

International Journal of Methods in Psychiatric Research  
*Int. J. Methods Psychiatr. Res.* 18(3): 159–168 (2009)  
Published online in Wiley InterScience  
([www.interscience.wiley.com](http://www.interscience.wiley.com)) DOI: 10.1002/mpr.288

# Psychometric properties of an interviewer-administered version of the Kessler Psychological Distress scale (K10) among Dutch, Moroccan and Turkish respondents

T. FASSAERT,<sup>1,2</sup> M.A.S. DE WIT,<sup>1</sup> W.C. TUINEBREIJER,<sup>3</sup> H. WOUTERS,<sup>4</sup> A.P. VERHOEFF,<sup>1,5</sup>  
A.T.F. BEEKMAN<sup>6,7</sup> & J. DEKKER<sup>6,8</sup>

1 Department of Epidemiology, Documentation and Health Promotion, Municipal Health Service, Amsterdam, The Netherlands

2 Department of Social Medicine, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

3 Department of Public Mental Health Care, Municipal Health Service, Amsterdam, The Netherlands

4 Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

5 Department of Sociology and Anthropology, University of Amsterdam, Amsterdam, The Netherlands

6 Department of Psychiatry, VU University, Amsterdam, The Netherlands

7 GGZ in Geest, Amsterdam, The Netherlands

8 Arkin, Amsterdam, The Netherlands

---

## Key words

ethnic groups, psychometrics, mass screening, anxiety disorders, depressive disorders

## Correspondence

T. Fassaert, GGD Amsterdam, Cluster EDG, Nieuwe Achtergracht 100, P.O. Box 2200, 1000 CE Amsterdam, The Netherlands.  
Telephone (+31) (0)20 555 5026  
Fax (+31) (0)20 555 5160  
Email: [tfassaert@ggd.amsterdam.nl](mailto:tfassaert@ggd.amsterdam.nl)

## Abstract

The Kessler Psychological Distress scale (K10) is an instrument that is widely used to screen for mental disorders, but information is lacking on its psychometric qualities in non-Western samples. This study used a population-based sample ( $N = 725$ ) to assess the reliability and validity of the K10 across ethnic groups in an urban area. The results were generally supportive of the K10 as a reliable and valid instrument to screen for anxiety and depression in all three groups. Cronbach's alpha was high (0.93) and the results indicated the existence of a solid single factor structure. Item bias in relation to ethnic background was minor. In each group, there was good criterion validity with respect to one-month DSM-IV diagnosis for depressive and/or anxiety disorder. The results nevertheless highlight the importance of cross-cultural validation, as we found different cut-off values for ethnic subgroups to obtain optimal sensitivity and specificity for detecting depressive and/or anxiety disorders. Copyright © 2009 John Wiley & Sons, Ltd.

Received 19 December 2007;  
revised 7 March 2008;  
accepted 7 March 2008

## Introduction

The Kessler Psychological Distress scale (K10; Kessler *et al.*, 2002) is becoming increasingly popular as a screening tool for anxiety and depressive disorders (Kessler *et al.*, 2003; Furukawa *et al.*, 2003; Cairney *et al.*, 2007; Hides *et al.*, 2007). Although strongly disabling and highly prevalent in the general population (Aydin and Uluşahin, 2001; Chwastiak and Von Korff, 2003; Kessler, 2007), both conditions are often unrecognized. The K10 was developed by Kessler *et al.* (2002) and consists of 10 items, rated on five-point Likert-type scales, which indicate the degree to which symptoms of psychological distress are present among individuals. There is strong evidence supporting the relationship between the K10 and the Composite International Diagnostic Interview (CIDI) for anxiety and depressive disorders in Western population samples (WHO, 1997; Andrews & Slade, 2001; Rodgers *et al.*, 2007). The CIDI is nowadays widely considered as a standard for determining psychological disorders.

However, the reliability and validity of the K10 have been established mainly with data from Western population samples. Information on cross-cultural validity of the K10 among non-Western populations is insufficiently available (Boufous *et al.*, 2005). In research where multiple nationalities or ethnicities are involved, such gaps in knowledge are often dealt with either by excluding minorities from epidemiological research or by simply assuming that cross-cultural validity exists (e.g. Streiner and Norman, 2003). Both options can be considered undesirable for several reasons.

The first decision tends to make studies less representative, thus reducing the generalizability of results (Streiner and Norman, 2003). In this context, it cannot be ignored that large and still increasing parts of the general population in Europe and the United States consist of migrants (e.g. Watters, 2002). Their mental health is often worse than that of the members of the host society (Bhugra, 2004), as is the case for depressive and anxiety disorders (Marwaha and Livingston, 2002). In the Netherlands, anxiety and depression are more prevalent among Turkish and Moroccan labour migrants (Bengi-Arslan *et al.*, 2002; van der Wurff *et al.*, 2004; de Wit *et al.*, 2009). Labour migration from Morocco and Turkey to Western Europe started halfway through the 1960s. Apart from the Netherlands, countries with large Turkish and Moroccan migrant populations are France, Germany and Belgium. In Amsterdam, 14% of the population are ethnic Moroccan or Turkish, which means they or at least one of their parents were born in Morocco or Turkey (O&S, 2007).

However, simply assuming cross-cultural validity without sufficient evidence to support this, disregards possible changes in reliability and validity that may occur when an instrument is translated (Streiner and Norman, 2003). It additionally neglects the influence of culture on the interpretation of items (Azocar *et al.*, 2001). It is important to realize that many migrants in Western countries, like Turkish and Moroccan labour migrants, have an Arabic and/or Muslim background. In the United States, for example, Muslims are in fact the fastest growing minority population (Ali *et al.*, 2005). Muslims tend to have different values with respect to psychological problems, which determines how problems are perceived and coped with (Al-Krenawi, 2005). Therefore, this study focuses on the psychometric qualities of the K10 in a sample of ethnic Dutch, Turkish and Moroccan participants. Our aim is to examine whether the K10 is an instrument with deviating psychometric properties among non-Western (i.e. Moroccan and Turkish) participants compared to ethnic Dutch.

## Methods

### Setting and procedures

This study is part of the larger Amsterdam Health Monitor (AHM), which is a cross-sectional population-based health survey, designed and conducted every four years by the Amsterdam Municipal Health Service. The AHM of 2004 was conducted in collaboration with the Dutch National Institute for Public Health and Environment (RIVM) and was based on a representative sample of 1736 people from the communal population register, stratified by age groups (18–34 years, 35–44 years, 45–54 years, 55–64 years and 65 years and older) and ethnicity (ethnic Dutch, Turkish, Moroccan, other). Turkish and Moroccan respondents were oversampled to ensure sufficient response. Information on the representativity of the first phase can also be found in Agyemang *et al.* (2006). The socio-economic status of the respondents of the first phase was comparable to that of the original sample. All respondents participated in a face-to-face interview on general health, in the language of their choice (i.e. Dutch, Turkish, Moroccan-Arabic or Berber). Thus, participants not fluent in the Dutch language were included as well.

Respondents who gave permission were invited to participate in the next phase, aimed specifically at common mental disorders (de Wit *et al.*, 2009). Again, structured interviews were performed in the language that was preferred by respondents. The interview included the K10 (Kessler *et al.*, 2002) and the CIDI version 2.1 (WHO, 1997) for depressive and anxiety disorders. The entire

interview could be completed within 1.5 hours. Interviewers were trained during a full-time week and monitored during the period of data-collection. To avoid non-response as a consequence of summer vacation, Christmas and Ramadan, all interviews were planned between February and June of 2005. The study procedures were approved by the medical ethical committee of the Academic Medical Centre of the University of Amsterdam. Compared to the general first phase of the AHM, we found no selection among participants in the second phase with respect to physical or psychological health indicators (de Wit *et al.*, 2009).

**Measures**

**The K10**

The K10 was developed as a short screening scale for psychological distress (Kessler *et al.*, 2002). Items for the K10 were selected from 612 questions that were derived from 18 existing instruments, including well-known instruments such as the self-rating depression scale (Zung, 1965), the Beck Depression Inventory (BDI; Beck *et al.*, 1961), and the Centre for Epidemiologic Studies – Depression Scale (CES-D; Radloff, 1977). The initial set of items was drastically reduced after extensive pilot testing. The final scale consisted of 10 items (see Figure 1). Each item of the K10 has five response categories: ‘none of the time’ (1), ‘a little of the time’ (2), ‘some of the time’ (3), ‘most of the time’ (4) and ‘all of the time’ (5). The total score is the sum of all responses. The scores thus range between 10 and 50. Only if participants had less than 10%

missing data, missing values were imputed using linear interpolation. Further information about the K10 can be found on the National Comorbidity Survey (NCS) website (<http://www.hcp.med.harvard.edu/ncs/index.php>).

At the time of conducting the AHM, an official Dutch translation of the K10 was already available (Alonso *et al.*, 2004). For Turkish and Moroccan participants, the K10 had to be translated.<sup>1</sup> To ensure conceptual equivalence, the instrument was translated and back-translated by two professional and independent translators. The back-translation was compared to the original translation, and no further alterations were considered necessary. These procedures were restricted to the Turkish version. A Moroccan translation was not feasible because the Moroccan population in Amsterdam speaks several languages (i.e. Northern/Southern Berber and Standard/Moroccan Arabic) and because Moroccan Arabic and Berber do not exist in written form. In the past, this problem was dealt with by translating instruments into standard Arabic, after which interviewers were asked to translate to Moroccan Arabic and Berber on the spot. However, the translation from Standard Arabic into

<sup>1</sup> A comment should be made at this point. After the data-collection for the AHM of 2004 was finished, another Turkish version of the K10 Self-administered Questionnaire was published at the website of the Australian Transcultural Mental Health Centre NSW (in 2005) (<http://www.dhi.gov.au/tmhc/resources/translations.htm>). To some extent, this version may differ from our version.

| Item | During the past 30 days, ...  |
|------|---|
| 1    | ...about how often did you feel tired out for no good reason?           |
| 2    | ...about how often did you feel nervous?                                |
| 3    | ... how often did you feel so nervous that nothing could calm you down? |
| 4    | ...about how often did you feel hopeless?                               |
| 5    | ...about how often did you feel restless or fidgety?                    |
| 6    | ...how often did you feel so restless that you could not sit still?     |
| 7    | ...about how often did you feel depressed?                              |
| 8    | ...how often did you feel so depressed that nothing could cheer you up? |
| 9    | ...about how often did you feel that everything was an effort?          |
| 10   | ...about how often did you feel worthless?                              |

**Figure 1** Items of the K10.

Moroccan Arabic or Berber is almost as difficult as the direct translation from Dutch. Moreover, only few Moroccans who nowadays live in Amsterdam are able to speak/read both Dutch and Standard Arabic fluently. Therefore, Moroccan interviewers generally used the Dutch version of the K10, with only core themes pre-translated, using standard Arabic words that are commonly used in Morocco. Thus, there is no 'Moroccan' translation in written form, but interviews have been digitally recorded.

### The gold standard

We defined the gold standard as the presence of a current anxiety and/or depressive disorder according to DSM-IV criteria (e.g. Furukawa *et al.*, 2003), as established with the CIDI version 2.1 (WHO, 1997). The CIDI has been translated into Dutch, Turkish and Arabic (Karam *et al.*, 1995; Simon *et al.*, 2002). For the interview with Moroccan participants, again the Dutch questionnaire was used in combination with core themes from the Arabic version. Based on the CIDI, DSM-IV [Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)] diagnoses were made for depressive disorders (major depressive disorder, dysthymia) and anxiety disorders (social phobia, agoraphobia, panic disorder and generalized anxiety disorder). For this study, only one-month prevalence figures were used, as the K10 applies to the four weeks preceding the interview.

### The World Health Organization Disability Assessment Schedule II (WHODAS II)

The World Health Organization Disability Assessment Schedule II (WHODAS II) (WHODAS, 2007) was included as a measure of health related disability, in this case related to mental health problems. The WHODAS II consists of 36 items that are used to rate difficulty with a set of activities. These activities are related to six domains, namely (a) understanding and communicating, (b) getting around, (c) self care, (d) getting along with people, (e) life activities, (f) participation in society and (g) daily work/school (if applicable). All items are rated on five-point Likert-type scales, with high scores indicating high levels of disability. The Turkish version is reliable and valid (Uluğ *et al.*, 2001). For more details on the WHODAS II the reader is referred to the WHODAS website (<http://www.who.int/icidh/whodas/index.html>).

### Socio-demographic characteristics

Socio-demographic measures that were used in this study were ethnicity, age, sex, educational level and health

insurance type. A participant was considered to be Turkish or Moroccan if he or she was born in Turkey or Morocco or if at least one parent was born in Turkey or Morocco (O&S, 2007). Respondents were considered ethnic Dutch if the respondent and both parents were born in the Netherlands. Educational level was divided into two categories, namely low (no education or only primary school) and high (all higher levels of education). Health insurance was used as a covariate for socioeconomic status (SES) because almost everybody in the Netherlands has medical insurance and, until January 2006, people with an income below a certain level had public insurance. Conversely, people with higher incomes were privately insured.

### Statistical analyses

We first examined the internal consistency of the scale, then the factor structure (internal validity), item bias towards ethnic groups (external validity), and predictive validity.

### Internal consistency

SPSS version 14 (SPSS, 2006) was used to calculate Cronbach's alpha and item total correlations. Cronbach's alpha had to be at least 0.70 in each ethnic group, while values of  $\alpha$  exceeding 0.90 and inter-item correlations higher than 0.70 were indicative of redundancy. Corrected item-total correlations, indicating the correlation of a particular item to the total score calculated without that certain item, had to be higher than >0.11 (Kojima *et al.*, 2002).

### Factor structure and item bias

Two predictions can be derived from the summation of the K10 items and its aim to screen for non-specific psychological distress. First, a single dimension of non-specific psychological distress explains the data. Second, no interference of the relationship between psychological distress and the K10 item responses by ethnicity is to be expected. From a modern psychometric perspective this interference is called differential item functioning (DIF; Fischer and Molenaar, 1995; Crane *et al.* 2006). DIF refers to the idea that only the single dimension, which in this case is psychological distress, explains the responses made by the participants. In case of DIF, which was undesired, participants with different ethnic backgrounds, but with equal underlying levels of psychological distress, would differ in the probabilities of assenting to K10 items and their categories. In case of absent DIF, which was desired, the probability of assenting to the K10 items and their categories would not differ across ethnic groups.

**Table 1** Socio-demographic characteristics of study sample

|   | Ethnic Dutch<br>( <i>N</i> = 321) | Moroccan<br>( <i>N</i> = 191) | Turkish<br>( <i>N</i> = 213) | <i>p</i> <sup>1</sup> |
|---|-----------------------------------|-------------------------------|------------------------------|-----------------------|
| <i>Sex</i>  |                                   |                               |                              |                       |
| Male (%)  | 41.7                              | 52.9                          | 39.9                         | 0.016                 |
| <i>Age</i>  |                                   |                               |                              |                       |
| Mean <sup>2</sup>   | 54.1 (14.6)                       | 49.6 (14.4)                   | 47.3 (14.2)                  | <0.001                |
| Range   | 19–92                             | 19–91                         | 20–82                        |                       |
| <i>Education</i>  |                                   |                               |                              |                       |
| Higher (%)  | 80.3                              | 39.1                          | 40.5                         | <0.001                |
| <i>Insurance</i>  |                                   |                               |                              |                       |
| Public (%)  | 63.2                              | 93.0                          | 80.7                         | <0.001                |
| <i>One-month CIDI-diagnosis</i>                             |                                   |                               |                              |                       |
| Positive (%)  | 7.5                               | 8.9                           | 20.7                         | <0.001                |
| Weighted percentage (%)                                     | 6.6                               | 9.8                           | 18.7                         |                       |
| <i>WHODAS II</i>  |                                   |                               |                              |                       |
| Mean <sup>2</sup>   | 52.0 (15.9)                       | 57.8 (23.8)                   | 67.2 (26.5)                  | <0.001                |
| Range   | 36–116                            | 36–144                        | 36–162                       |                       |
| <i>K10</i>  |                                   |                               |                              |                       |
| Number of participants with one missing value               | 1                                 | 1                             | 4                            |                       |
| Number of participants with > one missing value             | 0                                 | 2                             | 2                            |                       |
| Raw sum score <sup>2,3</sup>                                | 15.0 (5.2)                        | 20.6 (9.7)                    | 18.7 (9.5)                   | <0.001                |
| Range <sup>3</sup>  | 10–39                             | 10–50                         | 10–50                        |                       |
| Percentage participants with minimum sum score <sup>3</sup> | 12.8                              | 13.3                          | 18.0                         |                       |
| Percentage participants with maximum sum score <sup>3</sup> | 0.0                               | 1.4                           | 0.5                          |                       |

<sup>1</sup> Means were tested with ANOVA, proportions with  $\chi^2$ -tests.

<sup>2</sup> Standard deviations are given in parentheses.

<sup>3</sup> Participants with more than one missing value were excluded.

Uni-dimensionality of the scale was examined in Mplus (Muthén and Muthén, 2007), with explorative and confirmative factor analysis for categorical data. The one parameter logistic model (OPLM; Verhelst and Glas, 1995), a Rasch type of model, was then used to estimate the levels of non-specific psychological distress measured by the answering categories of the items. Subsequently, the fit of the items to the OPLM model was examined. Poor fit, as a consequence of differences between items in measurement precision or non-hierarchically ranking of answering categories, could be corrected for by weighting the items or collapsing the answering categories respectively. Only if poor fit of the data was caused by a particular item not measuring the same construct as the other items, this item was excluded. The significance of ethnic DIF was examined for the answering categories with  $\chi^2$  distributed fit statistics (alpha set at 0.01 for multiple testing). Because significant results are not necessarily

relevant as well, items with significant DIF were graphically inspected to see whether members from different ethnic groups, but with equal estimated levels of psychological distress, had substantial different probabilities of assenting the categories of each K10 item.

### Predictive validity

Receiver Operating Characteristic (ROC) analysis was carried out in SPSS version 14 to assess sensitivity, specificity, positive predictive value (PPV) and the area under the curve (AUC). Furthermore, based on the well-established association between common mental disorders and disability, we hypothesized that a strong relation existed between the K10 and the WHODAS II, and that this association was similar across ethnic groups. Therefore, additional support for cross-cultural validity was

**Table 2** Results of the ROC-analysis: sensitivity, specificity, positive predictive value (PPV) and area under the curve (AUC)

| Cut-off value                 | Ethnic Dutch        |             |      | Moroccan            |             |      | Turkish             |             |      |
|-------------------------------|---------------------|-------------|------|---------------------|-------------|------|---------------------|-------------|------|
|                               | Sensitivity         | Specificity | PPV  | Sensitivity         | Specificity | PPV  | Sensitivity         | Specificity | PPV  |
| 14.5                          | 0.875               | 0.667       | 18.1 | 0.941               | 0.518       | 16.2 | 0.909               | 0.428       | 31.1 |
| 15.5                          | 0.833               | 0.712       | 19.6 | 0.941               | 0.576       | 18.0 | 0.909               | 0.472       | 32.8 |
| 16.5                          | 0.792               | 0.768       | 22.4 | 0.941               | 0.612       | 19.3 | 0.909               | 0.541       | 35.9 |
| 17.5                          | 0.750               | 0.796       | 23.7 | 0.941               | 0.653       | 21.1 | 0.909               | 0.585       | 38.2 |
| 18.5                          | 0.708               | 0.818       | 24.8 | 0.941               | 0.671       | 21.9 | 0.886               | 0.616       | 39.4 |
| 19.5                          | 0.625               | 0.842       | 25.0 | 0.941               | 0.694       | 23.2 | 0.841               | 0.642       | 39.8 |
| 20.5                          | 0.583               | 0.870       | 27.5 | 0.941               | 0.735       | 25.8 | 0.795               | 0.686       | 41.6 |
| 21.5                          | 0.500               | 0.902       | 30.0 | 0.882               | 0.741       | 25.0 | 0.795               | 0.736       | 45.7 |
| 22.5                          | 0.417               | 0.926       | 32.3 | 0.824               | 0.776       | 26.4 | 0.795               | 0.748       | 46.8 |
| 23.5                          | 0.417               | 0.951       | 41.7 | 0.765               | 0.800       | 27.1 | 0.705               | 0.786       | 47.8 |
| 24.5                          | 0.375               | 0.954       | 40.9 | 0.765               | 0.818       | 28.9 | 0.636               | 0.811       | 48.4 |
| 25.5                          | 0.333               | 0.965       | 44.4 | 0.588               | 0.859       | 28.6 | 0.568               | 0.843       | 50.0 |
| AUC (95% confidence interval) | 0.853 (0.786–0.920) |             |      | 0.875 (0.802–0.947) |             |      | 0.803 (0.725–0.881) |             |      |

derived from absent interaction between ethnicity and K10 sum score, using regression analyses in SPSS version 14.

## Results

### Sample characteristics

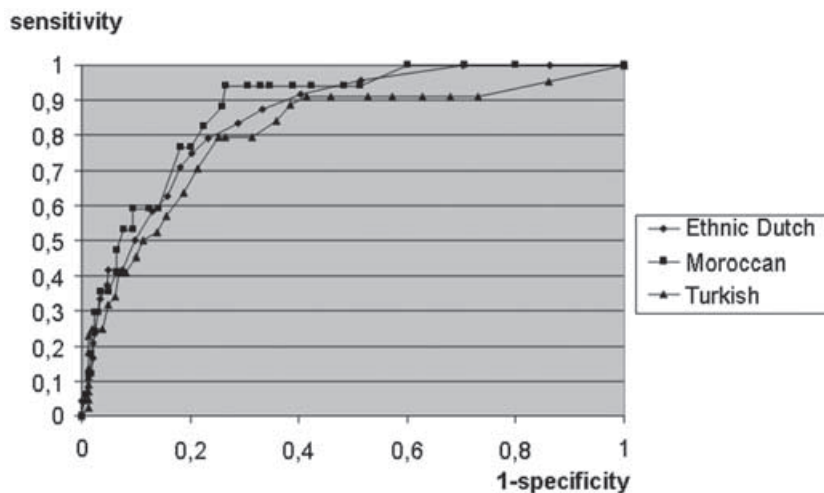
Table 1 shows characteristics of the study sample. There were relatively few women among the Moroccan participants compared to ethnic Dutch and Turkish participants. Furthermore, the average age and educational levels were higher for ethnic Dutch respondents than for participants with a migrant background. With respect to the outcome measures, the highest level of 1-month diagnoses for depressive and anxiety disorders was found among Turkish respondents. Furthermore, Turkish and Moroccan participants both had higher scores on the WHODAS II compared to ethnic Dutch. With respect to the K10, only few respondents had missing values. Both the sum scores and the range of scores were significantly higher for both Moroccan and Turkish respondents compared to the ethnic Dutch group.

### Internal consistency

Internal consistency of the K10 was supported by a high Cronbach's alpha of 0.93, well above the generally accepted minimum of 0.70. Items of the K10 were highly inter-correlated (0.43–0.74). Highly correlated item-pairs ( $r > 0.70$ ) were items two and three and items seven and nine. Additionally, item-total correlations were high (0.61–0.79).

### Factor structure and item bias

Explorative factor analysis revealed a dominant first factor which explained 70% of the variance. Factor loadings ranged from 0.67 to 0.89. Confirmatory factor analysis showed sufficient fit of the data to a single factor model. The CFI index was 0.97 ( $CFI > 0.95$  indicates good fit) and the RMSEA was 0.09 ( $0.08 > RMSEA < 0.10$  indicates mediocre fit). After weighting the items for their measurement precision and correcting for non-hierarchically ranked answering categories, there was good fit of the items to the OPLM model. There was only minor ethnic DIF. Items one, five and nine had significant DIF and were graphically inspected. In case of similar levels of non-specific psychological distress, the ethnic Dutch respondents had a larger probability of affirming item five than the Turkish. Conversely, ethnic Dutch had a lower probability than the Turkish and the Moroccans to affirm item nine. Though significant, the DIF findings from item one were judged as not relevant.



**Figure 2** ROC-curves for the K10 predicting one-month depressive and/or anxiety disorder.

**Predictive validity**

Table 2 shows the results for the ROC analysis, indicating good results for all ethnic groups. The AUCs were generally high, varying between 0.80 (Turkish respondents) and 0.88 (Moroccans). Figure 2 clearly shows that differences were small. Except for the Turkish group, PPVs were generally low. Table 2 furthermore shows that, to obtain a similar balance between sensitivity and specificity across ethnic groups, cut-off scores varied between ethnic Dutch on the one hand and Turkish and Moroccan participants on the other hand. For example, with comparable sensitivity (~0.80) and specificity (~0.75), the cut-off score for Moroccan and Turkish participants (22.5) was higher than for ethnic Dutch participants (16.5). The PPV was generally highest in the Turkish group.

Finally, there was a strong relation between K10 sum score and disability, regardless of ethnicity (Table 3). There were no significant interaction effects (i.e. between the K10 and ethnicity) suggesting that the association between depression/anxiety, as measured by the K10, and disability is not influenced by ethnic background.

**Discussion**

In response to the rather general comment that ‘there is no evidence to date which indicates that a translated K-10 has similar calibration properties than the original English version’ (Boufous *et al.*, 2005, p. 172), the present study addressed the psychometric properties of the K10 in a multi-ethnic population sample in the Netherlands. The results showed that the K10 appears to be an adequate instrument that has good psychometric properties among Dutch, Turkish, and Moroccan respondents. The results

**Table 3** Linear regression for the association between K10, ethnicity and the WHODAS (*N* = 714)

|                          | Beta  | Standard error | <i>t</i> -Value | Significance |
|--------------------------|-------|----------------|-----------------|--------------|
| K10-sumscore             | 2.00  | 0.17           | 11.64           | 0.00         |
| K10 * Moroccan ethnicity | -2.37 | 1.75           | -1.36           | 0.18         |
| K10 * Turkish ethnicity  | 1.02  | 1.74           | 0.59            | 0.56         |

Note: Also included as main effects (but not depicted): age, sex, education, insurance type and ethnicity.

indicated the existence of a solid single factor structure with largely absent item bias (DIF), suggesting that the non-specific psychological distress as measured by the K10 is negligibly biased towards the ethnic groups examined in this study. Additionally, sensitivity and specificity of the K10 with respect to a one-month CIDI-diagnosis for depressive and/or anxiety disorders were good in all subgroups, as AUCs greater than 0.80 can be judged as indicating good validity (e.g. Hansson *et al.*, 2005). Finally, the results suggested that the K10 is as good in predicting disability among Turkish and Moroccan respondents as it is among ethnic Dutch. However, our study shows that, when applying the K10 in a multi-ethnic sample, different cut-off scores are necessary to achieve comparable sensitivity and specificity across ethnic groups. Within one ethnic group cut-off points may vary according to the

purpose of screening. For example, if one is interested only in severe psychological distress (i.e. moderate cases are to be filtered out), higher cut-off scores may be employed in order to reach high specificity.

The results suggested some redundancy of K10 items, indicated by a high Cronbach's alpha. In that respect the existence of the K6 (Kessler *et al.*, 2002), which is the same instrument as the K10 minus four items (items one, three, six and seven), is interesting. Considering the combinations of highly correlated items as discussed in the results section, removal of these four items would eliminate at least some redundancy. Secondary analyses (not reported) pointed out that Cronbach's alpha for the K6 (0.89) is lower, although marginally, compared to alpha for the K10 (0.93). A simple measure to handle (minor) redundancy among K10 items could therefore be to use the K6 instead of the K10. This would also be in line with Furukawa *et al.* (2003), who reported to prefer the K6 in screening for any mood or anxiety disorder, in terms of brevity and consistency across sub-samples. However, they noted that the K10 was better in screening for severe disorders. Therefore, which version is to be administered will primarily depend on the purpose of the screening.

There are a few potential limitations to this study that do need to be taken into account. Firstly, the results showed that the data fitted the OPLM model, but only after correcting for differences between items in measurement precision and non-hierarchically ranking of answering categories. Actually, we should have calculated the sum scores with the item scales as revised by OPLM for the remaining validity analyses. This, however, would have compromised the comparability of our results with other studies, which generally use raw sum scores. We therefore did not revise the sum scores, but instead checked our results by rerunning the analyses using the revised K10 sum scores (results not reported). Compared to the results presented in this paper, the revised K10 sum score yielded only marginally different results on predictive validity. Secondly, selection bias was tested for with non-response analyses, but selection may have occurred on one or more unknown variable(s). Notable in that respect is the low response among Moroccan women. Finally, a true golden standard for assessment of psychological morbidity in a transcultural setting was lacking. Although the CIDI is widely used and well validated, the World Health Organization (WHO) has formulated some drawbacks of the CIDI if applied in non-Western populations, since most methodological studies on the CIDI have been carried out in Western countries (WHO Consortium, 2004).

Overall, the results of our study were quite favourable. This may suggest that core symptoms of depression and anxiety are not very different across cultures (e.g. Sartorius, 1980; Simon *et al.*, 2002), though we acknowledge cultural variations in clinical expression of depression and anxiety (Kirmayer, 2001). Additional support for this hypothesis might be found in recent findings from the same dataset, by Schrier *et al.* (2007), who showed similarities in symptom profiles for depressive disorder among ethnic Dutch, Turkish and Moroccan participants. Our results appear to be well in line with these findings in a sense that the construct of psychological distress, which the K10 was designed to measure, proved to be invariant across three ethnic groups while item bias was largely absent. We therefore encourage further use of the K10 in the context of inter-ethnic population studies.

### Acknowledgments

The data-collection for this study was funded by the Municipality of Amsterdam and the mental health care institutions Arkin (formerly JellinekMentrum and AMC de Meren) and GGZ inGeest (formerly Stichting Buitenamstel Geestgronden). Additional funding was obtained from the academic collaboration between the Amsterdam Municipal Health Service and the Academic Medical Centre of the University of Amsterdam. This academic collaboration is funded by its partners and a grant from the Netherlands Organisation for Health Research and Development (ZonMw).

### Declaration of interests statement

The authors have no competing interests.

### References

- Agyemang C., Ujcic-Voortman J., Uitenbroek D., Foets M., Droomers M. (2006) Prevalence and management of hypertension among Turkish, Moroccan and native Dutch ethnic groups in Amsterdam, the Netherlands: The Amsterdam Health Monitor Survey. *Journal of Hypertension*, **24**, 2169–2176.
- Ali O.M., Milstein G., Marzuk P.M. (2005) The Imam's role in meeting the counselling needs of Muslim communities in the United States. *Psychiatric Services*, **56**, 202–205.
- Al-Krenawi A. (2005) Mental health practice in Arab countries. *Current Opinion in Psychiatry*, **18**, 560–564.
- Alonso J., Angermeyer M.C., Lépine J.P. (2004) The European Study of the Epidemiology of Mental Disorders (ESEMeD) project: an epidemiological basis for informing mental health policies in Europe. *Acta Psychiatrica Scandinavica*, **109**(Suppl. 420), 5–7.
- Andrews G., Slade T. (2001) Interpreting scores on the Kessler Psychological Distress Scale (K10). *Australian and New*



- Zealand Journal of Public Health*, **25**, 494–497, DOI: 10.1111/j.1467-842X.2001.tb00310.x
- Aydin I.O., Uluşahin A. (2001) Depression, anxiety comorbidity, and disability in tuberculosis and chronic obstructive pulmonary disease patients: applicability of GHQ-12. *General Hospital Psychiatry*, **23**, 77–83, DOI: 10.1016/S0163-8343(01)00116-5
- Azocar F., Aréan P., Miranda J., Muños R.F. (2001) Differential item functioning in a Spanish translation of the Beck Depression Inventory. *Journal of Clinical Psychology*, **57**, 355–365, DOI: 10.1002/jclp.1017
- Beck A.T., Ward C.H., Mendelson M., Mock J., Erbaugh J. (1961) An inventory for measuring depression. *Archives of General Psychiatry*, **4**, 561–571.
- Bengi-Arslan L., Verhulst F.C., Crijnen A.A. (2002) Prevalence and determinants of minor psychiatric disorder in Turkish immigrants living in the Netherlands. *Social Psychiatry and Psychiatric Epidemiology*, **37**, 118–124, DOI: 10.1007/s001270200003
- Bhugra D. (2004) Migration, distress and cultural identity. *British Medical Bulletin*, **69**, 129–141, DOI: 10.1093/bmb/ldh007
- Boufous S., Silove D., Bauman A., Steel Z. (2005) Disability and health service utilization associated with psychological distress: the influence of ethnicity. *Mental Health Services Research*, **7**, 171–179, DOI: 10.1007/s11020-005-5785-2
- Cairney J., Veldhuizen S., Wade T.J., Kurdyak P., Streiner D.L. (2007) Evaluation of 2 measures of psychological distress as screeners for depression in the general population. *Canadian Journal of Psychiatry*, **52**, 111–120.
- Chwastiak L.A., Von Korff M. (2003) Disability in depression and back pain. Evaluation of the World Health Organization Disability Assessment Schedule (WHO DAS II) in a primary care setting. *Journal of Clinical Epidemiology*, **56**, 507–514, DOI: 10.1016/S0895-4356(03)00051-9
- Crane P.K., Gibbons L.E., Jolley L., van Belle G. (2006) Differential item functioning analysis with ordinal logistic regression techniques. DIFdetect and difwithpar. *Medical Care*, **44**(Suppl. 3), 115–123.
- de Wit M.A.S., Tuinebreijer W.C., Dekker J., Beekman, A.T.F., Gorissen W.H.M., Schrier A.C., Penninx B.W.J.H., Komproe I.H., Verhoeff, 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**, 905–912, DOI: 10.1007/s00127-008-0382-5
- Fischer G.H., Molenaar I.W. (1995) *Rasch Models: Foundations, Recent Developments, and Applications*, Springer Verlag.
- Furukawa T.A., Kessler R.C., Slade T., Andrews G. (2003) The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. *Psychological Medicine*, **33**, 357–362, DOI: 10.1017/S0033291702006700
- Hansson S.L., Svanström Røjvall A., Rastam M., Gillberg C., Gillberg C., Anckarsäter H. (2005) Psychiatric telephone interview with parents for screening of childhood autism – tics, attention-deficit hyperactivity disorder and other comorbidities (A-TAC): preliminary reliability and validity. *British Journal of Psychiatry*, **187**, 262–267.
- Hides L., Lubman D.I., Devlin H., Cotton S., Aitken C., Gibbie T., Hellard M. (2007) Reliability and validity of the Kessler 10 and Patient Health Questionnaire among injecting drug users. *Australian and New Zealand Journal of Psychiatry*, **41**, 166–168.
- Karam E.G., Yabroudi P., Nasser-Karam A., Mansour C., Sabah S., Al Atrash R. (1995) The Arabic Composite International Diagnostic Interview (CIDI). *Arabic Journal of Psychiatry*, **6**, 19–29.
- Kessler R.C. (2007) The global burden of anxiety and mood disorders: putting ESEMeD findings into perspective. *Journal of Clinical Psychiatry*, **68**(Suppl. 2), 10–19.
- Kessler R.C., Andrews G., Colpe L.J., Hiripi E., Mroczek D.K., Normand S.-L.T., Walters E.E., Zaslavsky A.M. (2002) Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, **32**, 959–976, DOI: 10.1017/S0033291702006074
- Kessler R.C., Barker P.R., Colpe L.J., Epstein J.F., Gfroerer J.C., Hiripi E., Howes M.J., Normand S.L., Manderscheid R.W., Walters E.E., Zaslavsky A.M. (2003) Screening for serious mental illness in the general population. *Archives of General Psychiatry*, **60**, 184–189.
- Kirmayer L.J. (2001) Cultural variations in the clinical presentation of depression and anxiety: implications for diagnosis and treatment. *Journal of Clinical Psychiatry*, **62**(Suppl. 13), 22–28.
- Kojima M., Furukawa T.A., Takahashi H., Kawai M., Nagaya T., Tokudome S. (2002) Cross-cultural validation of the Beck Depression Inventory-II in Japan. *Psychiatry Research*, **110**, 291–299.
- Marwaha S., Livingston G. (2002) Stigma, racism or choice. Why do depressed ethnic elders avoid psychiatrists? *Journal of Affective Disorders*, **72**, 257–265, DOI: 10.1016/S0165-0327(00)00381-5
- Muthén L.K., Muthén B.O. (2007) *Mplus User's Guide*, 4th ed., Muthén & Muthén.
- O&S website. (2007) Available: <http://www.os.amsterdam.nl/> [29 May 2007].
- Radloff L.S. (1977) The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*, **1**, 385–401.
- Rodgers B., Parslow R., Degenhardt L. (2007) Affective disorders, anxiety disorders and psychological distress in non-drinkers. *Journal of Affective Disorders*, **99**, 165–172, DOI: 10.1016/j.jad.2006.09.006

- Sartorius N., Jablensky A., Gulbinat W., Ernberg G. (1980) WHO collaborative study: assessment of depressive disorders. *Psychological Medicine*, **10**, 743–749.
- Schrier A.C., de Wit M.A.S., Dekker J., Kupka R.W., Beekman A.T.F. (2007) Similar depressive symptom patterns in a population based study of immigrants from Turkey and Morocco and Dutch natives, Abstract presented at the American Psychiatric Association Annual Meeting, San Diego.
- Simon G.E., Goldberg D.P., Von Korff M., Üstün T.B. (2002) Understanding cross-national differences in depression prevalence. *Psychological Medicine*, **32**, 585–594, DOI: 10.1017/S0033291702005457
- SPSS Inc. (2006) SPSS for Windows, Version 14.0.2, SPSS Inc.
- Streiner D.L., Norman G.R. (2003) *Health Measurement Scales. A Practical Guide to their Development and Use*, Oxford University Press.
- Uluğ B., Ertuğrul A., Göğüs A., Kabakci E. (2001) Reliability and validity of the Turkish version of the World Health Organization disability assessment schedule (WHO–DAS-II). *Turkish Journal of Psychiatry*, **12**, 121–130.
- van der Wurff F.B., Beekman A.T.F., Dijkshoorn H., Spijker J.A., Smits C.H.M., 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**, 33–41, DOI: 10.1016/j.jad.2004.04.009
- Verhelst N.D., Glas C.A.W. (1995) The one parameter logistic model. In: *Rasch Models: Foundations, Recent Developments, and Applications* (eds Fischer G.H., Molenaar I.W.), pp. 215–238, Springer Verlag.
- Watters C. (2002) Migration and mental health care in Europe: report of a preliminary mapping exercise. *Journal of Ethnic and Migration Studies*, **28**, 153–172.
- World Health Organization (WHO). (1997) *Composite International Diagnostic Interview (CIDI)*, version 2.1, WHO.
- WHO World Mental Health Survey Consortium. (2004) Prevalence, severity, and unmet need for treatment of Mental Disorders in the World Health Organization World Mental Health Surveys. *JAMA*, **291**, 2581–2590.
- WHODAS Website. (2007) Available: <http://www.who.int/icidh/whodas/> [accessed 16 August 2007].
- Zung W.W. (1965) A self rating depression scale. *Archives of General Psychiatry*, **12**, 63–70.

Copyright of *International Journal of Methods in Psychiatric Research* is the property of John Wiley & Sons, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.