vs. White	Hispanic vs. White	African American vs. Hispanic	Nonspouse vs. spouse caregiver	Male vs. female caregiver
1. Made me feel more useful	—	—	—	1.97	—
2. Made me feel good about myself	—	—	—	—	—
3. Made me feel needed	0.45	—	—	0.54	—
4. Made me feel appreciated	—	—	—	—	—
5. Made me feel important	—	—	—	—	—
6. Made me feel strong and confident	—	—	—	0.58	—
7. Given more meaning to my life	—	—	—	—	—
8. Enabled me to learn new skills	—	—	0.53	—	—
9. Enabled me to appreciate life more	2.33	—	2.38	1.90	—
10. Enabled me to develop a more positive attitude toward life	3.12	—	2.04	—	—
11. Strengthened my relationships with others	—	—	—	—	—

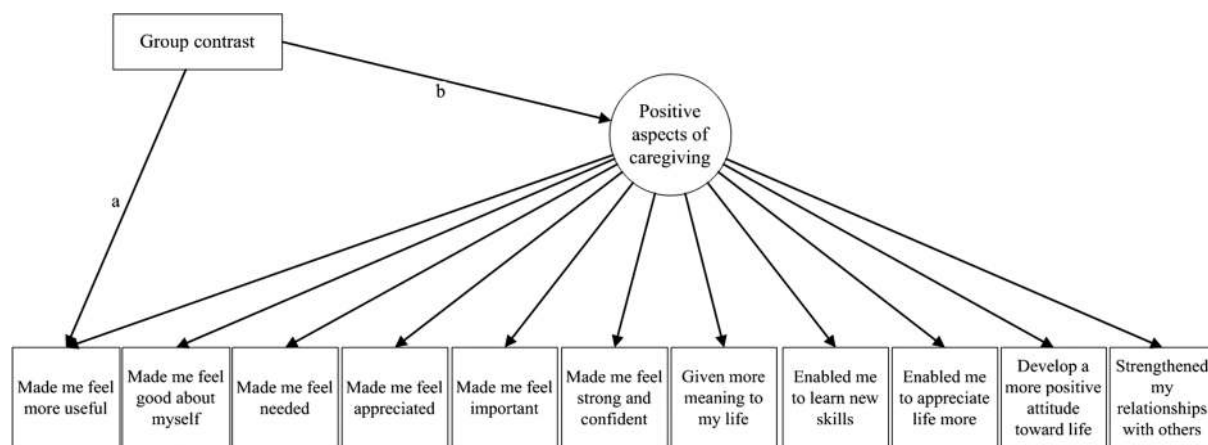


Figure 1. A multiple indicators, multiple causes model for testing differential item functioning in the positive aspects of caregiving measures.

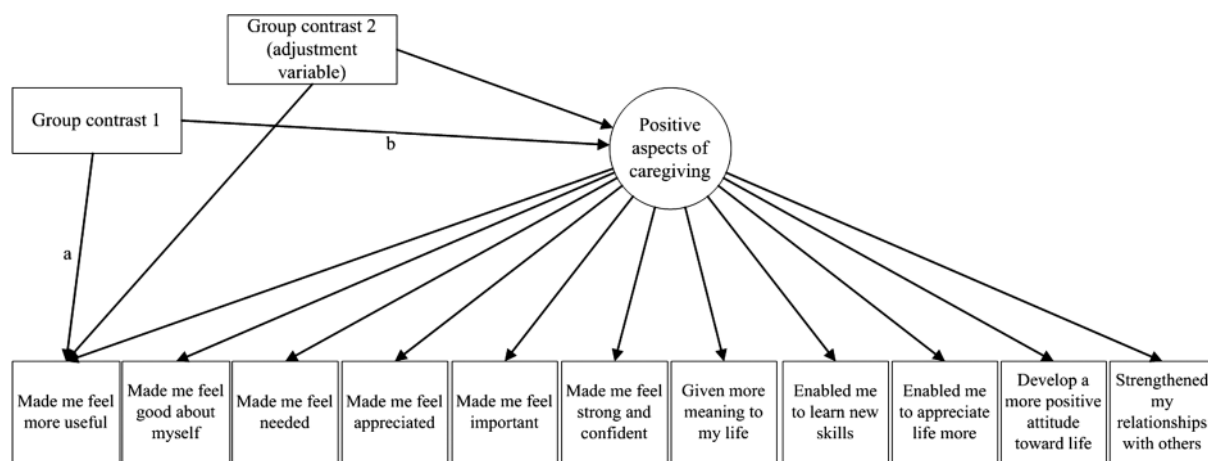


Figure 2. Modified multiple indicators, multiple causes model for testing whether differential item functioning (DIF) identified for one group contrast is independent from DIF identified for a second group contrast.

considerable agreement with the positive statements, on average. The multiple regression analysis of the predictors of the overall 11-item PAC scores is summarized in Table 2. That analysis indicated an overall omnibus effect by race ($F(2, 636) = 28.61, p < .0001$), such that both African American (adjusted $M = 33.24$) and Hispanic (adjusted $M = 34.81$) caregivers had higher scores than White caregivers (adjusted $M = 27.47$). PAC scores also differed significantly by caregiver gender ($t(636) = 6.95, p = .009$) but not by caregiver age or caregiving relationship. Men (adjusted $M = 33.29$) reported significantly more PAC than women (adjusted $M = 30.39$), whereas both spouses and nonspouses reported similar PAC levels.

The underlying 1-factor model estimated prior to any DIF testing was found to have consistently high standardized factor loadings that ranged from 0.67 to 0.87 across the 11 items. These uniformly high loadings indicate that each item was a sensitive indicator of the underlying latent construct, as assessed by the PAC items as a whole. Inspection of the normalized residuals revealed only 2 of the 55 residuals had absolute values >2.0 , indicating that the model-estimated polychoric correlations deviated significantly from the observed sample correlations at only chance levels. Thus, the overall fit of the 1-factor model was judged to be adequate for further DIF testing.

The results of the MIMIC DIF analyses for each binary group contrast on the 11 items of the PAC are summarized in Table 1.

This table lists the statistically significant ORs associated with the a-paths from Figure 1 for each group comparison after applying the Bonferroni-adjusted p -value ($p < .004$). Using these criteria, only 10 of 55 possible instances of DIF were observed to be statistically significant, and only 5 of those 10 instances showed a degree of DIF that was also considered to be practically meaningful according to our a priori OR cut points.

The five instances of significant and meaningful DIF all involved comparisons between African American caregivers and either White or Hispanic caregivers. After adjusting for the impact of the overall latent factor, African American caregivers were less likely than White caregivers to report that caregiving made them “feel needed.” In comparison with both White and Hispanic caregivers, African American caregivers reported that caregiving enabled them to “appreciate life more” and to develop “a positive attitude toward life.” As summarized in Table 1, four items showed statistically significant DIF ($p < .004$) between spouses and nonspouses, but the ORs for these instances of DIF did not exceed our threshold for practical significance. No statistically significant or practically meaningful DIF was observed for any item comparisons by gender or between Hispanics and Whites.

Two items—appreciating life more and developing a more positive attitude toward life—were found to have statistically significant and practically meaningful DIF in both the African American versus

White and African American versus Hispanic comparisons. For the “appreciate life more” item, statistically significant DIF was found in the nonspouse versus spouse comparison, with nonspouses showed the same direction of increased endorsement as did the African Americans. Because African American caregivers were much less likely to be providing care to a spouse than White caregivers (30.3% vs. 57.1%, respectively, $p < .0001$), subsequent MIMIC models examined DIF due to one group contrast (e.g., race) after controlling for another group contrast (e.g., relationship) as illustrated in Figure 2. The results of these adjusted DIF analyses showed that both of the items with African American versus White and African American versus Hispanic DIF still showed statistically significant and practically meaningful DIF after controlling for the nonspouse versus spouse differences: Appreciating life more (African American vs. White OR = 2.12, African American vs. Hispanic OR = 2.24); more positive attitude toward life (African American vs. White OR = 3.18, African American vs. Hispanic OR = 2.00).

Scale Score Comparisons

In some applications, items with significant and meaningful DIF might be dropped from a scale to avoid the potential measurement bias that they can create. Table 3 presents descriptive data for the summary scores from the full 11-item PAC and for a 9-item PAC that eliminates the two items with significant and meaningful DIF from the African American versus White and African American versus Hispanic analyses (items 9 and 10). Both the 11-item and 9-item versions had excellent internal consistency, and both African American and Hispanic caregivers had significantly higher summary scores than White caregivers on both versions ($p < .05$). Although African Americans and Hispanics did not differ significantly on the 11-item PAC, after removing the two items with race DIF, Hispanics were found to have significantly higher scores on the 9-item PAC than African Americans ($t(421) = 2.06, p = .04$). Dividing mean differences by the SD from the total sample yielded standardized mean differences in standard deviation units (SDUs). African Americans had scores that were 0.54 SDUs higher than Whites on the 11-item measure and 0.46 SDUs on the 9-item measure. Therefore, including the two items with DIF in the 11-item version compared with the

9-item version increases the overall standardized difference between African Americans and Whites by approximately 17% ($[0.54 - 0.46]/0.46 * 100$). Conversely, African Americans scored lower than Hispanics by 0.14 SDUs on the 11-item measure and by 0.20 SDUs on the 9-item measure, so including the two items with DIF on the African American versus Hispanic comparison decreased the overall standardized mean difference on the summary score by 30% ($[0.14 - 0.20]/0.20 * 100$). Hispanics scored higher than Whites by 0.67 and 0.65 SDUs on the 11-item and 9-item PAC measures, respectively, suggesting minimal changes on this comparison.

Discussion

Simple summary score comparisons on the 11-item PAC scale used in REACH II indicated that both Hispanics and African Americans reported significantly more positive experiences associated with dementia caregiving than Whites. Statistically significant DIF was observed in only 10 of 55 possible instances for the 11 PAC items examined across race, gender, and relationship comparisons, and in only five of these 10 instances was the DIF also considered to be practically meaningful based on our a priori criteria adapted from Cole et al. (2000). These findings indicate that, in general, the 11-item PAC used in the REACH II trial is mostly free from major item biases across major demographic subgroups, and summary scores should provide valid comparisons that are not overly biased by item content that is differentially sensitive to specific demographic subgroups. However, two items were identified as being particularly sensitive to African American caregivers such that African Americans endorsed these two items at significantly higher levels than would be expected based on their scores on the latent construct (i.e., their responses to the other nine items). Standardized comparisons of the mean differences on the 11-item and 9-item versions of the PAC suggested some measurement bias is introduced by these two items that accentuate African American versus White differences but attenuate African American versus Hispanic differences when the 11-item PAC is used. The 9-item PAC scale that removes items 9 and 10 as listed in Table 1, therefore, might provide some advantages over the 11-item version for assessing positive experiences more uniformly across multiple race and ethnicity subgroups.

Table 2. Multiple Regression Analysis of the Summary Score of the Positive Aspects of Caregiving Measure as a Function of Demographic Factors

Predictor	Regression coefficient	Standard error	<i>t</i> (<i>df</i> = 636)	<i>p</i> Value
Intercept	26.17	2.60	10.05	<.0001
Age (years)	0.06	0.04	1.45	.1471
Gender (female = 1; male = 0)	-2.90	1.10	-2.64	.0086
Relationship (spouse = 1; nonspouse = 0)	-1.54	1.09	-1.42	.1565
Race				
African American vs. White	5.78	1.03	5.61	<.0001
Hispanic vs. White	7.35	1.02	7.22	<.0001

Table 3. Descriptive Statistics for the 11-item and a 9-item Positive Aspects of Caregiving Measure by Caregiver Race

	Cronbach's α	Total sample (<i>N</i> = 642)		African Americans (<i>N</i> = 211)		Hispanics (<i>N</i> = 212)		Whites (<i>N</i> = 219)	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
PAC-11	0.92	30.95	10.92	32.46 ^a	10.19	33.95 ^a	10.88	26.60 ^b	10.29
PAC-9	0.90	24.96	9.00	25.78 ^a	8.63	27.54 ^b	8.90	21.66 ^c	8.45

Notes. *SD* = standard deviation.

Race groups means that have different superscripts differ significantly from each other ($p < .05$).

Analyses of DIF are useful not only for identifying items that contribute potential sources of measurement bias, but also for exploring content areas that might tap culturally specific experiences. In our analyses, African Americans and Hispanics both reported more positive caregiving experiences on items that were not sensitive to cultural differences, but African Americans also responded disproportionately to two items that speak more directly to them—about feeling “enabled” to “appreciate life” and to “develop a positive attitude toward life.” These items may tap a long-standing tradition in the African American community that is cultivated in early life through racial socialization and maintained in adulthood and later life through spiritual and cultural beliefs that support positive attitudes about life, especially when facing adversity or overcoming hardships (McLoyd, Hill, & Dodge, 2005).

This adaptive sense of self has been further explored in studies describing how African Americans develop active coping styles and create attitudes of resilience, which are used to buffer negative life experiences such as caregiving to older relatives with dementia (Dilworth-Anderson et al. 2005; McCallum, Sorocco, & Fritsch, 2006; Merritt, McCallum, & Fritsch 2011). The present findings reinforce previous recommendations to conduct more research that captures cultural beliefs and attitudes that will further our understanding of the long-term effects of active coping, resilience, and positive attitudes on life among African American caregivers (Napoles, Chadiha, Eversley, & Moreno-John, 2010). We suggest that coping and having a positive attitude, especially while experiencing long-term adversity and stress, may create what James (1994) described as John Henryism—an orientation where work, determination, and a “can do” positive attitude provide the impetus to succeed yet might also put one at risk for adverse health consequences.

There are some limitations to this study’s findings that should be kept in mind. Although a multisite, national sample was obtained in REACH II, some sites enrolled more minority participants than others, and effects by race, therefore, might partly represent effects by geographic region. In addition, an inclusion criterion for dementia caregivers in REACH II is that they had to be experiencing at least some distress, so the results may not generalize to those caregivers who are experiencing no stress or strain from the caregiving experience.

In conclusion, we examined REACH II data and evaluated the items from the PAC scale to determine if certain items were differentially sensitive across race and other demographic groups. Several different criteria adapted from previous studies were used to examine DIF including statistical significance and practical meaningfulness. Although there was some evidence of significant and meaningful DIF for two items that were specifically sensitive to African Americans in comparison with both Whites and Hispanics, the 11-item PAC appears to be suitable for assessing individual differences in the positive experiences associated with caregiving across gender and relationship subgroups. Specific items were indicative of content domains that are particularly meaningful to African American caregivers and reinforce the need to conduct additional research that better captures the cultural aspects of caregiving for this group.

Funding

This work was supported by a grant from the National Institute on Aging (R03 AG046493).

References

- Belle, S. H., Burgio, L., Burns, R., Coon, D., Czaja, S. J., Gallagher-Thompson, D.,...Zhang, S. (2006). Enhancing the quality of life of dementia caregivers from different ethnic or racial groups: A randomized, controlled trial. *Annals of Internal Medicine*, 145, 727–738. doi:10.7326/0003-4819-145-10-200611210-00005
- Bollen K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley-Interscience. doi:10.1002/9781118619179
- Brown, R. M., & Brown, S. L. (2014). Informal caregiving: A reappraisal of effects on caregivers. *Social Issues and Policy Review*, 8, 74–102. doi:10.1111/sipr.12002
- Brown, S. L., Smith, D. M., Schulz, R., Kabeto, M. U., Ubel, P. A., Poulin, M.,... Langa, K. M. (2009). Caregiving behavior is associated with decreased mortality risk. *Psychological Science*, 20, 488–494. doi:10.1111/j.1467-9280.2009.02323.x
- Calderón, V., & Tennstedt, S. (1998). Ethnic differences in the expression of caregiver burden: Results of a qualitative study. *Journal of Gerontological Social Work*, 30, 159–178. doi:10.1300/j083v30n01_10
- Centers for Disease Control and Prevention. (2010). Caregiving: A public health priority. Retrieved from <http://www.cdc.gov/aging/caregiving>
- 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. *Journal of Clinical Epidemiology*, 53, 285–289. doi:10.1016/s0895-4356(99)00151-1
- Dilworth-Anderson, P., Brummett, B. H., Goodwin, P., Williams, S. W., Williams, R. B., & Siegler, I. C. (2005). Effect of race on cultural justifications for caregiving. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 60, S257–S262. doi:10.1093/geronb/60.5.s257
- Dilworth-Anderson, P., Pierre, G., & Hilliard, T. S. (2012). Social justice, health disparities, and culture in the care of the elderly. *The Journal of Law, Medicine & Ethics*, 40, 26–32. doi:10.1111/j.1748-720X.2012.00642.x
- Dilworth-Anderson, P., Williams, I. C., & Gibson, B. E. (2002). Issues of race, ethnicity, and culture in caregiving research: A 20-year review (1980–2000). *The Gerontologist*, 42, 237–272. doi:10.1093/geront/42.2.237
- Fredman, L., Cauley, J. A., Hochberg, M., Ensrud, K. E., & Doros, G. (2010). Mortality associated with caregiving, general stress, and osteoporotic fractures. *Journal of the American Geriatrics Society*, 58, 937–943. doi:10.1111/j.1532-5415.2010.02808.x
- Goins, R. T., Garrouette, E. M., Fox, S. L., Dee Geiger, S., & Manson, S. M. (2011). Theory and practice in participatory research: Lessons from the Native Elder Care Study. *The Gerontologist*, 51, 285–294. doi:10.1093/geront/gnq101
- Haley, W. E., Gitlin, L. N., Wisniewski, S. R., Mahoney, D. F., Coon, D. W., Winter, L.,...Ory, M. (2004). Well-being, appraisal, and coping in African-American and Caucasian dementia caregivers: Findings from the REACH study. *Aging & Mental Health*, 8, 316–329. doi:10.1080/13607860410001728998
- Haley, W. E., Roth, D. L., Coleton, M. I., Ford, G. R., West, C. A., Collins, R. P., & Isobe, T. L. (1996). Appraisal, coping, and social support as mediators of well-being in black and white family caregivers of patients with Alzheimer’s disease. *Journal of Consulting and Clinical Psychology*, 64, 121–129. doi:10.1037//0022-006x.64.1.121
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates. doi:10.4324/9780203357811
- James, S. A. (1994). John Henryism and the health of African-Americans. *Culture, Medicine and Psychiatry*, 18, 163–182. doi:10.1007/bf01379448
- Jones, R. N. (2006). Identification of measurement differences between English and Spanish language versions of the Mini-Mental State Examination. Detecting differential item functioning using MIMIC modeling. *Medical Care*, 44, S124–S133. doi:10.1097/01.mlr.0000245250.50114.0f
- Knight, B. G., & Sayegh, P. (2010). Cultural values and caregiving: The updated sociocultural stress and coping model. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 65B, 5–13. doi:10.1093/geronb/gbp096

- Lin, I. F., Fee, H. R., & Wu, H. S. (2012). Negative and positive caregiving experiences: A closer look at the intersection of gender and relationships. *Family Relations*, 61, 343–358. doi:10.1111/j.1741-3729.2011.00692.x
- McCallum, T. J., Sorocco, K. H., & Fritsch, T. (2006). Mental health and diurnal salivary cortisol patterns among African American and European American female dementia family caregivers. *The American Journal of Geriatric Psychiatry*, 14, 684–693.
- McLoyd, V.C., Hill, N.E., & Dodge, K.A. (Eds.). (2005). *African American family life: Ecological and cultural diversity*. New York, NY: Guilford. doi:10.5860/choice.43-6868
- Merritt, M. M., McCallum, T. J., & Fritsch, T. (2011). How much striving is too much? John Henryism active coping predicts worse daily cortisol responses for African American but not white female dementia family caregivers. *The American Journal of Geriatric Psychiatry*, 19, 451–460. doi:10.1097/JGP.0b013e3181eaffa4
- Muthén, L. K., & Muthén, B. O. (2013). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Napoles, A. M., Chadiha, L., Eversley, R., & Moreno-John, G. (2010). Reviews: Developing culturally sensitive dementia caregiver interventions: Are we there yet? *American Journal of Alzheimer's Disease and Other Dementias*, 25, 389–406. doi:10.1177/1533317510370957
- O'Reilly, D., Connolly, S., Rosato, M., & Patterson, C. (2008). Is caring associated with an increased risk of mortality? A longitudinal study. *Social Science & Medicine*, 67, 1282–1290. doi:10.1016/j.socscimed.2008.06.025
- Ory, M. G., Hoffman, R. R., Yee, J. L., Tennstedt, S., & Schulz, R. (1999). Prevalence and impact of caregiving: A detailed comparison between dementia and nondementia caregivers. *The Gerontologist*, 39, 177–185. doi:10.1093/geront/39.2.177
- Picot, S. J., Debanne, S. M., Namazi, K. H., & Wykle, M. L. (1997). Religiosity and perceived rewards of black and white caregivers. *The Gerontologist*, 37, 89–101. doi:10.1093/geront/37.1.89
- Pinquart, M., & Sörensen, S. (2003). Differences between caregivers and non-caregivers in psychological health and physical health: A meta-analysis. *Psychology and Aging*, 18, 250–267. doi:10.1037/0882-7974.18.2.250
- Pinquart, M., & Sörensen, S. (2005). Ethnic differences in stressors, resources, and psychological outcomes of family caregiving: A meta-analysis. *The Gerontologist*, 45, 90–106. doi:10.1093/geront/45.1.90
- Pinquart, M., & Sörensen, S. (2006). Gender differences in caregiver stressors, social resources, and health: An updated meta-analysis. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 61, P33–P45. doi:10.1093/geronb/61.1.p33
- Pinquart, S., & Sörensen, S. (2011). Spouses, adult children, and children-in-law as caregivers of older adults: A meta-analytic comparison. *Psychology and Aging*, 26(1), 1–14. doi:10.1037/a0021863
- Pruchno, R., & Gitlin, L. N. (2012). Family caregiving in late life: Shifting paradigms. In R. Blieszner & V. H. Bedford (Eds.), *Handbook of families and aging* (pp. 515–541). Westport, CT: Praeger Publishers. doi:10.5860/choice.50-4731
- Roff, L. L., Burgio, L. D., Gitlin, L., Nichols, L., Chaplin, W., & Hardin, J. M. (2004). Positive aspects of Alzheimer's caregiving: The role of race. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 59, P185–P190. doi:10.1093/geronb/59.4.p185
- Roth, D. L., Fredman, L., & Haley, W. E. (2015). Informal caregiving and its impact on health: A reappraisal from population-based studies. *The Gerontologist*, 55, 309–319. doi:10.1093/geront/gnu177
- Roth, D. L., Haley, W. E., Hovater, M., Perkins, M., Wadley, V. G., & Judd, S. (2013). Family caregiving and all-cause mortality: Findings from a population-based propensity-matched analysis. *American Journal of Epidemiology*, 178, 1571–1578. doi:10.1093/aje/kwt225
- Roth, D. L., Haley, W. E., Owen, J. E., Clay, O. J., & Goode, K. T. (2001). Latent growth models of the longitudinal effects of dementia caregiving: A comparison of African American and White family caregivers. *Psychology and Aging*, 16, 427–436. doi:10.1037/0882-7974.16.3.427
- Roth, D. L., Perkins, M., Wadley, V. G., Temple, E. M., & Haley, W. E. (2009). Family caregiving and emotional strain: Associations with quality of life in a large national sample of middle-aged and older adults. *Quality of Life Research*, 18, 679–688. doi:10.1007/s11136-009-9482-2
- Schulz, R., & Martire, L. M. (2004). Family caregiving of persons with dementia: Prevalence, health effects, and support strategies. *The American Journal of Geriatric Psychiatry*, 12, 240–249. doi:10.1097/00019442-200405000-00002
- Stewart, A. L., & Nápoles-Springer, A. M. (2003). Advancing health disparities research: Can we afford to ignore measurement issues? *Medical Care*, 41, 1207–1220. doi:10.1097/01.MLR.0000093420.27745.48
- Tarlow, B. J., Wisniewski, S. R., Belle, S. H., Rubert, M., Ory, M. G., & Gallagher-Thompson, D. (2004). Positive aspects of caregiving: Contributions of the REACH project to the development of a new measure for Alzheimer's caregiving. *Research on Aging*, 26, 429–453. doi:10.1177/0164027504264493
- Woods, C. M. (2009). Empirical selection of anchors for tests of differential item functioning. *Applied Psychological Measurement*, 33, 42–57. doi:10.1177/0146621607314044
- Yang, F. M., & Jones, R. N. (2007). Center for Epidemiologic Studies-Depression Scale (CES-D) item response bias found with Mantel-Haenszel method was successfully replicated using latent variable modeling. *Journal of Clinical Epidemiology*, 60, 1195–1200. doi:10.1016/j.jclinepi.2007.02.008
- Yee, J. L., & Schulz, R. (2000). Gender differences in psychiatric morbidity among family caregivers: A review and analysis. *The Gerontologist*, 40, 147–164. doi:10.1093/geront/40.2.147
- Zarit, S. H. (2012). Positive aspects of caregiving: More than looking on the bright side. *Aging & Mental Health*, 16, 673–674. doi:10.1080/13607863.2012.692768