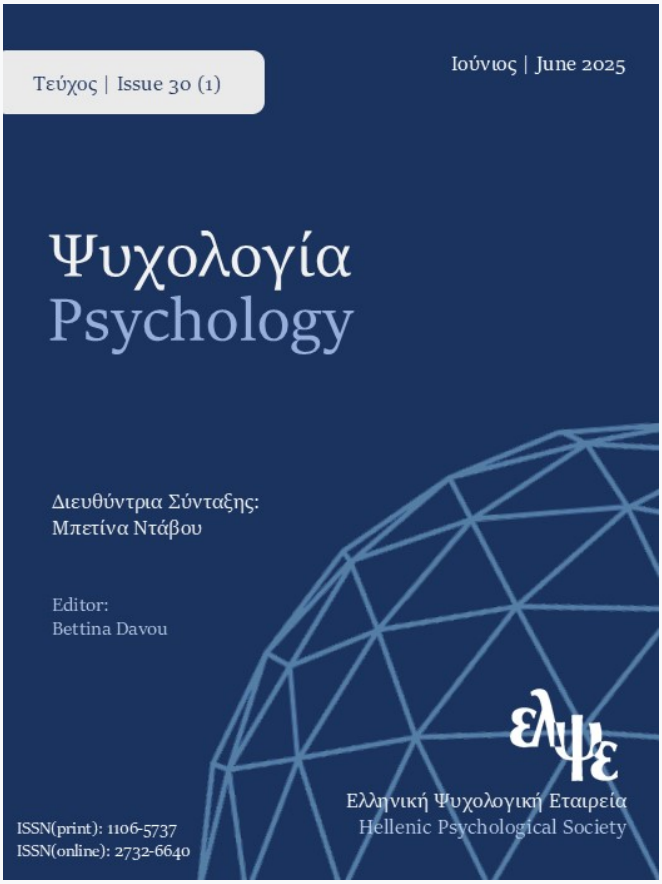
Psychology: the Journal of the Hellenic Psychological Society

Vol 30, No 1 (2025)

June 2025



Translation and validation of the Greek version of the Outcome Questionnaire-45.2

Leto Katakis, Paris Vogazianos, Panagiotis Parpottas, Marianna Anagnostou, Maria Chryssicopoulou, Elena Dragioti, George Tsitsas

doi: 10.12681/psy_hps.39687

To cite this article:

Katakis, L., Vogazianos, P., Parpottas, P., Anagnostou, M., Chryssicopoulou, M., Dragioti, E., & Tsitsas, G. (2025). Translation and validation of the Greek version of the Outcome Questionnaire-45.2. *Psychology: The Journal of the Hellenic Psychological Society*, *30*(1), 141–155. https://doi.org/10.12681/psy_hps.39687

ΨΥΧΟΛΟΓΙΑ | PSYCHOLOGY 30(1), 141-155

Ελληνική Ψυχολογική Εταιρεία
Hellenic Psychological Society

ΕΜΠΕΙΡΙΚΗ ΕΡΓΑΣΙΑ | RESEARCH PAPER

Translation and validation of the Greek version of the Outcome Questionnaire-45.2

Leto KATAKIS[1,2], Paris VOGAZIANOS[1], Panagiotis PARPOTTAS[1], Marianna ANAGNOSTOU[3], Maria CHRYSSICOPOULOU[2], Elena DRAGIOTI[4], George TSITSAS[3]

[1] Department of Social and Behavioral Sciences, European University Cyprus
[2] Laboratory for the Study of Human Relations
[3] Center of Student Counseling, Harokopio University
[4] Research Laboratory Psychology of Patients, Families and Health Professionals, Department of Nursing, School of Health Sciences, University of Ioannina

KEYWORDS

Outcome Questionnaire
Translation
Greek
Confirmatory factor analysis
Validity
Reliability

CORRESPONDENCE

Panagiotis Parpottas
European University Cyprus
6 Diogenis str, 2404 Engomi,
Nicosia, Cyprus
p.parpottas@euc.ac.cy

ABSTRACT

The present study aimed to translate and validate the original English version of the Outcome Questionnaire (OQ-45.2) for use in the Greek context. A sample of 385 psychotherapy trainees (102 men and 283 women), aged 18 to 65 years, was recruited from a systemic/family psychotherapy training center in Greece. Participants completed the Greek version of the OQ-45.2 (OQ-45.2-GR), the World Health Organization Quality-of-Life Scale (WHOQOL-BREF), and a demographic questionnaire. Confirmatory factor analysis supported a three-factor structure consistent with the original English version, yielding satisfactory robust fit indices (χ^2/df, p = .001; RMSEA = .060; SRMR = .091). The OQ-45.2-GR demonstrated high internal consistency (Cronbach's α = .912) and acceptable test-retest reliability (r = .64, p < .01). Concurrent validity was evidenced through significant negative correlations between the OQ-45.2-GR and WHOQOL-BREF total and subscale scores (r ranging from -.29 to -.69, p < .01). These findings provide preliminary evidence for the reliability and validity of the OQ-45.2-GR as a measure of psychotherapy outcomes in Greek-speaking populations. The OQ-45.2 is a widely utilized instrument for evaluating psychotherapy outcomes and serves additionally as a screening tool to assess client needs at therapy initiation, thereby informing clinical decision-making and optimizing treatment planning. The availability of a validated Greek version of the OQ-45.2 is expected to enhance clinical assessment practices, contribute to evidence-based treatment planning, and facilitate further research into psychotherapy processes and outcomes within Greece.

Introduction

The field of evidence-based practice in mental health has witnessed significant growth in recent decades. This expansion has spurred a specific interest in the evaluation of therapy effectiveness with a great amount of research focusing on therapy outcomes. However, the literature reveals different conceptualizations of evidence-based practice (Laska et al., 2014) which naturally lead to the question of how therapy outcome is assessed. Burlingame and Beecher (2008), delineate three principal methodologies within evidence-based practice: empirically supported treatments, practice guidelines, and practice-based evidence. The first two predominantly utilize treatment protocols for specific diagnoses, often referencing the American Psychiatric Association's classifications. In contrast, practice-based evidence emphasizes the appraisal of patients' outcomes and therapy

https://doi.org/10.12681/psy_hps.39687

Psychology: The journal of the Hellenic Psychological Society
https://ejournals.epublishing.ekt.gr/index.php/psychology

processes and specifically evaluates the treatment's efficacy for each unique client receiving it. A pivotal advantage of monitoring outcomes, is that the early detection of negative effects can inform timely revisions to treatment plans, while positive developments can affirm the efficacy of the chosen treatment approach.

The scholarly landscape reveals a great number of conceptualizations and methods for assessing evidence-based therapeutic outcomes. More specifically, when reviewing the literature, one encounters numerous different known practice-based outcome measures (Lambert, 2013; Lambert & Barley, 2001; Lambert et al., 2013). The majority of measures include an assessment of clients' subjective internal experiences including general discomfort, symptomatology for diagnostic purposes, social functioning, adjustment levels and the perceived quality of interpersonal relationships (Burlingame et al., 2013; Lambert & Barley, 2001; Lambert et al., 2013). Basic measures are at first, the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) (Evans et al, 2019), a standardized self-report questionnaire developed to evaluate psychological distress, functioning, and well-being in individuals receiving psychological therapy. Designed for routine clinical use, CORE-OM captures changes over time and helps assess the effectiveness of therapeutic interventions. The second is the Brief Symptom Inventory (BSI) (Derogatis & Melisaratos, 1983), a widely used psychological self-report tool designed to measure psychological distress and psychiatric symptoms. The third is the Outcome Questionnaire (OQ-45.2) which serves as a comprehensive tool throughout the therapeutic process. More specifically, during the assessment phase it can be used as a screening tool, while it can contribute to decision-making during therapeutic sessions. Lastly, it can be utilized at the end of therapy to evaluate the therapeutic outcome. The OQ-45 has a long history in measuring therapy outcomes, beginning with its use to demonstrate treatment effects beyond placebo, and later, its development in evaluating differences between therapeutic approaches. From there on, OQ-45 was used as a mean to enhance treatment outcomes through continuous monitoring of clients' feedback. Notably, the OQ-45 is the first outcome system recognized as an evidence-based measure in USA (NRE – National Registry of Evidenced based programs and practice).

The OQ-45 assesses three aspects of clients' subjective internal experience: (a) subjective discomfort (SD), (b) interpersonal relationships (IR) and (c) social role performance (SP). The SD dimension, includes items that address depressive and anxiety symptoms, commonly associated with a range of psychological disorders (Feldman, 1993; Lambert et al., 2013), as well as with symptoms of affective and adjustment disorders and substance abuse (Gillaspy et al., 2002; Hatfield & Ogles, 2004). The second dimension, IR, evaluates issues related to satisfaction with and problems in interpersonal relationships. The perceived satisfaction plays a major role for the maintenance of person's sense of well-being (Diener, 1984; Zautra, 1983). Problems deriving from intimate relationships may cause personal discomfort, which could be linked to psychopathological symptoms (Horowitz et al., 1988). Finally, the questionnaire examines the SP dimension, which reflects the individual's ability to work, study and engage in leisure activities, thereby gauging their capacity to fulfill various social roles (Lambert et al., 2013).

The OQ-45 exhibits robust psychometric properties, especially in its total scores (Lambert et al., 2004, as cited in Lo Coco et al., 2012). Its three-week test-retest reliability for non-treatment samples is acceptable as is estimated at r = .84, with its internal consistency reliability at α = .93 (Lambert et al., 2013). Additionally, according to its manual (Lambert et al., 2013), the OQ-45 demonstrates satisfactory concurrent validity with other outcome measures such as the Beck Depression Inventory (BDI), the Symptom Checklist–90–Revised (SCL-90-R), the Zung Self-Rating Anxiety Scale (ZSAS), and others (from r = .78 to r = .88, p < .05). Moreover, in alcohol treatment studies, the OQ-45 has shown excellent reliability, with a Cronbach's α coefficient .94 for baseline and follow-up measures respectively, as well as concurrent validity with other established outcome measures such as the BDI and the SCL-90-R (Beck & Steer, 1993; Derogatis, 1977; Gillaspy et al., 2002; Lambert et al., 2013). Furthermore, the OQ-45's high internal consistency and reliability, coupled with its strong concurrent validity, have led to its translation and validation in various languages and cultural contexts.

The OQ-45 has been translated into several languages and validated in a number of countries. Namely Dutch (de Beurs et al., 2005; de Jong et al. 2007), Italian (Lo Coco et al., 2008), German (Huag, 2004), Chinese (Li &

Luo, 2009), Japanese (Takara et al., 2017) and French (Brosseau-Liard et al., 2020), with positive results for their psychometric properties. More specifically, the Dutch translation showed test-retest reliability ranging from r = .70 to r = .83 which is comparable to American samples, and Cronbach's α ranging from .77 to .92 with the SP scale providing weaker internal consistency (de Jong et al., 2007). Moreover, the Chinese translation produced Cronbach α coefficients ranging from .91 to .92 (Li & Luo, 2009), and finally the Italian version (Lo Coco et al., 2008) demonstrated a very good reliability with a Cronbach's α = .93 for the scale's total score, α = .92 for SD, α = .77 for IR and α = .66 for SP.

Among the existing outcome measures, we chose to explore the OQ-45 and more specifically its latest version (OQ-45.2; Lambert et al., 2013). The OQ-45.2 was selected for the current study because it was developed on the basis of the abovementioned areas of functioning and due to the absence of a reliable outcome measure in Greek language. Moreover, OQ-45.2 has robust psychometric properties (Lambert, 2013; Lambert et al., 2013; Yusof & Carpenter, 2014) and is at a manageable length, giving an important advantage for use in research. Furthermore, it offers a balanced approach, capturing both symptom distress and role functioning, and is especially useful for routine outcome monitoring in psychotherapy, where the emphasis lies in tracking client change over time and informing adjustments to therapeutic interventions. Therefore, the aim of the current study was to translate OQ-45.2 in Greek and evaluate its psychometric properties.

Despite its widespread use, the OQ-45.2 has yet to be translated and tested in Greek. Thus, the primary aim of this study was to translate the OQ-45.2 into Greek and examine its psychometric properties, including its factor structure, criterion, convergent validity, internal consistency and test-retest reliability. Furthermore, this study objective was to determine whether the OQ-45.2 is a useful tool for the measurement of therapy outcomes in the Greek cultural context. Our hypothesis is that the Greek version will mirror the factorial structure of the original and other validated translations (Lambert et al., 2013) and also demonstrate good internal consistency. Additionally, it is hypothesized that the OQ-45.2 will correlate with an established measure of quality of life, the Greek version of the World Health Organization Quality-of-Life Scale (WHOQOL-BREF) (Ginieri-Coccossis et al., 2012).

Method

Participants

The sample was drawn from a pool of Systemic/Family psychotherapy trainees in the Greater Athens area in Greece and recruited from one Systemic/Family psychotherapy center. A total of 386 participants were recruited to investigate the validity and reliability of the scale, with one participant not fully completing the questionnaire and being excluded from the analysis. Hence, our final sample consisted of 385 participants, with 102 men (26.5%) and 283 women (73.5%). Participants aged 18 to 65 years, with 42 participants being in the group of 18-24 years (10.9%), 140 in the group of 25-34 (36.5%), 92 in the group of 35-44 (24.0%), 70 in the group of 45-54 (18.2%), 34 in the group of 55-64 (8.9%) and six were aged 65 (1.6%).

Procedure

The study protocol was approved by the EUC Institutional Ethics Committee (2022-01) and all procedures were in accordance with the ethical standards of the institution and with the 1964 Helsinki declaration for studies involving human participants.

Before commencing the study, the OQ-45.2 was translated into Greek following the established guidelines for the translation and cultural adaptation of self-reported outcomes (Wild et al., 2005). All 10 steps of the translation procedure were implemented as follows: 1. Preparation; 2. Forward Translation; 3. Reconciliation; 4. Back Translation; 5. Back Translation Review; 6. Harmonization; 7. Cognitive Debriefing; 8. Review of Cognitive Debriefing Results and Finalization; 9. Proofreading and 10. Final Report. The completed translation was then submitted to OQ Measures, LLC, which had requested strict adherence to this process. After reviewing the final report, the provider granted formal approval for the use of the Greek version. Although validation was not a

prerequisite for this approval, the researchers informed OQ Measures of the planned validation study and received methodological input based on their expertise.

Concerning the study, after researchers were granted access to a Systemic/Family psychotherapy institute, they informed instructors both orally and in writing regarding the scope and procedures of the research. Prospect participants were psychotherapy trainees and were approached by their instructors at the time and location of their classes and were asked to participate in the study on a voluntarily basis. All questionnaires were self-reported and completed online via Google Forms during class. Instructors were responsible for ensuring that interaction among participants was avoided and it was also made sure that trainees understood the requirements and instructions of participation in the study. Prior to the administration of the online questionnaires, standardized oral instructions were provided in a structured format by the instructors. These instructions, which were also available to participants in written form online, clarified the purpose of the study, emphasized the importance of independent completion of the measures, and explicitly prohibited discussion or exchange of information between participants during the assessment. The total duration for the whole battery of questions administered was 15-20 minutes. A personal identification code was requested, so that researchers could be able to match participants' answers for the re-test. Participants were prompted to fill in the re-test questionnaire 15 days later. Finally, it has been explained that there were no correct or wrong answers and participants were encouraged to answer honestly. The wording of all the above explanations and instructions were predetermined and consistent for all participants, and were available for them in written before the completion of questionnaires in Google Forms. After completion, data was extracted to an Excel file, then transformed to an SPSS file and finally stored in an encrypted and password protected USB which was only available to the researchers.

Measures

Demographics. Participants provided basic demographic information, including age and gender, with all responses collected anonymously to ensure confidentiality. Given the sample's relative homogeneity —as all participants were psychotherapy trainees with academic backgrounds in psychology, social work, or related fields— additional items were included to capture their therapeutic experience. Specifically, participants were asked whether they had ever sought help from a mental health professional, the total duration of such experiences, and the length of their participation in psychotherapy groups.

The Outcome Questionnaire - OQ-45.2 (Lambert et al., 2013). The OQ-45.2 is a self-report questionnaire consisting of 45 items, designed to measure general distress and general functioning. The OQ-45.2 is specifically designed to assess and monitor a client's progress throughout psychotherapy. The OQ–45 is the first outcome system which was recognized as an evidence-based measure in the United States of America National Registry of Evidenced (NRE) based programs and practice. Permission to use the latest version of the questionnaire and guidelines for its translation, administration and validation was requested and provided by OQ Measures, LLC. The questionnaire utilizes a 5-point Likert scale that ranges from 0 (never) to 4 (almost always) and typically takes around 5 minutes to complete across three distinct dimensions: a) Subjective Discomfort, b) Interpersonal Relationships, and c) Social Role Performance. Examples of characteristic items for each dimension include "I feel lonely" for Subjective Discomfort, "I have trouble getting along with friends and close acquaintances" for Interpersonal Relationships and "I feel angry enough at work/school to do something I might regret" for Social Role Performance.

The OQ-45.2 total score ranges from 0 to 180, with higher scores suggesting greater levels of distress and scores at or above 64 indicate a status alert (clinical distress). It incorporates 9 positively worded items, which are reversed-scored in addition to the rest of the symptom distress items that are included in the scale. Moreover, according to its manual (Lambert et al, 2013, p.7), the OQ-45.2 is designed to be used as a baseline-screening instrument for treatment assignment decisions and for treatment planning, but is not designed to be used for

patient diagnosis. Its validity across many different populations and diagnoses has been tested and affirmed widely over the last four decades.

The Greek version of the World Health Organization Quality-of-Life Scale - WHOQOL-BREF (Skevington, 2004; Ginieri-Coccossis et al., 2012). The WHOQOL-BREF (Skevington, 2004), is a shorter version of the WHOQOL-100 (Power & Kuyken, 1998), developed with the aim of monitoring the outcome (appraisal and reappraisal) of care interventions cross culturally. It is included in this research as a measure of criterion validity of the OQ-45.2. The WHOQOL-BREF consists of 24 items corresponding to 24 QoL (thematic) facets, and two items comprising an overall quality of life/general health rating (OQOL). Items are organized into four domains: physical health psychological health, social relationships and environment. It has been cross culturally tested in 23 countries and has been validated in the Greek language by Ginieri-Coccossis et al. (2012). Mean scores of items within each domain are used. The Greek version of the instrument consists of 30 items (26 from the original English version, plus four national items selected from a pool of 23 national items produced by focus groups). Permission to use the WHOQOL-BREF Greek version was requested and given by the Department of Psychiatry of the University of Athens, Aiginiteio University Hospital.

Data Analysis

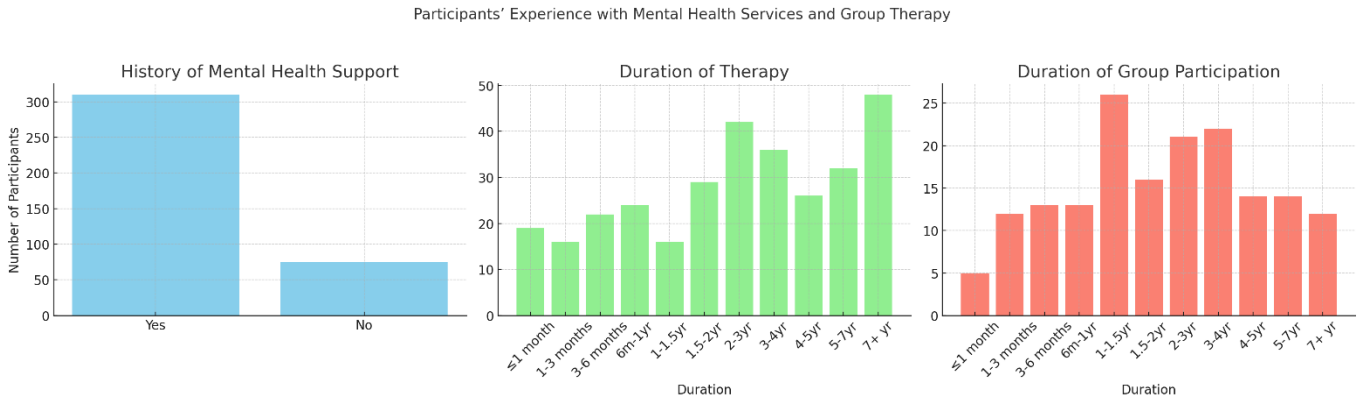
The scores for all measures were calculated according to the instructions given in the respective manuals. Means and standard deviations were calculated, and different demographic and other sample sub-groups were examined to investigate possible differences. The statistical analyses were conducted using the SPSS Statistics version 29.0 for Windows, with a significance level of $p < .05$. Construct validity for the OQ-45.2-GR was assessed through confirmatory factor analysis (CFA) in order to try and confirm the factorial structure of the original English version, using the R project software and utilizing lavaan package (Rosseel, 2012). The Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR) were used to evaluate the overall goodness-of-fit of the model. According to Hu and Bentler (1999), typically, CFI and TLI values equal to or greater than .95, RMSEA values less than or equal to .06 and SRMR values less than or equal to .08 indicate a good fit. Additionally, a statistically non-significant chi-square to degrees of freedom ratio (χ^2/df) also suggests a good fit. The internal consistency of all scales as well as subscales were assessed using Cronbach's α coefficient. Test-retest reliability as well as convergent and criterion validity of all measures were determined using the Pearson correlation. Prior to all inferential investigation, the Kolmogorov-Smirnov Test of normality was conducted to assess parametricity.

Results

Participants' Experience with Mental Health Services and Group Therapy

Most participants (80.5%) reported having sought help from a mental health professional at some point in their lives. The reported duration of therapy ranged from less than one month to over seven years, with 15.5% indicating ongoing involvement of seven years or more. Regarding group therapy, 43.6% of the participants reported participation, with durations also ranging broadly, and 15.5% indicating 1–1.5 years of experience. These findings are illustrated in Figure 1.

Figure 1. *Participants' experience with mental health services and group therapy.*



**Note.* The first panel shows the proportion of participants who have sought help from a mental health professional at some point in their lives. The second panel displays the duration of individual engagement in therapy, while the third panel shows the duration of participation in psychotherapy groups

Construct Validity and Confirmatory Factor Analysis

The CFA model was specified to include a predetermined number of factors and their corresponding indicators. Specifically, the hypothesized model comprised of three factors, each represented by a set of observed items (Figure 2). The specified CFA model was estimated using the Weighted Least Square Mean and Variance Adjusted method. The Robust CFA in a three-factor model (Table 1) provided satisfactory fit indices. Specifically, in our three-factor model, χ^2/df was 10.6 (p = .001), CFI= .863, TLI= .856, RMSEA= .060, and SRMR= .091.

Descriptive Indices and Internal Consistency

Descriptive statistics for the OQ-45.2-GR and the WHOQOL-BREF are presented in Table 2. These include means, standard deviations, and score ranges for the total score and the three subscales of the OQ-45.2-GR—Symptom Distress, Interpersonal Relationships, and Social Role Performance—as recommended by Lambert et al. (2013). Corresponding indices for the WHOQOL-BREF domains (Physical Health, Psychological Health, Social Relationships, and Environment) are also reported.

The internal consistency for the total score of the OQ-45.2-GR was found to be excellent (α = .912). Similarly, a moderate coefficient was observed for the Symptom Distress (α = .897) subscale as well as for the Interpersonal Relationships (α = .773) subscale. The Social Role Performance subscale demonstrated a low coefficient (α = .577), which may be attributed to the small number of items comprising this scale (Ponterotto & Ruckdeschel, 2007).

Table 1. *Standard and Robust Confirmatory Factor Indices for the OQ-45.2-GR*

	Standard				**Robust**			
Model	CFI	TLI	RMSEA	SRMR	CFI	TLI	RMSEA	SRMR
3 factor	.941	.938	.073	.091	.863	.856	.060	.091
2 factor	.930	.927	.080	.095	.842	.834	.064	.095
1 factor	.924	.921	.083	.097	.829	.820	.067	.097

Figure 2. *Factor structure of the OQ-45.2-GR*

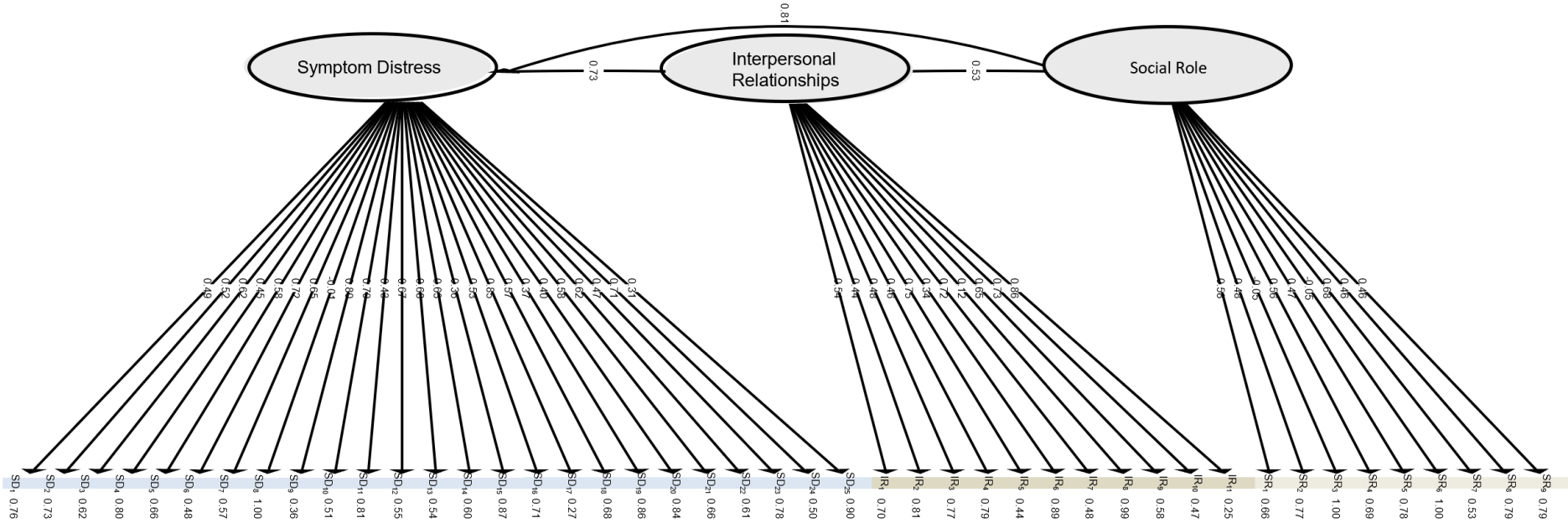


Table 2. *Descriptive Indices for the OQ-45.2-GR and WHOQOL-BREF in the Total Sample*

	Mean ±SD	**Range**
OQ-45.2-GR Symptom Distress	30.71 ±11.21	3 - 80
OQ-45.2-GR Interpersonal Relationships	13.15 ±5.17	1 - 28
OQ-45.2-GR Social Role Performance	11.19 ±3.54	2 - 23
OQ-45.2-GR Total Score	55.06 ±17.09	7 - 125
WHOQOL-BREF Physical Health	77.67 ±13.21	35.71 - 100
WHOQOL-BREF Psychological Health	67.10 ±13.58	12.50 - 91.66
WHOQOL-BREF Social Relationships	67.33 ±17.64	1 - 100
WHOQOL-BREF Environment	67.79 ±11.16	21.86 - 96.88

**Note*. OQ-45.2-GR=Outcome Questionnaire; WHOQOL-BREF=World Health Organization Quality of Life-Bref; Total sample size: *N* = 385

Convergent and Criterion Validity

The next step in the analysis was to assess the correlations between the OQ-45.2-GR scores and the subscale scores of the WHOQOL-BREF. As expected, all OQ-45.2-GR subscales were significantly negatively correlated with WHOQOL-BREF domains, indicating that higher psychological distress was associated with lower perceived quality of life. The analysis between the OQ-45.2-GR domains revealed significant correlations and this was also the case between the OQ-45.2-GR total score and WHOQOL-BREF subscales (Table 3).

Table 3. *Correlations Between OQ 45.2-GR and WHOQOL-BREF*

	1	2	3	4	5	6	7	8
1. OQ 45.2-GR Symptom Distress	1.00	.551**	.607**	.944**	-.626**	-.689**	-.325**	-.387**
2. OQ 45.2-GR Interpersonal Relationships		1.00	.379**	.745**	-.336**	-.499**	-.677**	-.261**
3. OQ 45.2-GR Social Role Performance			1.00	.711**	-.525**	-.456**	-.280**	-.333**
4. OQ 45.2-GR Total Score				1.00	-.624**	-.697**	-.478**	-.397**
5. WHOQOL-BREF Physical Health					1.00	.569**	.339**	.382**
6. WHOQOL-BREF Psychological Health						1.00	.434**	.383**
7. WHOQOL-BREF Social Relationships							1.00	.148**
8. WHOQOL-BREF Environment								1.00

**Note*. OQ.45.2-GR=Outcome Questionnaire; WHOQOL-BREF=World Health Organisation Quality of Life-Bref; Total sample size: *N* = 385

** Correlation is significant at the 0.01 level (2-tailed)

Test-retest Reliability

At least two weeks later, the OQ-45.2-GR along with the WHOQOL-BREF, were re-administered to a randomly selected subgroup of the original sample (n = 49) in order to assess test-retest reliability (Table 4). Re-administering both instruments allowed us to evaluate the stability of participants' responses and the consistency of the relationship between the two measures over time. The results indicated that the test-retest reliability of the OQ-45.2-GR total score was satisfactory.

Table 4. *Test-retest Reliability*

	Pearson's r
OQ-45.2-GR Symptom Distress	.560**
OQ-45.2-GR Interpersonal Relationships	.666**
OQ-45.2-GR Social Role Performance	.441**
OQ-45.2-GR Total Score	.638**
WHOQOL-BREF Physical Health	.508**
WHOQOL-BREF Psychological Health	.561**
WHOQOL-BREF Social Relationships	.572**
WHOQOL-BREF Environment	.228

**Note*. OQ-45.2-GR=Outcome Questionnaire; WHOQOL-BREF=World Health Organisation Quality of Life-Bref; Test-retest sample size: *N* = 49

** Correlation is significant at the 0.01 level (2-tailed)

Discussion

The primary objective of this study was to examine the quintessential psychometric properties of the OQ-45.2 (Lambert et al., 2013) and more specifically, its factor structure, criterion, convergent validity, internal consistency and reliability. This study stands as a pioneering effort in validating the OQ-45.2 within a non-Anglo-Saxon cultural milieu. This was achieved through the careful translation of the original English version of the OQ-45.2 and a thorough examination of the psychometric properties of its Greek iteration, thereby offering an invaluable instrument for evaluating the nuances of the therapeutic process.

Despite its widespread application and comprehensive coverage across multiple dimensions, the OQ-45.2, with its sophisticatedly concise three-factor structure, had not been previously validated in the Greek language. The English version of the OQ-45.2, characterized by its three-factor structure encompassing Subjective Discomfort, Interpersonal Relationships, and Social Role Performance, has been widely recognized for its satisfactory internal and test-retest reliability (Lambert et al., 2013). Several studies have confirmed the three-factor structure of the OQ-45.2, while cross-cultural validations, such as the comparison of the Duch and American versions (de Jong et al., 2007) and between the Hebrew and Arabic versions (Raz Gross et al., 2015), have further reinforced the robustness of this structure. Moreover, subsequent research including studies by Bludworth et al. (2010) and Kim et al. (2010), have supported the stability of the Symptom Distress, Interpersonal Relations, and Social Role Performance factors across clinical and non-clinical samples. The findings from our study revealed that the Greek version mirrors this factor structure and maintains comparable psychometric properties, thereby endorsing this tool as a valid and reliable instrument for research in Greek-speaking samples. Specifically, the results from CFA unambiguously revealed a coherent three-factor structure, in line with extant research on the OQ-45.2 (Lambert et al., 2013). These three facets of functioning exhibited moderate positive correlations amongst themselves. The CFI, along with other indices like TLI, RMSEA, and SRMR, endorsed a three-factor model with satisfactory fit indices.

Moreover, the study indicated that the OQ-45.2-GR exhibits noteworthy convergent and criterion validity. Notably, significant negative correlations were observed between the OQ-45.2-GR and the WHOQOL-BREF, particularly in the domains of physical and psychological health, underscoring the intricate interplay between a client's challenges and their overall well-being. However, these associations may also be influenced by cultural and contextual factors specific to the Greek population. Greek society places a strong emphasis on close-knit family and social relationships; disruptions in these areas may carry a heightened emotional burden, thereby affecting self-reported measures of distress and well-being (Ginieri-Coccossis et al., 2012; Kafetsios, 2004). Additionally, the prevalent stigma surrounding mental health issues in Greece can influence individuals' willingness to report psychological distress and seek help, potentially impacting the validity of self-reported assessments (Tzouvara & Papadopoulos, 2014; Tzouvara et al., 2016). Generally, the total score and all three factors of the OQ-45.2-GR negatively correlated with every aspect of the WHOQOL-BREF, corroborating the findings of Burlingame et al. (2018) regarding the high concurrent validity of the OQ-45.2 with other outcome measures.

Moreover, previous studies have demonstrated the concurrent validity of the OQ-45.2 by reporting significant correlations with established outcome measures such as the Beck Depression Inventory (Beck & Steer, 1993) and the Symptom Checklist-90-Revised (Derogatis, 1977). The test-retest reliability results align with previous evidence (Lambert et al., 2013), attesting to the measure's satisfactory short-term stability. In this study, the level of consistency between the test and retest phases was commendable (test-retest reliability at .64 for the total score), albeit slightly lower than in prior studies (e.g., Lambert et al., 2013).

Overall, the OQ-45.2-GR is distinguished by its sound psychometric properties and high reliability. The internal consistency of the total score of the OQ-45.2-GR, as well as its three subscales, was found to be satisfactory in this study, albeit with a moderate coefficient observed for the social role subscale, potentially influenced by the limited number of items constituting this scale. This observation aligns with findings from other language adaptations of the instrument. For example, the Hungarian version reported a Cronbach's α of .67 for this subscale (Matavovszky et al., 2024), while the Polish version showed values between .78 and .80 (Simon et al., 2015). Such results suggest that this pattern may reflect an inherent limitation of the subscale, rather than a translation-related issue. The Social Role scale encompasses a diverse range of social functioning domains, which may introduce conceptual variability and reduce internal consistency. Cultural differences in how social role performance is interpreted could further contribute to response variation (Simon et al., 2013). Supporting evidence for the OQ-45.2 robust psychometric properties and reliability is also found in its Italian version, which demonstrated an acceptable Cronbach's α of .82 (Lo Coco et al., 2008, as cited in Kivlighan et al., 2017), and in the Kivlighan et al. (2017) study, where the OQ-45.2 total score's Cronbach's α was reported as .92.

The availability of a psychometrically robust measure of therapeutic outcome can profoundly enhance the holistic approach to therapy. However, this study is not without limitations. All participants were psychotherapy trainees enrolled in a specific systemic/family therapy training program at a single center in Greece. While this sampling strategy was appropriate for initial validation, it limits the generalizability of the findings to broader populations, such as practicing clinicians, individuals from other therapeutic orientations, and clinical groups. Moreover, the test-retest reliability analysis was conducted on a relatively small subsample (n = 49), which may affect the stability of those estimates. The absence of participants with diagnosed mental health conditions also limits the evaluation of the instrument's clinical sensitivity and applicability.

In light of these limitations and the observed psychometric patterns—particularly the moderate reliability of the Social Role subscale—future research should aim to replicate the validation of the OQ-45.2-GR using larger, more diverse, and clinically representative samples from varied training backgrounds and therapeutic settings. In addition, exploring cultural influences—through qualitative methods or cross-cultural comparisons—could provide deeper insight into how contextual factors shape self-reported outcomes and item interpretation.

Furthermore, the study did not assess measurement invariance across key demographic subgroups (e.g., gender or age), which limits conclusions about the scale's equivalence and comparability across populations. Future studies should include invariance testing to examine whether the factor structure holds consistently across diverse groups. These steps would enhance the scale's psychometric validation and inform the development of culturally appropriate clinical cut-off scores.

Finally, the study did not collect data on participants' specific professional roles (e.g., psychologist, social worker) or years of work experience. Although the sample was relatively homogeneous, gathering such information in future studies could support more detailed subgroup analyses and increase the practical relevance of findings.

In conclusion, the development of a valid screening tool capable of estimating therapeutic outcomes and detecting early signs of challenges during the therapeutic process is paramount. In an era where psychotherapy is becoming increasingly widespread, the need for valid, culturally adapted assessment tools that are carefully developed and applied in clinical practice is more critical than ever. Thus, this study achieved its fundamental goal to translate and validate the original English version of the OQ-45.2 for the Greek population. Overall, the findings affirmed that the factorial structure of the Greek version resonates with that of the original questionnaire and that the OQ-45.2-GR is a valid and reliable instrument for studying therapeutic outcomes within the Greek cultural context.

References

Beck, A. T., & Steer, R. A. (1993) *Beck Depression Inventory Manual*. Psychological Corporation.

Bludworth, J. L., Tracey, T. J., & Glidden-Tracey, C. (2010). The bilevel structure of the Outcome Questionnaire-45. *Psychological assessment*, *22*(2), 350–355. https://doi.org/10.1037/a0019187

Brosseau-Liard, P. É., Vandette, M.-P., Jamshidi, P., Kogan, C. S., & Aubry, T. (2020). Propriétés psychométriques de la version française (Mesure d'impact; MI-45) du Outcome Questionnaire-45 (OQ-45) en milieu clinique et universitaire [Psychometric properties of the French version (Measure of Impact; MI-45) of the Outcome Questionnaire-45 (OQ-45) in clinical and university settings]. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement, 52*(1), 78–83. https://doi.org/10.1037/cbs0000138

Burlingame, G. M., & Beecher, M. E. (2008). New directions and resources in group psychotherapy: Introduction to the issue. *Journal of Clinical Psychology, 64*(11), 1197–1205. https://doi.org/10.1002/jclp.20534

Burlingame, G. M., Strauss, B. M., & Joyce, A. (2013). Change mechanisms and effectiveness of small group treatments. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change (*pp. 640–689). John Wiley & Sons.

Burlingame, G. M., Whitcomb, K. E., Woodland, S. C., Olsen, J. A., Beecher, M., & Gleave, R. (2018). The effects of relationship and progress feedback in group psychotherapy using the Group Questionnaire and Outcome Questionnaire-45: A randomized clinical trial. *Psychotherapy (Chicago, Ill.)*, *55*(2), 116–131. https://doi.org/10.1037/pst0000133

de Beurs, E., den Hollander-Gijsman, M., Buwalda, V., Trijsburg, W., & Zitman, F. (2005). De Outcome Questionnaire OQ-45 Psychodiagnostisch gereedschap [The Outcome Questionnaire (OQ-45): Measuring psychiatric symptoms and interpersonal functioning]. *Psycholoog, 40*(7-8), 393–400.

de Jong, K., Nugter, M.A., Polak, M.G., Wagenborg, J.E.A., Spinhoven, P. and Heiser, W. (2007). The Outcome Questionnaire-45 in a Dutch population: A cross cultural validation. *Clinical Psychology & Psychotherapy, 14*, 288–301. https://doi.org/10.1002/cpp.529

Derogatis, L. R. (1977). *SCL-90-R: Administration, scoring and procedures manual*. Clinical Psychometric Research.

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, *13*(3), 595–605. https://doi.org/10.1017/S0033291700048017

Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95*(3), 542–575. https://doi.org/10.1037/0033-2909.95.3.542

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical outcomes in routine evaluation. *Journal of Mental Health, 9*(3), 247–255. http://doi.org/10.1080/jmh.9.3.247.255.

Feldman, L. A. (1993). Distinguishing depression and anxiety in self-report: Evidence from confirmatory factor analysis on nonclinical and clinical samples. *Journal of Consulting and Clinical Psychology, 61*(4), 631–638. https://psycnet.apa.org/doi/10.1037/0022-006X.61.4.631

Gillaspy, J. A., Wright, A. R., Campbell, C., Stokes, S., & Adinoff, B. (2002). Group alliance and cohesion as predictors of drug and alcohol abuse treatment outcomes. *Psychotherapy Research, 12*(2), 213–229. https://doi.org/10.1093/ptr/12.2.213

Ginieri-Coccossis, M., Triantafillou, E., Tomaras, V., Soldatos, C., Mavreas, V., & Christodoulou, G. (2012). Psychometric properties of WHOQOL-BREF in clinical and health Greek populations: Incorporating new culture-relevant items. *Psychiatrike = Psychiatriki*, *23*(2), 130–142.

Hatfield, D. & Ogles, B.M. (2004). The use of outcome measures by psychologists in clinical practice. *Professional Psychology, Research and Practice, 35*, 485–491. http://dx.doi.org/10.1037/0735-7028.35.5.485

Horowitz L. M. (1979). On the cognitive structure of interpersonal problems treated in psychotherapy. *Journal of Consulting and Clinical Psychology*, *47*(1), 5–15. https://doi.org/10.1037//0022-006x.47.1.5

Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of interpersonal problems: psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*, *56*(6), 885–892. https://doi.org/10.1037//0022-006x.56.6.885

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huag, S., Puschner, B., Lambert, M. J., & Kordy, H. (2004). Veraderungs messung in der psychotherapie mit dem Ergebnis Frage Bogen (EB-45). Assessment of change in psychotherapy with the German version of the Outcome Questionnaire (OQ-45). *Zeitschrift fur Differenttielle und Diagnosfische Psychologie*, *25*(3), 141–151. https://doi.org/10.1024/0170-1789.25.3.141

Kafetsios, K. (2004). Attachment and emotional intelligence abilities across the life course. *Personality and Individual Differences, 37*(1), 129–145. https://doi.org/10.1016/j.paid.2003.08.006

Kim, S.-H., Beretvas, S. N., & Sherry, A. R. (2010). A validation of the factor structure of OQ-45 scores using factor mixture modeling. *Measurement and Evaluation in Counseling and Development, 42*(4), 275–295. https://doi.org/10.1177/0748175609354616

Kivlighan, D. M., Jr., Lo Coco, G., Oieni, V., Gullo, S., Pazzagli, C., & Mazzeschi, C. (2017). All bonds are not the same: A response surface analysis of the perceptions of positive bonding relationships in therapy groups. *Group Dynamics: Theory, Research, and Practice, 21*(3), 159–177. https://doi.org/10.1037/gdn0000071

Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy, 50*(1), 42–51. https://doi.org/10.1037/a0030682

Lambert, M. J., & Barley, D. E. (2001). Research summary on the therapeutic relationship and psychotherapy outcome. *Psychotherapy: Theory, Research, Practice, Training, 38*(4), 357–361. https://doi.org/10.1037/0033-3204.38.4.357

Lambert, M. J., Kahler, M., Harmon, C., Burlingame, G., Shimokawa, K., & White, M. (2013). *Administration and scoring manual outcome questionnaire– 45*. OQ Measures.

Laska, K. M., Gurman, A. S., & Wampold, B. E. (2014). Expanding the lens of evidence-based practice in psychotherapy: A common factors perspective. *Psychotherapy*, *51*(4), 467–481. https://doi.org/10.1037/a0034332

Li, Y.-J., & Luo, H. (2009). The reliability and validity of the Outcome Questionnaire-Chinese version. *Chinese Mental Health Journal, 23*(2), 105–107.

Lo Coco, G., Chiappelli, M., Bensi, L., Gullo, S., Prestano, C. & Lambert, M. J. (2008). The factorial structure of the Outcome Questionnaire-45: A study with an Italian sample. *Clinical Psychology & Psychotherapy. 15*(6), 418–23. https://doi.org/10.1002/cpp.601

Lo Coco, G., Gullo, S., & Kivlighan, D. M. (2012). Examining patients' and other group members' agreement about their alliance to the group as a whole and changes in patient symptoms using response surface analysis. *Journal of counseling psychology, 59*(2), 197–207. https://doi.org/10.1037/a0027560

Matavovszky, D., Nguyen Luu, L. A., & Karner, O. (2024). Development of an instrument to assess the mental health of university students: validation of the Outcome Questionnaire-45 in a Hungarian sample. *Frontiers in Psychology, 15*, Article 1334615. https://doi.org/10.3389/fpsyg.2024.1334615

Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and Motor Skills, 105*(3), 997–1014. https://doi.org/10.2466/pms.105.3.997-1014

Power, M., & Kuyken, W. (1998). World Health Organization Quality of Life Assessment (WHOQOL): Development and general psychometric properties. *Social Science and Medicine, 46*(12), 1569–1585. https://doi.org/10.1016/S0277-9536(98)00009-4

Raz Gross, M. D., Glasser, S., Jacobson, D. M., Levitan, G., & Ponizovsky, A. M. (2015). Validation of the Hebrew and Arabic versions of the Outcome Questionnaire (OQ-45). *Israel Journal of Psychiatry, 52*(1), 33.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Simon, W., Śliwka, P., Sobański, J. A., Klasa, K., Sala, P., Żak, W., Busath, G., & Lambert, M. J. (2015). The orthogonal-oblique bi-level model of the Outcome Questionnaire (OQ-45.2): Polish adaptation based on factor analysis. *Psychiatria polska, 49*(5), 1043–1070. https://doi.org/10.12740/PP/59064

Skevington, S. M., Lotfy, M., & O'Connell, K. A. (2004). The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial a report from the WHOQOL Group. *Quality of Life Research, 13*(2), 299–310. https://doi.org/10.1023/B:QURE.0000018486.91360.00

Takara, R., Beecher, M. E., Okiishi, J. C., Shimokawa, K., Lambert, M. J., & Griner, D. (2017). Translation of the Outcome Questionnaire-45 (OQ) into Japanese: A cultural adaptation. *Psychotherapy Research, 27*(2), 154–166. https://doi.org/10.1080/10503307.2015.1080876

Tzouvara, V., & Papadopoulos, C. (2014). Public stigma towards mental illness in the Greek culture. *Journal of Psychiatric and Mental Health Nursing, 21*(10), 931–938. https://doi.org/10.1111/jpm.12146

Tzouvara, V., Papadopoulos, C., & Randhawa, G. (2016). Systematic review of the prevalence of mental illness stigma within the Greek culture. *The International Journal of Social Psychiatry, 62*(3), 292–305. https://doi.org/10.1177/0020764016629699

Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR task force for translation and cultural adaptation. *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research, 8*(2), 94–104. https://doi.org/10.1111/j.1524-4733.2005.04054.x

Yusof, Y., & Carpenter, J. (2014). The Impact of Family Therapists' Adult Attachment Styles on Their Career Choice and Approach to Therapy: An Interpretive Phenomenological Analysis. *Journal of Social Work Practice, 29* (4), 395–412. https://doi.org/10.1080/02650533.2014.922059

Zautra, A. J. (1983). Social resources and the quality of life. *American Journal of Community Psychology, 11*(3), 275–289. https://doi.org/10.1007/BF00893368

Ελληνική Ψυχολογική Εταιρεία
Hellenic Psychological Society

ΕΜΠΕΙΡΙΚΗ ΕΡΓΑΣΙΑ | RESEARCH PAPER

Μετάφραση και στάθμιση της Κλίμακας Μέτρησης Αποτελεσματικότητας 45.2 στα Ελληνικά (OQ-45.2-GR/ΚΛΙΜΑ-45.2)

Λητώ ΚΑΤΑΚΗ[1,2], Πάρις ΒΟΓΑΖΙΑΝΟΣ[1], Παναγιώτης ΠΑΡΠΟΤΤΑΣ[1], Μαριάννα ΑΝΑΓΝΩΣΤΟΥ[3], Μαρία ΧΡΥΣΙΚΟΠΟΥΛΟΥ[2], Έλενα ΔΡΑΓΚΙΩΤΗ[4], Γιώργος ΤΣΙΤΣΑΣ[3]

[1] Τμήμα Κοινωνικών Επιστημών και Επιστημών Συμπεριφοράς, Ευρωπαϊκό Πανεπιστήμιο Κύπρου
[2] Εργαστήριο Διερεύνησης Ανθρώπινων Σχέσεων
[3] Κέντρο Ψυχολογικής και Συμβουλευτικής Υποστήριξης, Χαροκόπειο Πανεπιστήμιο
[4] Εργαστήριο Ψυχολογίας Ασθενών, Οικογενειών και Επαγγελματιών Υγείας, Τμήμα Νοσηλευτικής/Σχολή Επιστημών Υγείας, Πανεπιστήμιο Ιωαννίνων

KEYWORDS IN GREEK

Κλίμακα Μέτρησης Αποτελεσματικότητας
Μετάφραση
Ελληνικά
Επιβεβαιωτική ανάλυση παραγόντων
Εγκυρότητα
Αξιοπιστία

CORRESPONDENCE

Παναγιώτης Παρπόττας
Ευρωπαϊκό Πανεπιστήμιο Κύπρου
Διογένους 6, 2404 Έγκωμη, Λευκωσία, Κύπρος
p.parpottas@euc.ac.cy

ABSTRACT IN GREEK

Η παρούσα έρευνα είχε ως στόχο την μετάφραση και στάθμιση της αγγλικής έκδοσης της Κλίμακας Μέτρησης Αποτελεσματικότητας (OQ-45.2) προς χρήση της στον ελληνικό πληθυσμό. Το δείγμα αποτέλεσαν 385 εκπαιδευόμενοι ψυχοθεραπευτές, (102 άντρες και 283 γυναίκες) ηλικίας 18 έως 65 χρόνων, προερχόμενοι από κέντρο εκπαίδευσης στη συστημική/οικογενειακή ψυχοθεραπεία στην Ελλάδα. Οι συμμετέχοντες συμπλήρωσαν την OQ-45.2-GR/ΚΛΙΜΑ-45.2 και την ελληνική έκδοση της Κλίμακας Ποιότητας Ζωής (WHOQOL-BREF-GR), όπως και αριθμό ανώνυμων δημογραφικών στοιχείων. Η Επιβεβαιωτική ανάλυση παραγόντων (CFA) κατέδειξε ότι η ελληνική έκδοση της OQ-45.2 αποτελούνταν από τρεις παράγοντες όπως η αγγλική της έκδοση, με επαρκείς δείκτες προσαρμογής (χ^2/df, p = .001, RMSEA = .060, and SRMR = .091). Επιπλέον, η κλίμακα επιδεικνύει καλές ψυχομετρικές ιδιότητες με υψηλή αξιοπιστία εσωτερικής συνέπειας με Cronbach α = .912 και αποδεκτή αξιοπιστία ελέγχου-επανελέγχου r = .64, p < .01. Επιπλέον, η αξιολόγηση της συντρέχουσας εγκυρότητας της OQ-45.2-GR/ΚΛΙΜΑ-45.2 με το WHOQOL-BREF-Gr επαλήθευσε την αναμενόμενη αρνητική συσχέτιση μεταξύ του συνολικού σκορ και των υποκλιμάκων των δύο ερωτηματολογίων (από r = -.29, p < .01, έως r = -.69, p < .01). Η προκαταρκτική αξιολόγηση της OQ-45.2-GR/ΚΛΙΜΑ-45.2 καταδεικνύει αποδεκτή αξιοπιστία και εγκυρότητα, ενώ θα μπορούσε να χρησιμοποιηθεί και σε μεγαλύτερα δείγματα στον ελληνικό πληθυσμό. Η OQ-45.2-GR/ΚΛΙΜΑ-45.2 εστιάζεται στην αξιολόγηση του ψυχοθεραπευτικού αποτελέσματος, ενώ μπορεί να χρησιμοποιηθεί ως εργαλείο ανίχνευσης των αναγκών των θεραπευομένων στην αρχή της θεραπείας και να συμβάλει στη λήψη αποφάσεων όπως και στην περεταίρω αξιοποίηση του θεραπευτικού σχεδιασμού. Η διαθεσιμότητα της OQ-45.2-GR/ΚΛΙΜΑ-45.2 στην ελληνική γλώσσα αναμένεται να βελτιώσει τις πρακτικές κλινικής αξιολόγησης, να συμβάλει σε τεκμηριωμένες πρακτικές θεραπευτικού προγραμματισμού και να προάγει περεταίρω την έρευνα στις διαδικασίες και την αποτελεσματικότητα της ψυχοθεραπείας στην Ελλάδα.

https://doi.org/10.12681/psy_hps.39687

Ψυχολογία: Το περιοδικό της Ελληνικής Ψυχολογικής Εταιρείας
https://ejournals.epublishing.ekt.gr/index.php/psychology