



Use of CES-D among 56–66 year old people of Dutch, Moroccan and Turkish origin: Measurement invariance and mean differences between the groups

Silvia Klokgieters^{1,2} · Lidwine Mokkink^{2,3} · Henrike Galenkamp^{2,4} · Aartjan Beekman^{2,5} · Hannie Comijs^{2,5}

Published online: 8 September 2018
© The Author(s) 2018

Abstract

When assessing depressive symptoms across ethnic populations it is important to ensure that items from a questionnaire are valued and interpreted similarly across groups. We aimed to examine measurement (in)variance of the Center for Epidemiological Studies Depression Scale (CES-D) among people of Dutch, Moroccan and Turkish origin in the Netherlands and to compare the level of depressive symptoms across these three groups. Data were used from the Longitudinal Aging Study Amsterdam, including 269 people from Turkish, 209 from Moroccan and 618 from Dutch origin (aged 55–65 years). A multi-group confirmatory factor analysis (MGCFA) was performed to test measurement invariance of the four-factor CES-D across the three cohorts. To compare scores across ethnic groups, we performed ANCOVA. The four subscales of the CES-D (depressed affect, positive affect, somatic symptoms, and interpersonal problems) appeared measurement invariant in people of Dutch, Moroccan and Turkish origin. Turkish and Moroccan participants reported more depressive symptoms on all four domains. The four subscales of the CES-D measure the same constructs in people of Dutch, Moroccan and Turkish origin. Higher levels of depressive symptoms in the migrants groups are therefore not due to measurement invariance, but point to increased mental health problems in these groups.

Keywords Depression · Immigrants · Measurement invariance · CES-D · Factor structure

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12144-018-9977-5>) contains supplementary material, which is available to authorized users.

✉ Silvia Klokgieters
s.s.klokgieters@vu.nl

¹ Department of Sociology, Faculty of Social Sciences, Vrije Universiteit, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

² Department of Psychiatry, GGZinGeest - VU University Medical Center, Amsterdam, The Netherlands

³ Department of Epidemiology and Biostatistics and Amsterdam Public Health Research Institute, VU University Medical Center, Amsterdam, The Netherlands

⁴ Department of Public Health and Amsterdam Public Health Research Institute, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

⁵ Department Psychiatry and Amsterdam Public Health Research Institute, VU University Medical Center, Amsterdam, The Netherlands

Introduction

In the Netherlands the number of Turkish and Moroccan older adults is increasing. Most of these immigrants are first generation immigrants, and came to the Netherlands in the 1960s and 1970s for work or family reunification. Immigrants are considered vulnerable groups in society due to an accumulation of risk factors including language barriers (Pot et al. 2018; Suurmond et al. 2016), acculturative stress (Denktaş 2011), and experiences of discrimination and segregation (Pettigrew et al. 1997). Moreover, these risk factors are usually experienced on top of traditional risk factors for depression, including, being female (van de Velde et al. 2010), poverty (Bruce and Hoff 1994) and physical health deterioration (Cole and Dendukuri 2003). Because these characteristics are known to be linked to mental health problems (Van Der Zwaan and Tolsma 2013), questions arise about the mental health of these immigrants. Previous research showed that depressive symptoms as measured with the Center for Epidemiological Studies Depression Scale (CES-D) were more prevalent among older immigrants as compared to older adults from the Dutch origin (Van Der Wurff et al. 2004). The prevalence of self-reported depressive symptoms was 33.6% for

people from Moroccan origin and 61.5% for people from Turkish origin, whereas the prevalence of depressive symptoms in the people from Dutch origin was 14.5% (Van Der Wurff et al. 2004). However, it is questionable whether the scores on the CES-D reflect real differences in depressive symptoms among these cultural groups. Depression might be expressed in different ways in different cultures, and items of the CES-D might be valued and interpreted differently across cultural groups (Guarnaccia et al. 1989; Kim et al. 2011; Kirmayer et al. 2017). In non-Western cultures the attitude towards mental illness seems more stigmatized than in Western cultures and the expression of depressed mood might be more devalued (Zhang et al. 2011). In some cultures there are no words for some of the key symptoms of depression, which illustrates that there may be fundamental cultural divides with regard to the experience and expression of depressive symptoms. As a result, migrants might emphasize somatic symptoms, such as sleep problems, having difficulty to get started, or feeling tired more than the typically psychological symptoms, such as feeling down, self-deprecation, and suicidal ideation (Spijker et al. 2004). Given the importance of cross-cultural comparisons of depression, it is important to have access to instruments with minimal measurement variance (MI) (Van Den Berg and Lance 2000).

The CES-D (Radloff 1977) has been translated in many different languages and is widely used across the world. Radloff (1977) examined the factor structure of the CES-D and identified four factors, including depressed affect, positive affect, somatic symptoms/retarded activity and interpersonal problems. This four-factor structure of the CES-D has been replicated many times (Kim et al. 2011; Chin et al. 2015). However, several other studies found a three-factor or even a two-factor structure (see Carleton et al. 2013). In addition, bifactor or higher order models were evaluated and were found to be invariant across different cultural groups (Gomez and McLaren 2015; Yang et al. 2009). In the Netherlands, Beekman et al. (1997a) confirmed the four-factor structure of the CES-D. More recently a study compared the factor structure of a native Dutch sample with that from a Chinese sample, and showed that in both samples, the four-factor model resulted in the best fit. They concluded that the four dimensions (somatic symptoms, depressed affect, positive affect, and interpersonal problems) of the CES-D seemed to be the most informative when assessing depressive symptoms in older adults compared to the other factor models (Zhang et al. 2011). In a meta-analysis Kim et al. (2011) replicated the original four-factor structure in four out of five ethnic groups, and concluded that the original four-factor structure may not be the best fit for all ethnic groups.

The four-factor model was the most informative when assessing depressive symptoms in older adults compared to the other factor models (Zhang et al. 2011), but one previous study showed violations of MI across ethnic groups (Kim et al. 2011). In addition, the sum score on the CES-D is widely used in research and practice. Therefore we

considered both the unidimensional and the four-factor model for measurement invariance testing.

When comparing different cultural groups, it is important that the CES-D measures depression the same way across cultural groups. In other words the CES-D has to be measurement invariant. Measurement variance may be caused by differences in the interpretation of items and by differences in response style due to cultural differences in conceptualization, meaning, and symptom expression of depression (Kim et al. 2011). For instance Zhang et al. (2011) found several items of the CES-D (for instance ‘feeling depressed’, ‘feeling fearful’, and ‘feeling good’) to be measurement non-invariant in a Chinese and Dutch sample of older adults. In addition, for different measures of depression, Smits and colleagues (Smits et al. 2005) found that questions about depressive symptoms were prone to interpreting issues among Turkish and Moroccan immigrants. When interviewing these immigrants, they found that depressive symptoms are often subject of shame among individuals or families and therefore more likely to be concealed. Physical symptoms, by contrast were more easily reported. Given the potential variety of interpretations, it is important to study measurement invariance before comparing groups with different cultural backgrounds.

The first aim of the present study was to examine measurement invariance of the CES-D regarding ethnicity. When we show that the CES-D is measurement invariant, we aim to compare the scores on the CES-D between people of Dutch, Moroccan and Turkish origin. For comparing the CES-D (subscale) scores across groups, we aim to eliminate the potential effect of other important risk factors for depression. More specifically, we included three potential risk factors, namely sex, income level and physical limitations. Each of these risk factors are known to be associated with depressive symptoms and ethnicity (Bruce and Hoff 1994; Beekman et al. 1997b; Cole and Dendukuri 2003; Van Der Wurff et al. 2004; Schellingerhout 2004; Reijneveld et al. 2007; Van De Velde et al. 2010; Denктаş 2011).

Methods

Sample

The present study included two samples collected in the context of LASA. The Longitudinal Aging Study Amsterdam (LASA) investigates determinants and consequences of ageing in social, cognitive, physical and emotional domains of functioning (Huisman et al. 2011). The people from Dutch origin ($n = 1023$) were included in 2012–2013. Respondents were drawn from the population registers of 11 Dutch municipalities that differ with regard to the degree of urbanization in the Netherlands. The respondents were born between 1948 and 1957 (cooperation rate 63%). The second sample included 269 people from Turkish and 209 from Moroccan origin, also

born between 1948 and 1957 (cooperation rate Total: 45%, Turkish immigrants: 50%, Moroccan immigrants: 40%) and was collected in 2013–2014. LASA focussed on these immigrant groups because they comprise the largest groups of labour immigrants that have settled in the Netherlands. There are a number of reasons why the people of Moroccan and Turkish origin are considered to be at risk of more rapid physical deterioration, social loneliness and depression (Schellingerhout 2004). Because people of Moroccan and Turkish origin predominantly live in the larger cities in the Netherlands, data collection took place by drawing from the registers of fifteen Dutch cities with a population size between 85 and 805 thousand.

In order to enhance the comparability between, we selected only those Dutch respondents who lived in urbanized areas of the Netherlands. To this end 317 native Dutch respondents to LASA were excluded from the original sample. Furthermore, 88 Dutch respondents were excluded because they were not born in the Netherlands. Similarly, 5 immigrants were excluded because they were not born in Turkey or Morocco. In total 618 native Dutch, 267 Turkish and 209 Moroccan respondents were included in the analysis.

The CES-D is part of the face-to-face interview. The interviews were conducted by trained interviewers, who were of the same ethnic background and gender as the respondent. If needed, translated interviews in Turkish, Tarafit or Moroccan dialect were available for the immigrant sample. These Turkish and Arabic versions were previously translated according to standard procedures (Spijker et al. 2004). First, all questions of the CES-D were translated into Turkish language and Moroccan Arabic dialect by certified translators. Second, all translated questions were translated back into Dutch by other certified translators. Third, experts reconciled both translations and decided upon a final translation of the CES-D. Afterwards the study conducted an analysis to validate the translation of the CES-D. Translations in Tarafit and all other questions in the interview, where done by professional translators and were subsequently evaluated through several pilot interviews.

Measurements

The CES-D was used to assess self-reported depressive symptoms, and consists of 20 items that are scored on a 4-point Likert scale. Higher scores on the CES-D indicate more depressive symptoms. The four original subscales of the CES-D include depressed affect (items 3, 6, 9, 10, 14, 17, 18), positive affect (items 4, 8, 12, 16), somatic symptoms (items 1, 2, 5, 7, 11, 13, 20), and interpersonal problems (items 15, 19) (see [supplementary table](#)).

Covariates

Country of origin (Dutch, Turkish or Moroccan) is defined according to the country of birth of the participants. Income level was based on the level of income in the household, and was dichotomized at the Dutch poverty line; below € 1425 for respondents with partner and below €1040 for respondents without partner. The number of physical limitations was based on the respondent's ability to perform seven daily tasks: going up and down a staircase, using own or public transportation, cutting toenails, dressing and undressing, sitting down and standing up from a chair, walking outside, and taking a shower or bathe. Five response options were included: no difficulty, some difficulty, much difficulty, only with help, and unable, coded as 0 to 4. Sum scores are calculated with range from 0 to 28. Cronbach's alpha was .81, .87, and .79 for Dutch, Turkish, and Moroccan respondents, respectively.

Statistical Analysis

To examine measurement invariance of the CES-D, first confirmatory factor analysis (MGCFA) was performed using polychoric correlations with a weighted least squares algorithm with mean and variance adjustment (WLSMV), using MPLUS version 7, to assess how well the data fit the competing models. First, the unidimensional model was tested in each cohort (i.e. the Dutch, Turkish and Moroccan group) because in many studies the CES-D is used as a unidimensional scale. Next, the original four-factor model was tested. Model fit was evaluated using the Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Standardized Root Mean Square Residual (SRMR). Generally, CFI and TLI >0.95, and SRMR <0.08 are considered as good (Hu and Bentler 1999). RMSEA between 0.05 and 0.08 is considered reasonable, and RMSEA ≥0.1 suggests poor fit (University of Connecticut 2014). Local independence was investigated by inspecting the residual correlations between items obtained from the confirmatory factor analysis. In addition, the internal consistency was computed for each unidimensional (sub)scale.

After having found one model that showed reasonable fit in each ethnic group, we incorporated this model into a multi-group MGCFA. Four nested models were tested, i.e. configural invariance, metric invariance, scalar invariance, and strict invariance, by comparing a more constrained model to the previous model. In the configural invariance model the same factor structure is imposed on the three groups, indicating that the clustering of items and the factors that they represent are similar across groups (Gregorich 2006). In the metric invariance model the factor loadings are constrained to be equal, implying that the magnitude of the loadings is similar across the groups (Gregorich 2006). In the scalar invariance model, both the factor

loadings, and the item thresholds are constrained across groups, indicating that item responses are not systematically higher or lower in one group compared to the other group(s) (Gregorich 2006). In strict invariance models, in addition to constraining the factor loadings and item thresholds, also the residual variances, or error terms, of each item are constrained across the groups (Gregorich 2006).

Following recommendations by Chen (2007) for comparing two nested models, cut-off values of $\Delta CFI < 0.01$ and $\Delta RMSEA < 0.015$ will be used for testing metric invariance, scalar invariance, as well as strict invariance. In the present study, models will be considered acceptable on condition that both indices met the above criteria.

Finally, in order to compare scores across ethnic groups, we performed univariate Analysis of Covariance (ANCOVA), adjusting for the effect of gender, income and physical limitations. The analysis for each of the four sub-scales was done separately. In addition, pairwise comparisons were performed through post-hoc testing.

Results

Sample Characteristics

Table 1 summarizes the sample characteristics including age, sex, household income level, and physical functioning for all origin groups separately. People from Turkish and Moroccan origin had lower income levels than people from Dutch origin. Turkish immigrants had the lowest physical functioning levels followed by Moroccan immigrants and lastly by older adults from Dutch origin.

Confirmatory Factor Analysis and Internal Consistency

First, we determined goodness-of-fit indices of the two models, i.e. the unidimensional model and the original four-factor model (Table 2) for each origin group separately. In all

three groups the four-factor model showed the best fit. The goodness-of-fit indices for the unidimensional model were largely insufficient. Table 3 depicts the Cronbach's alpha of all four domains of the CES-D for each group separately. The internal consistency of the somatic symptoms, the depressed affect and the positive affect subscales is sufficient in all three groups. The internal consistency of interpersonal contacts sub-scale is insufficient, but this sub-scale consist of only two items.

Measurement Invariance Analysis

Table 4 shows the results from the multi-group CFA. Adding constraints for equal factor structure, equal factor loadings, equal item thresholds and equal residual variances, respectively (i.e. configural invariance, metric invariance, strong invariance and strict invariance) did not lead to a reduced model fit when testing measurement invariance for the group variable. While the RMSEA value for country of origin increased after adding constraints in the metric and strong model (i.e. $\Delta RMSEA = 0.007$, $\Delta RMSEA = 0.000$) it did not fall above the critical value of 0.08. Moreover, while the RMSEA for country of origin indicated acceptable fit the CFI indicated good fit. The final strict model had a RMSEA of 0.056, a CFI of 0.958 and a TLI of 0.960, indicating reasonable and good fit, respectively. Measurement invariance regarding country of origin was confirmed for the four factors of the CES-D.

Comparison of CES-D Scores between the Migrant and Dutch Groups

After adjusting for income, sex and physical limitations, descriptive statistics indicate that native Dutch and immigrant groups differ on all four domains of de CES-D (Table 5). On all four domains, average scores of people from Turkish and Moroccan origin indicated more symptoms than native Dutch. Only for positive affect, Turkish and Moroccan groups differed from each other. Turkish immigrants had lower levels of positive affect than Moroccan immigrants.

Table 1 Demographic characteristics of the groups

	All N = 1091	Turkish immigrants N = 267	Moroccan immigrants N = 209	Native Dutch N = 618	p value [‡]
Mean age in years (SD)	60.9 (3.0)	60.7 (3.1)	61.1 (2.9)	60.5 (2.9)	*
Female (%)	528 (48)	120 (45)	81 (39)	327 (53)	**
Household income (SD)					
Poverty line [†]	338 (31)	135 (51)	135 (69)	68 (11)	***
Physical functioning (SD)	2.7 (4.6)	5.9 (6.2)	3.4 (4.5)	1.3 (2.8)	***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[†] Below €1425 for respondents with partner and below €1040 for respondents without partner

[‡] Based on ANOVA, Chi-square test, Kruskal-Wallis tests, as appropriate

Table 2 Results of the confirmatory factor analysis for the unidimensional and four-factor model of the CES-D in the three groups

	Unidimensional model			Four factor model [†]		
	Dutch (n = 618)	Turkish (n = 260)	Moroccan (n = 206)	Dutch (n = 618)	Turkish (n = 260)	Moroccan (n = 206)
Chi-square (df)	831.752 (170)	500.383 (170)	437.105 (170)	475.184 (165)	261.083 (165)	220.837 (165)
p value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0024
RMSEA	0.079	0.086	0.071	0.055	0.047	0.041
CFI	0.924	0.924	0.957	0.964	0.978	0.986
TLI	0.915	0.915	0.952	0.959	0.975	0.984

RMSEA, Root Mean Square Error of Approximation; CFI, Comparative Fit Index; TLI, Tucker-Lewis Index

[†] Somatic symptoms includes items 1, 2, 5, 7, 11, 13, 20; depressed affect includes items 3, 6, 9, 10, 14, 17, 18; positive affect includes items 4, 8, 12, 16; interpersonal problems includes items 15, 19

Discussion

The four subscales of the CES-D (depressed affect, positive affect, somatic symptoms, and interpersonal problems) are measurement invariant in older adults from Dutch and Turkish and Moroccan origin. This implies that the four subscales of the CES-D measure the same constructs in the three groups. Therefore we are able to make meaningful comparisons between the three groups. It has to be noted that the internal reliability of the interpersonal subscale was insufficient, but this subscale only contains two items. Nevertheless, we should be careful in interpreting this subscale.

When comparing the three groups our results showed that the Turkish and Moroccan immigrants reported more depressive symptoms on all four domains, compared to adults of Dutch origin. These results are in line with prior studies (Van Der Wurff et al. 2004; Schrier et al. 2010), using the CES-D as a unidimensional scale (Van Der Wurff et al. 2004) or the Symptom Checklist-90-revised (SCL-90-R), which did not examine measurement invariance of the scales. Except for positive affect, people from Turkish and Moroccan origin scored similarly on the subscales. The Turkish group had the lowest levels of positive affect of all groups. Although we expected more depressive symptoms in the immigrant groups as compared to the Dutch origin group, we expected them to score especially high on the somatic subscale, and lower on the depressed affect subscale as was suggested by Acartürk and colleagues (Acartürk et al. 2011).

The MGCFA showed that the original four-factor structure of the CES-D had the best fit in all three groups. The goodness of fit of an unidimensional model was insufficient in the three groups, meaning that the use of a total score is not grounded in our study sample. Our results show the importance of studying the factor structure and measurement invariance before comparing groups with a different cultural background. In addition to the unidimensional and 4-factor model, we fitted a bifactor model because Chen and colleagues (Chen et al. 2006) argued that there are mathematical advantages of the bi-factor model, which include estimating fewer parameters and reduced model complexity. However, in the Turkish sample, this model could not be identified, and two items (i.e. item 7; I felt that everything I did was an effort and item 18; I felt sad) showed extremely high factor loadings both for the general factor as well as for the group factor compared to other factor loadings. It might be that for Turkish immigrants item 7 signifies physical health problems, which would explain why they were more likely to endorse item 7. In the Dutch sample, the model could be identified, but item 18 also showed proportionately high factor loadings compared to the other factor loadings. The fact that item 18 showed high factor loadings for both Dutch and the Turkish group might indicate similarity rather than difference between the structure of the CES-D in both origin groups. Because of these difficulties, we decided not to use this model.

Our study has several strengths and limitations. Strength is that we included a group of older adults from Turkish,

Table 3 Cronbach’s alpha’s for four CES-D factors of Turkish and Moroccan immigrants and native Dutch

		Turkish immigrants α	Moroccan immigrants α	Native Dutch α
Factor 1	Somatic symptoms	0.79	0.80	0.75
Factor 2	Depressed affect	0.86	0.90	0.86
Factor 3	Positive affect	0.79	0.73	0.78
Factor 4	Interpersonal problems	0.58	0.61	0.56

Table 4 Goodness-of-fit indices for models tested in the native Dutch, and Turkish and Moroccan immigrants for the four-factor model of the CES-D

	Configural	Metric	Strong	Strict
Chi-square (df)	963.927 (497)	1168.049 (523)	1242.606 (563)	1297.145 (603)
<i>p</i> value	0.0000	0.0000	0.0000	0.0000
RMSEA	0.051	0.058	0.058	0.056
CFI	0.972	0.961	0.959	0.958
TLI	0.968	0.958	0.959	0.960

RMSEA, Root Mean Square Error of Approximation; CFI, Comparative Fit Index; TLI, Tucker-Lewis Index

Moroccan and Dutch from urban areas in the The Netherlands, where most of the immigrant groups live. Also we were able to include persons who did not speak Dutch or who were illiterate by using interviewers that spoke the language of the respondent, and by using questionnaires translated by professional translators available in Moroccan dialects and Turkish language that were verbally executed. A limitation is that among the immigrant groups the cooperation rate was low. This seems a common problem, across Europe studies report challenges with cooperation rates among older migrants (Lipson and Meleis 1989), as well as in the Netherlands (Reijneveld et al. 2007). Little is known about the characteristics of the persons who do not participate in this study. This makes it difficult to estimate how this has affected our results. The use of only self-report measures may also be considered a limitation. Some ethnic minority groups may be more inclined to give socially desirable answers (Reijneveld 1998; Reijneveld and Stronks 1999), which may have resulted in underreporting depressive symptoms. It is suggested that underreporting could be based on socio-economic factors, health problems, trust issues and time constraints (Spijker et al. 2004).

It is worrying that the immigrant groups report much more depressive symptoms than their native Dutch peers. A previous Dutch study among adult immigrant groups showed already in 2008 higher prevalence rates for depressive disorders among Turkish and Moroccan immigrants as compared to their native Dutch peers (De Wit et al. 2008). Nonetheless, it has been shown that Turkish and Moroccan

immigrants make relatively little use of mental health care facilities (Kamperman et al. 2007). It is therefore questionable whether these symptoms are adequately recognised by health care practitioners or by the Turkish and Moroccan immigrants themselves. One reason might be that language barriers occur between health practitioners and immigrants (Fransen et al. 2013; Rademakers 2014). Language barriers go beyond the ability to speak Dutch but include cultural differences in the interpretation of mental illness (i.e. of the taboos resting on mental illness) as well as an individual's own knowledge and understanding of mental illness (Fransen et al. 2013). Based on the results of this study, it is important that health professionals are aware of the high prevalence rates of depression among older immigrants and that the individual depressive symptoms are systematically asked for in these migrants when they are suspect for depression. In addition, the finding that Turkish immigrants might display lower positive affect than Moroccan immigrants might be an important lead for health professionals to pay extra attention to.

In conclusion, the four subscales of the CES-D to measure depressive symptoms were measurement invariant in older adults of Dutch, Turkish and Moroccan origin. Given this finding, we were able to compare the prevalence of depressive symptoms across these three groups and found that the Turkish and Moroccan immigrants aged 55–65 years, reported more depressive symptoms on all four domains, compared to the Dutch. Health care professionals should be sensitive for mental health problems among these older immigrant groups.

Table 5 Scores on the four CES-D factors of Turkish and Moroccan immigrants and native Dutch group with confidence intervals adjusted for income, sex and physical functioning

	Subscale (range)	Turkish immigrants <i>N</i> = 267	Moroccan immigrants <i>N</i> = 209	Native Dutch <i>N</i> = 618	<i>p</i>
Factor 1	Somatic symptoms (0–14)	4.8 (4.4–5.2) [§]	5.3 (4.9–5.8) [§]	3.5 (3.0–4.0) ^{†‡}	***
Factor 2	Depressed affect (0–14)	4.1 (3.7–4.5) [§]	3.7 (3.3–4.2) [§]	2.2 (1.7–2.6) ^{†‡}	***
Factor 3	Positive affect (0–12)	5.8 (5.5–6.2) ^{‡,§}	7.0 (6.6–7.4) ^{†,§}	8.8 (7.9–8.7) ^{†‡}	***
Factor 4	Interpersonal problems (0–2)	0.6 (0.6–0.7) [§]	0.6 (0.5–0.7) [§]	0.3 (0.2–0.4) ^{†‡}	***

Note: Pairwise comparisons of proportions are post-hoc ANCOVA using Bonferroni correction

p* < 0.05, *p* < 0.01, ****p* < 0.001,

† = different from Turkish immigrants, ‡ = different from Moroccan immigrants, § = different from native Dutch

Funding The Longitudinal Aging Study Amsterdam is supported by a grant from the Netherlands Ministry of Health Welfare and Sports, Directorate of Long-Term Care. The data collection among the migrant groups was financially supported by the Netherlands Organization for Scientific Research (NWO) in the framework of the project “New Cohorts of young old in the twenty-first century” (file number 480–10-014).

Compliance with Ethical Standards

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Acartürk, C. Z., Nierkens, V., Agyemang, C., & Stronks, K. (2011). Depressive symptoms and smoking among young Turkish and Moroccan ethnic minority groups in the Netherlands: A cross-sectional study. *Substance Abuse Treatment, Prevention, and Policy*, 6(1), 5. <https://doi.org/10.1186/1747-597x-6-5>.
- Beekman, A. T. F., Deeg, D. J. H., Schoevers, R. A., Smit, J. H., Hooijer, C., & Tilburg, W. V. (1997a). Depressie bij ouderen in de Nederlandse bevolking: Een onderzoek naar de prevalentie en risicofactoren [depression among older adults in the Dutch population: A study in the prevalence of riskfactors]. *Tijdschrift voor Psychiatrie*, 39(4), 294–308.
- Beekman, A. T., Deeg, D. J., Van Limbeek, J., Braam, A. W., De Vries, M. Z., & Van Tilburg, W. (1997b). Criterion validity of the center for epidemiologic studies depression scale (CES-D): Results from a community-based sample of older subjects in The Netherlands. *Psychological Medicine*, 27(1), 231–235. <https://doi.org/10.1017/s0033291796003510>.
- Bruce, M. L., & Hoff, R. A. (1994). Social and physical health risk factors for first-onset major depressive disorder in a community sample. *Social Psychiatry and Psychiatric Epidemiology*, 29(4), 165–171. <https://doi.org/10.1136/ebmh.5.4.125>.
- Carleton, R. N., Thibodeau, M. A., Teale, M. J., Welch, P. G., Abrams, M. P., Robinson, T., & Asmundson, G. J. (2013). The center for epidemiologic studies depression scale: A review with a theoretical and empirical examination of item content and factor structure. *PLoS One*, 8(3), e58067. <https://doi.org/10.1371/journal.pone.0058067>.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>.
- Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, 41(2), 189–225. https://doi.org/10.1207/s15327906mbr4102_5.
- Chin, W. Y., Choi, E. P., Chan, K. T., & Wong, C. K. (2015). The psychometric properties of the center for epidemiologic studies depression scale in Chinese primary care patients: Factor structure, construct validity, reliability, sensitivity and responsiveness. *PLoS One*, 10(8), e0135131. <https://doi.org/10.1371/journal.pone.0135131>.
- Cole, M. G., & Dendukuri, N. (2003). Risk factors for depression among elderly community subjects: A systematic review and meta-analysis. *American Journal of Psychiatry*, 160(6), 1147–1156. <https://doi.org/10.1176/appi.ajp.160.6.1147>.
- De Wit, M. A., Tuinebreijer, W. C., Dekker, J., Beekman, A. J., Gorissen, W. H., Schrier, A. C., Penninx, B. W., Komproe, I. H., & Verhooff, A. P. (2008). Depressive and anxiety disorders in different ethnic groups: A population based study among native Dutch, and Turkish, Moroccan and Surinamese migrants in Amsterdam. *Social Psychiatry and Psychiatric Epidemiology*, 43(11), 905–912. <https://doi.org/10.1007/s00127-008-0382-5>.
- Denktaş, S. (2011). Health and health care use of elderly immigrants in the Netherlands: A comparative study. Dissertation, Erasmus University Medical Center, Rotterdam.
- Fransen, M., Harris, C., & Essink-Bot, M. L. (2013). Beperkte gezondheidsvaardigheden bij patiënten van allochtone herkomst: Alleen een tolk inzetten en is meestal niet genoeg [Limited health skills among patients of immigrant background: Using a translator only is usually not enough]. *Nederlands Tijdschrift Geneeskunde*, 157, A5581.
- Gomez, R., & McLaren, S. (2015). The center for epidemiologic studies depression scale: Support for a bifactor model with a dominant general factor and a specific factor for positive affect. *Assessment*, 22(3), 351–360. <https://doi.org/10.1177/1073191114545357>.
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, 44(3), S78–S94. <https://doi.org/10.1097/01.mlr.0000245454.12228.8f>.
- Guarnaccia, P. J., Angel, R., & Worobey, J. L. (1989). The factor structure of the CES-D in the Hispanic health and nutrition examination survey: The influences of ethnicity, gender and language. *Social Science and Medicine*, 29(1), 85–94. [https://doi.org/10.1016/0277-9536\(89\)90131-7](https://doi.org/10.1016/0277-9536(89)90131-7).
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>.
- Huisman, M., Poppelaars, J., Van Der Horst, M., Beekman, A. T. F., Brug, J., & Deeg, D. J. H. (2011). Cohort profile: The longitudinal aging study Amsterdam. *International Journal of Epidemiology*, 40(4), 868–876. <https://doi.org/10.1093/ije/dyq219>.
- Kamperman, M., Komproe, I. H., & de Jong, J. T. V. M. (2007). Migrant mental health: A model for indicators of mental health and health care consumption. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, 26(1), 96–104. <https://doi.org/10.1037/0278-6133.26.1.96>.
- Kim, G., Decoster, J., Huang, C. H., & Chiriboga, D. A. (2011). Race/ethnicity and the factor structure of the center for epidemiologic studies depression scale: A meta-analysis. *Cultural Diversity and Ethnic Minority Psychology*, 17(4), 381–396. <https://doi.org/10.1037/a0025434>.
- Kirmayer, L. J., Gomez-Carrillo, A., & Veissière, S. (2017). Culture and depression in global mental health: An ecosocial approach to the phenomenology of psychiatric disorders. *Social Science and Medicine*, 183, 163–168. <https://doi.org/10.1016/j.socscimed.2017.04.034>.

- Lipson, J. G., & Meleis, A. I. (1989). Methodological issues in research with immigrants. *Medical Anthropology*, 12(1), 103–115. <https://doi.org/10.1080/01459740.1989.9966014>.
- Pettigrew, T. F., Jackson, J. S., Brika, J. B., Lemaine, G., Meertens, R. W., Wagner, U., Zick, A. (1997). Outgroup prejudice in Western Europe. *European Review of Social Psychology*, 8, 241–273. <https://doi.org/10.1080/14792779843000009>.
- Pot, A., Keijzer, M., & De Bot, K. (2018). The language barrier in migrant aging. *International Journal of Bilingual Education and Bilingualism*, 1–19. <https://doi.org/10.1080/13670050.2018.1435627>.
- Rademakers, J. (2014). Kennissynthese. Gezondheidsvaardigheden: niet voor iedereen vanzelfsprekend [knowledge synthesis. Health skills: not self-evident for everyone]. Utrecht, NIVEL.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1(3), 385–401. <https://doi.org/10.1177/014662167700100306>.
- Reijneveld, S. A. (1998). Reported health, lifestyles, and use of health care of first generation immigrants in the Netherlands: Do socio-economic factors explain their adverse position? *Journal of Epidemiology and Community Health*, 52(5), 298–304. <https://doi.org/10.1136/jech.52.5.298>.
- Reijneveld, S. A., & Stronks, K. (1999). The impact of response bias on estimates of health care utilization in a metropolitan area: The use of administrative data. *International Journal of Epidemiology*, 28(6), 1134–1140. <https://doi.org/10.1093/ije/28.6.1134>.
- Reijneveld, S. A., Spijker, J., & Dijkshoorn, H. (2007). Katz' ADL index assessed functional performance of Turkish, Moroccan, and Dutch elderly. *Journal of Clinical Epidemiology*, 60(4), 382–388. <https://doi.org/10.1016/j.jclinepi.2006.02.022>.
- Schellingerhout, R. (2004). *Gezondheid en welzijn van allochtone ouderen*. Den Haag: The Netherlands Institute for Social Research.
- Schrier, A. C., De Wit, M. A., Rijmen, F., Tuinebreijer, W. C., Verhoeff, A. P., Kupka, R. W., Dekker, J., & Beekman, A. T. (2010). Similarity in depressive symptom profile in a population-based study of migrants in the Netherlands. *Social Psychiatry and Psychiatric Epidemiology*, 45(10), 941–951. <https://doi.org/10.1007/s00127-009-0135-0>.
- Smits, C. H., de Vries, W. M., & Beekman, A. T. (2005). The CIDI as an instrument for diagnosing depression in older Turkish and Moroccan labour migrants: An exploratory study into equivalence. *International Journal of Geriatric Psychiatry*, 20(5), 436–445. <https://doi.org/10.1002/gps.1303>.
- Spijker, J., Van Der Wurff, F. B., Poort, E. C., Smits, C. H., Verhoeff, A. P., & Beekman, A. T. (2004). Depression in first generation labour migrants in Western Europe: The utility of the center for epidemiologic studies depression scale (CES-D). *International Journal of Geriatric Psychiatry*, 19(6), 538–544. <https://doi.org/10.1002/gps.1122>.
- Suurmond, J., Rosenmöller, D. L., El Mesbahi, H., Lamkaddem, M., & Essink-Bot, M. L. (2016). Barriers in access to home care services among ethnic minority and Dutch elderly—a qualitative study. *International Journal of Nursing Studies*, 54, 23–35. <https://doi.org/10.1016/j.ijnurstu.2015.02.014>.
- University of Connecticut. (2014). CT, USA. Available from: www.davidakenny.net/cm/fit.html [last accessed 8 Jun 2014].
- Van de Velde, S., Bracke, P., Levecque, K., & Meuleman, B. (2010). Gender differences in depression in 25 European countries after eliminating measurement bias in the CES-D. *Social Science Research*, 39 (3), 396–404. <https://doi.org/10.1016/j.ssresearch.2010.01.002>.
- Van den berg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4–70. <https://doi.org/10.1177/109442810031002>.
- Van Der Wurff, F. B., Beekman, A. T., Dijkshoorn, H., Spijker, J. A., Smits, C. H., Stek, M. L., & Verhoeff, A. (2004). Prevalence and risk-factors for depression in elderly Turkish and Moroccan migrants in the Netherlands. *Journal of Affective Disorders*, 83(1), 33–41. <https://doi.org/10.1016/j.jad.2004.04.009>.
- Van Der Zwaan, R., & Tolsma, J. (2013). Integration and depression; The relationship between socio-cultural integration and depression in Turkish and Moroccan Dutch. (In Dutch: Integratie en depressie; De relatie tussen sociaal-culturele integratie en depressieklachten bij Turkse en Marokkaanse Nederlander). *Mens en Maatschappij*, 88(2), 177–205. <https://doi.org/10.5117/mem2013.2.zwan>.
- Yang, F. M., Tommet, D., & Jones, R. N. (2009). Disparities in self-reported geriatric depressive symptoms due to sociodemographic differences: An extension of the bi-factor item response theory model for use in differential item functioning. *Journal of Psychiatric Research*, 43(12), 1025–1035. <https://doi.org/10.1016/j.jpsychires.2008.12.007>.
- Zhang, B., Fokkema, M., Cuijpers, P., Li, J., Smits, N., & Beekman, A. (2011). Measurement invariance of the center for epidemiological studies depression scale (CES-D) among Chinese and Dutch elderly. *BMC Medical Research Methodology*, 11(74), 11–74. <https://doi.org/10.1186/1471-2288-11-74>.