Scientific Research Publishing

Open Journal of Social Sciences, 2021, 9, 240-256
https://www.scirp.org/journal/jss
ISSN Online: 2327-5960
ISSN Print: 2327-5952

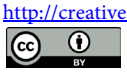
Assessment of Migrants', Refugees' and Asylum Seekers' Hard Skills: Cultural Adaptation and Psychometric Properties of the NADINE Hard Skill Tests

Maria Tountopoulou[1], Nikos Drosos[1,2], Fotini Vlachaki[1]

[1]Ison Psychometrica, Athens, Greece
[2]European University Cyprus, Nicosia, Cyprus
Email: martounto@yahoo.gr

How to cite this paper: Tountopoulou, M., Drosos, N., & Vlachaki, F. (2021). Assessment of Migrants', Refugees' and Asylum Seekers' Hard Skills: Cultural Adaptation and Psychometric Properties of the NADINE Hard Skill Tests. *Open Journal of Social Sciences, 9,* 240-256.
https://doi.org/10.4236/jss.2021.97017

Received: June 9, 2021
Accepted: July 18, 2021
Published: July 21, 2021



Abstract

The present paper aims at presenting the development of the NADINE Hard Skill Tests for migrants, refugees and asylum seekers. The NADINE tests are assessing the following skills: Observation, Language learning aptitude, Numeracy, Accuracy, Mechanical, ICT skill, Decision making, Problem solving, Dexterity, Spatial ability, Cognitive/analytical skills and Technical skills. The paper focuses on the linguistic and cultural adaptation and equivalency of the English, Farsi, French, Arabic and Spanish versions of the tests. To determine the degree to which the items, the response scale and the instructions are comprehensible and culturally relevant to the target population, four focus groups were formed, rating all items in terms of clarity and cultural relevance. The review resulted in minor changes. Furthermore, the Farsi, Arabic and French versions of the tests were administered to 120 migrants/refugees/asylum seekers and we calculated the difficulty level of each item and its contribution to the scale's internal reliability coefficient. Items with extremely low or high difficulty level and items that reduced reliability were eliminated. The remaining items had difficulty levels between 0.20 and 0.80 (with few exceptions), while all reliability coefficients were above 0.70 (the vast majority being above 0.80). Finally, the new versions of the test were administered to 427 participants and factor analyses were conducted to ensure that each test was measuring only one factor. Moreover, we calculated Tucker's congruence coefficient for each test to ensure cross-cultural equivalency. All Tucker's phi coefficients were above the 0.90 cut-off point. Results of this study support the case of linguistic and cultural equivalence between the different language versions of the tests, which constitute the NADINE Hard Skill Tests valid and re-

liable instruments that can be used by career practitioners to accurately assess the skills of migrants, refugees and asylum seekers.

Keywords

Cultural Adaptation, Hard Skills' Assessment, Refugees, Migrants, Work Integration

1. Introduction

Since 2015, Europe experienced a major migrant/refugee crisis characterised by a dramatic increase of migrants and refugees crossing its borders seeking asylum (Clayton, 2015). The vast majority of the newcomers were fleeing war conflicts in Syria, Afghanistan, Libya or other areas that posed a major threat for their life and integrity (Cosgrave, Hargrave, Foresti, & Massa, 2016). Although the sense of emergency for the refugee crisis that was strongly felt in 2015-2019 has dissipated (Kushnir, Kilkey, & Strumia, 2020; Rankin, 2019), more than 2.5 million people have entered EU (Cosgrave et al., 2016) and they need to be integrated the European labour market. Furthermore, the EU is still recovering from the aftermaths of the 2008 financial crisis with some countries having more than 15% unemployment rate (Greece and Spain) (Statista, 2021). Moreover, the EU has to anticipate the impact of the United Kingdom's withdrawal from the Union and more importantly the consequences of the current pandemic in the labour market. The International Monetary Fund has characterised the pandemic period as the "Great Lockdown" highlighting that its economic impact is expected to be analogous to the Great Depression and far graver than the 2008 financial crisis (Gopinath, 2020). This interaction between the refugee crisis and the other crises has put migrants, asylum seekers and refugees among the social groups that are most impacted by the pandemic and suffer disproportionately from its economic and psychosocial consequences (UN, 2020).

Work integration of migrants and refugees is essential for achieving social justice and it is recognised as a priority of the European Commission (EC, 2016, 2020) and the International Labour Organization (ILO, 2016). The European Commission has stated that assisting work integration of refugees is paramount for Europe's social cohesion and highlighted the need for accurate skill assessment to ensure the best use of their skills and talents (EC, 2020). Furthermore, it is a prerequisite for achieving the objectives of the Global Compact on Refugees (UN, 2018) and the 2030 Agenda for Sustainable Development (UN, 2015). We should note that migrants/refugees face many additional difficulties in entering the labour market due to limited access to socioeconomic resources (Udayar, Fedrigo, Durante, Clot-Siegrist, & Masdonati, 2020). Being forced to arrive in a new country without prior knowledge of the language and the culture makes their transition very difficult and the need for work integration may become a major source of stress. Some of the most usual obstacles are the loss of previous-

ly held professional credentials, and professional status, and change from professional to service and manual labour (Desiderio, 2016). Career counselling can play a major role in assisting work integration, but requires a good understanding of this population's unique and complex characteristics, and practical tools designed to address their needs.

The development of career assessment tools that are designed specifically for this population is crucial in order to provide career counselling services to vulnerable migrants, asylum seekers and refugees. The NADINE project, a three-year Horizon 2020 innovation project, aims at supporting migrant integration across Europe, and addresses the aforementioned issue by creating a series of tests designed to assess hard skills for migrants (the NADINE hard skills tests), refugees and asylum seekers in the framework of the broader scope to facilitate their labor market integration (see also Tountopoulou, Karaminas, Drosos, & Vlachaki, 2020). Employment and access to resources that facilitate work acquisition are considered as a prerequisite for smooth social inclusion of migrants and refugees in their new societies.

2. The NADINE Hard Skills Tests

2.1. Description and Theoretical Background

The NADINE Hard Skill Tests aim at assessing a vast variety of different skills in order to provide an accurate profile of each migrant/refugee/asylum seeker. Hard skills are strictly job-specific, closely connected with knowledge, easily observed, measured and trained. They constitute the core occupational requirements of a job (CEDEFOP, 2020) and mainly represent cognitive abilities. The "hard skills" constructs' operationalization was the result of thorough literature review and focus groups with employers and migrants/refugees/asylum seekers. Various European and national frameworks on skills were used, such as ESCO skills classification, OECD competency frameworks, definitions and descriptions used at the Survey of Adult Skills (OECD PIAAC), the OECD Learning Framework 2030 (OECD, 2005, 2014, 2018a, 2018b), ILO competency standards (ILO, 2008, 2015, 2018), the EU Key Competencies Framework (2006, 2018), the EU Skills Profile Tool for Third Country Nationals, as well as international surveys and studies on skills needed to enable migrants labour market integration in the host countries (e.g. ILO, 2017, The Adecco Group, 2017a, 2017b).

Additionally, the development of the NADINE Hard Skill Tests was largely based on the Cattell-Horn-Carroll Theory of Cognitive Abilities. The Cattell-Horn-Carroll (CHC) theory is one of the most comprehensive and empirically supported theories regarding the structure of cognitive abilities to date. It represents the integrated works of Raymond Cattell, John Horn, and John Carroll (Alfonso, Flanagan, & Radwan, 2005; Horn & Blankson, 2005; McGrew, 2005; Schneider & McGrew, 2012). Raymond Cattell made a distinction between "fluid" and "crystallized" intelligence. He defined fluid intelligence as the ability to perceive relationships independent of previous specific practice or instruction

concerning those relationships. Fluid intelligence involves being able to think and reason abstractly and solve problems. On the other hand, crystallized intelligence involves knowledge that comes from prior learning and past experiences and therefore is based upon facts, skills, and information learned in school or from past experience (Cattell, 1957, 1971). Horn expanded Catell's theory into an eight-factor model which included: Fluid intelligence (Gf), Quantitative knowledge (Gq), Crystallised intelligence (Gc), Reading/Writing (Grw), Visual processing (Gv), Short-term memory (Gsm), Long-term retrieval (Glr), Processing speed (Gs), Decision speed (CDs) and Auditory processing (Ga) (Horn, 1991; Horn and Blankson, 2005). Finally, Carroll (1997) differentiated factors or abilities into three strata that varied according to the "relative variety and diversity of variables" included at each level. The various G abilities are the most prominent and recognized abilities of the model. They are classified as broad or stratum II abilities and include abilities such as Gf and Gc, the two original factors. Broad abilities represent "basic constitutional and long standing characteristics of individuals that can govern or influence a great variety of behaviors in a given domain", while narrow abilities "represent greater specializations of abilities, often in quite specific ways that reflect the effects of experience and learning, or the adoption of particular strategies of performance". The broadest or most general level of ability in the Gf-Gc model is represented by stratum III, located at the apex of Carroll's (1993, 2005) hierarchy. An integrated Cattell-Horn and Carroll model was proposed by McGrew (1997), while several revisions of the model were made (McGrew & Flanagan, 1998; Flanagan et al., 2000), with the most recent the one being proposed by Schneider and McGrew (2012).

NADINE Tests' development methodology follows the relevant and internationally used scientific statistical and analytical psychometric procedures for psychological testing (Nunnally, 1978, American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999, Society for Industrial and Organizational Psychology, 2003), so as to minimize measurement errors and guarantee the quality, suitability, and usefulness of the tool to be developed. Such standardized psychometric tools are widely used to objectively measure individuals' mental abilities and skills. A standardized test requires all test takers to answer the same questions in the same way, and is scored in a "standard" or consistent manner, which makes it possible to compare the relative performance of individuals. One main consideration was the tests' length. In the first stages we included a large number of questions having in mind that a significant number was likely to be rejected as inappropriate in the next stages. The NADINE tests' first version included more than 10 to 48 items per test with the aim to reduce items for use in the final version after the piloting.

Items were developed taking into consideration the following: 1) thorough review of relevant literature to ensure content adequacy, 2) make sure items ad-

dress a single issue, and 3) make sure statements are culturally appropriate. For each skill to be measured, a working group of experts was composed, consisted of 2 - 3 experts in the field, such as mathematicians, ICT experts, physicians, engineers, neurologists, so as to propose the appropriate items for each relevant field. Moreover, the group included 2 psychologists, 1 sociologist and an expert in psychometrics. The test development team specified an exhaustive list of potential indicators of each scale and the representative sample of indicators was selected from this list. Then, the team proceeded to the development of items that supposedly measure the selected constructs. Each member of the working team proposed several items measuring the construct under evaluation. The next step was to evaluate each of the proposed items in terms of representativeness and content redundancy. The process resulted in a shorter, but comprehensive enough pool of items, following the rule of including 2 or 3 times larger number of items (Devellis, 2017), with the aim to reduce those finally used after the item analysis/factor analysis. Dichotomous and Multiple choice questions are used for different skills evaluation. Dichotomous questions are used for the assessment of Spatial ability and of Accuracy, where the options/given answers are "Different - Same". Multiple choice questions are used for all the other hard skills, with 4 or 5 options/given answers, so as to distinguish successfully those who are less capable in a field and reduce the chance of randomly choosing the right answer (25% or 20%).

Taking into consideration the aforementioned criteria the NADINE team developed 10 tests to evaluate the specific hard skills. Moreover, scores are provided for two combined scales as well. The tests are presented in **Table 1**.

In **Figure 1** and **Figure 2** are presented two examples of test items.

The present study has a two-folded aim: 1) to investigate whether the different language versions of the NADINE hard skill tests are linguistically and culturally equivalent and 2) to assess their psychometric properties. The procedures that we followed to investigate the aforementioned are described below.

Table 1. The NADINE hard skills tests.

Cognitive/Analytical Skills	Observational Test
	Decision Making Test
	Problem Solving Test
Technical Skills	Dexterity Test
	Spatial Ability Test
	Accuracy Test
Language Learning Aptitude Test (only combined score is provided)	Memory
	Phonetic Spelling
	Literacy
Numeracy Test	
Mechanical Test	
ICT Skill Test	

The test that follows consists of small number problems.

Each problem is followed by five choices. You are to choose the correct answer out of the five suggested solutions.

You can do any figuring on a scratch paper. You are not allowed to use a calculator.

Work *as quickly as you can to finish* as many problems as possible.
You will have *5½ minutes.*

Do them for practice:

Add: 194 + 47

- 231
- 241
- 153
- 261
- 243

A student achieved 96 marks out of a possible 120 in an English test.
What percentage mark did the student achieve for the test?

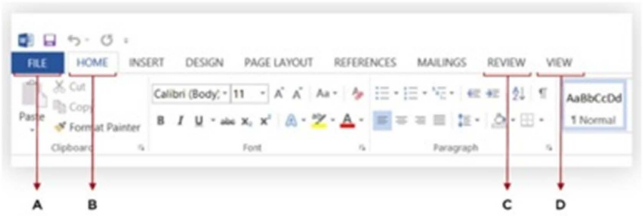
- 0.60
- 0.65
- 0.75
- 0.80
- 0.90

Figure 1. Numeracy test example.

In this test you have to answer to questions referring to the use of ICT.

Here are 2 examples:

1. In order to print a Word document, I have to select:



- A
- B
- C
- D

2. Which two websites offer free email service?
   - Yahoo and Amazon
   - Yahoo and Google
   - I-Tunes and Twitter
   - Library website and Hotmail

Figure 2. ICT test example.

2.2. Items' Review by Expert Groups

Before the translation and adaptation of the tests the items had to undergo a process of review in order to assess the content validity. It is widely acknowledged that the most common and arguably useful form of review is the use of expert groups (DeMaio & Landreth, 2004; Presser & Blair, 1994; Willis, Schechter, & Whitaker, 2000). The expert groups that were assigned for the review of

the NADINE skill assessment tools consisted of different experts, experienced in the field of the construct under measurement, such as experts in ICT, physics, mathematics, language learning, as well as a sociologist, a cognitive psychologist, a neurologist and an organizational psychologist. In any case, 4 to 5 persons, who were experts in each subject and 1 expert in test design participated in each expert focus group. The major concern of "subject matter experts" was to provide input about the initial pool of items regarding their relevance to the content domain, accuracy, & simplicity, and to evaluate potential item bias. The expert in test design was mainly concerned with assessing whether items conform to commonly accepted rules of good item design.

The items' review procedure was as follows: each expert panel member was given a list of the items along with the content dimension they belong. They were asked to evaluate each item on a 4-point scale based on relevance, clarity, and simplicity as 1 (not relevant), 2 (somewhat relevant), 3 (relevant), or 4 (very relevant). Only items that scored a 3 or 4 were considered relevant and therefore used to calculate the actual Content Validity Index (CVI) (Polit & Beck, 2006). The formula of content validity ratio is I-CVR = Ne/N, in which the Ne is the number of panelists indicating "essential" and N is the total number of panelists, while for calculating S-CVI the sum of I-CVIs is divided by the total number of items. The higher score indicates further agreement of panel members on the necessity of an item in an instrument. The results from the experts' panels in scale-level CVIs ranged from 0.95 to 0.99, and item-level CVIs ranged from 0.80 to 1.0, higher than 0.79 which is considered the cut-off point for needing revision. According to the review results, no items were deleted from the developed tools, while only minor revisions were suggested regarding the clarity or wording of the items, which were incorporated into the tools.

3. Linguistic and Cultural Adaptation

3.1. Translation and Adaptation of the Tests

The first phase of linguistic and cultural adaptation involved the translation of NADINE assessment tools from English to the target languages: Farsi, Arabic, French and Spanish. The second phase refers to the item review by target population, while the third phase refers to psychometric evaluation of the different language versions of the tests to ensure cultural equivalence.

First, in order to translate the NADINE questionnaires from English to Farsi, Arabic, French, Greek and Spanish, the working team combined the back-translation design (Hambleton, 2005; Sireci et al., 2006) and the Committee Approach as well (Harkness, Pennell, & Schoua-Glusberg, 2004), in order to benefit from both methods advantages. Particularly, the English version of the test was first translated into the target languages: Arabic, French, Farsi, and Spanish by two translators per language, working individually. The translators were fully proficient in both languages of interest and familiar with the cultures associated with the respective languages of each. Such characteristics enable the

translation process to consider the nuances of the language for which the instrument is intended, which ensures a greater cultural fit of the adaptation process group (Hambleton, 2005). Translators were given the instruction not to pursue the literal translation of items because that often results in incomprehensible statements or rather limited target language fluency. Therefore, they tried to work towards a balanced treatment of linguistic, cultural, and contextual information, so as the people to whom the tools are targeting to be able to understand them in the same way (Tanzer, 2005; Kankaraš & Moors, 2010). To successfully adapt a test in another language, the translators have to ensure the use of concepts, words and expressions that are culturally equivalent in the second language. After the two translators finished the translations of items, a combined translation came out in a way that best represents the views of both of versions. The result of the synthesis of the translations was then given to two different translators, in order to perform the back translation of the items into English following the same instructions. The original English version of the hard-skills assessment tool was then compared to the produced English versions after the back translation in terms of meaning of each item. No revision was needed.

3.2. Cultural Relevance and Review by Migrants, Asylum Seekers and Refugees' Groups

This next stage of the linguistic and cultural adaptation process aimed to verify whether the items, the response scale and the instructions are comprehensible and cultural relevant for the target population. Thus, this procedure aimed to investigate whether the instructions are clear, whether the terms found in the items are appropriate, whether the expressions correspond to those used by the group, and other aspects. For the process of the tests review by the target population, three focus groups were formed with 9, 9, 6 and 7 migrants, asylum seekers and refugees living in Greece. The first group with Farsi speakers, was comprised of 3 men and 6 women coming from Syria, Lebanon and Sudan, while the group of Arabic speakers included 5 men and 2 women from Iran and Afghanistan. The French speaking group consisted of 5 men and 4 women from Congo and Ivory Coast, while for the Spanish speakers 3 were women and 3 men from Venezuela and Mexico. All participants in the focus groups were selected in cooperation with NGOs in Greece working with migrant groups. The responsible persons in the NGOs were informing their "clients" on the process and the scope of it and those who expressed their interest in participating in the focus groups. The process of review was assisted each time by a culture mediator speaking Arabic, Farsi, French of Spanish, so as to facilitate the exchange of ideas and feedback on the tests items.

To obtain rich information about comprehension and cultural relevance of the translated items semi-structured cognitive interviewing was conducted (Oremus et al., 2005). The group discussion focused on clarifying the meaning and relevance of the items' content to the target culture. The semi-structured

cognitive interviewing technique was applied for all items comprising the assessment tools. The set of questions guiding the group discussion elicit: 1) comprehension of the item content (i.e. What does this item mean to you and people of your community?), and 2) cultural relevance of the item content (i.e. Do you and people in your community believe and/or experience what this item reflects? Is the content of the item offensive to people of your community? What words in your language can be used to express the same idea or indicator as in this item?). Additionally, for each item the participants were asked to rate in a 10-point scale the comprehensiveness (comprehension index) and the cultural relevance (cultural relevance index). When a given item was not clear to the participants, the facilitators were encouraged to provide synonyms that best exemplify the vocabulary of the target group.

The review process aimed at ensuring that all items are culturally relevant and clear and at identifying items that require modification to enhance their relevance. Items that are deemed irrelevant or unclear would be omitted. The rule for accepting, modifying or omitting an items was the following: the items that are rated as clear to understand (i.e., comprehension index > 5) and culturally relevant to the target culture (i.e., cultural relevance index ≥ 78%; Polit & Beck, 2006) are maintained in the tool, item modification is necessary if the item was rated as not easy to understand (i.e., comprehension index < 5) but relevant (i.e., cultural relevance index ≥ 78%) and exclusion of items are foreseen if they are rated as not easy to understand (i.e., comprehension index < 5) and not relevant to the target culture (i.e., cultural relevance index ≤ 78%) (McGorry, 2000). The review resulted in just minor changes in the wording of the instructions of the tests and of some items in numeracy and mechanical skills.

3.3. Preliminary Assessment of Psychometric Properties

To assess the psychometric properties of the items we proceeded in calculating the difficulty level of each item and its contribution to the scale's internal reliability coefficient. The sample consisted of 120 asylum seekers, vulnerable migrants and refugees currently residing in refugee structures in Greece: 41 participants were given the Farsi version of the tests, 41 participants the Arabic version and 36 participants the French version. With respect to gender participants 66 (55%) were male and 54 (45%) female. Regarding age 31.7% was 17 - 25, 40% was 26 - 40 and 28.3% was 41+ years old. Moreover, 30.8% had attended only primary education, 36.7% had finished secondary education and 28.3% had a bachelor degree. The participants were selected again in cooperation with NGOs in Greece working with migrant groups. The NGOs having in mind that we need representativeness of the different groups and variables (refugees, vulnerable migrants, asylum seekers speaking the three different languages) chose among their users, those who could provide us with the appropriate characteristics for the process.

We used the Farsi, Arabic and French versions of the NADINE Hard Skill Tests. Dichotomous and Multiple choice questions are used for different skills

evaluation. Dichotomous questions are used for the assessment of Spatial ability and of Accuracy, where the options/given answers are "Different-Same". Multiple choice questions are used for all the other hard skills, with 4 or 5 options/given answers. An impromptu questionnaire consisted of 5 items was used to assess demographic information. The scale was administrated in the refugee camps by the researchers or by camp staff who has been trained by the researchers. Participation was voluntary and this was made clear to participants who gave their written consent. They were asked to contribute in a study about developing valid and reliable career assessment instruments by filling out the research tests. They were instructed to answer all questions that they could.

We calculated item frequencies and descriptive measures to assess whether the items have adequate difficulty level to distinguish between individuals with high levels of the construct being measured from those with lower levels. Moreover, each item's contribution in internal reliability coefficient's of the corresponding scale. Thirty-nine items seemed to have extremely low or high difficulty level or reduced the reliability coefficients and were eliminated. The remaining items had difficulty level between 0.20 and 0.80 (with few exceptions), while all reliability coefficients were above 0.60; and the vast majority above 0.70. The study resulted in revising the NADINE Hard Skill Tests, and the items of each test were put in order according to their difficulty level beginning from the easiest and resulting to the most difficult.

4. Ensuring Cultural Equivalence

A major concern when translating and adapting a test in other languages is cultural equivalence, which is necessary for scores obtained by different language versions of the test to provide comparable scores. Van de Vijver and Leung (1997) refer to structural equivalence, wherein correlations between variables are identical in the different groups. In other words, structural equivalence shows that the same construct is measured across cultures despite using different versions of the same test. Therefore, a significant stage in adapting a new test is performing statistical analysis to determine construct equivalence showing that scores of the adapted versions measure the same construct. Therefore, we proceeded with the check of cross-cultural equivalency.

Subjects

The sample consisted of 427 participants (141 vulnerable migrants, 149 asylum seekers and 137 refugees) currently residing in Greece and Spain: 117 participants were given the Farsi version of the tests, 112 participants the Arabic version, 115 participants the French version and 83 the Spanish version. With respect to gender participants 217 were (50.8%) male and 210 (49.2%) female. Regarding age 29% was 17 - 25, 41% was 26 - 40 and 30% was 41+ years old. Moreover, 30.2% had attended only primary education, 36.8% had finished secondary education and 33% had a bachelor degree. The engagement of participants was made through NGOs working with migrants in Greece and Spain, who were responsible for the organization of the testing administration.

Method & Materials

We used the Farsi, Arabic, French and Spanish versions of the NADINE Hard Skills Tests. The number of items per test is presented in Table 2.

The scales were administrated in the NGOs' premises by the researchers and trained facilitators. Participation was voluntary and this was made clear to participants who gave their written consent. As mentioned before dichotomous questions were used for the assessment of Spatial ability and of Accuracy, where the options/given answers are "Different-Same", and multiple choice questions were used for all the other hard skills, with 4 or 5 options/given answers.

Results

The equivalence of the structure in different cultures was checked by means of calculations of Tucker's phi, which is the congruence index of two sets of factor solutions (van de Vijver & Leung, 1997). Exploratory factor analyses were conducted to categorize the scale's items into larger homogeneous groups, in order to better describe the internal structure of the participants' responses. In Table 3 we present the Factor analyses results for the ICT Skill Test for the three language groups, as an example of the procedure we followed. Principal Component Analysis was used as extraction method with a fixed number of factors. In this example, the final solution (having as the cut point loading the 0.25) revealed the one dimension we expected for all groups. For these analyses we had the following results: Farsi version: KMO = 0.88, χ^2 for the Bartlett's test of sphericity = 455.23 (df = 66, $p < 0.001$); Arabic version: KMO = 0.81, χ^2 for the Bartlett's test of sphericity = 384.64 (df = 66, $p < 0.001$); French version: KMO = 0.87, χ^2 for the Bartlett's test of sphericity = 531.11 (df = 66, $p < 0.001$) and Spanish version: KMO = 0.82, χ^2 for the Bartlett's test of sphericity = 301.28 (df = 66, $p < 0.001$). All tucker's phi coefficients were above the cut-point of 0.90, which indicate cultural equivalency of the four versions of the tests.

Table 2. Number of items per hard skills' test.

Main Scale Subscales		No of items
Cognitive/ Analytical Skills	Observational Test	10
	Decision Making Test	10
	Problem Solving Test	10
Technical Skills	Dexterity Test	15
	Spatial Ability Test	30
	Accuracy Test	48
Language Learning Aptitude Test (only combined score is provided)	Memory	10
	Phonetic Spelling	15
	Literacy	8
Numeracy Test		10
Mechanical Test		10
ICT Skill Test		12

Table 3. Factor analyses (PCA) for the 12 items of the "NADINE" ICT Skills Test according to the language used.

Items	Factors			
	Factor 1 (Farsi language)	Factor 1 (Arabic language)	Factor 1 (French language)	Factor 1 (Spanish language)
- Item 1	**0.47**	**0.40**	**0.60**	**0.53**
- Item 2	**0.54**	**0.59**	**0.61**	**0.52**
- Item 3	**0.58**	**0.36**	**0.54**	**0.67**
- Item 4	**0.60**	**0.55**	**0.48**	**0.52**
- Item 5	**0.55**	**0.53**	**0.60**	**0.72**
- Item 6	**0.67**	**0.53**	**0.58**	**0.59**
- Item 7	**0.76**	**0.65**	**0.71**	**0.69**
- Item 8	**0.78**	**0.73**	**0.77**	**0.64**
- Item 9	**0.59**	**0.61**	**0.57**	**0.40**
- Item 10	**0.62**	**0.59**	**0.77**	**0.68**
- Item 11	**0.75**	**0.78**	**0.82**	**0.70**
- Item 12	**0.64**	**0.61**	**0.65**	**0.59**
Variance explained	40.38	35.63	42.27	37.31
Internal Consistency Coefficients	0.86	0.82	0.87	0.84

Based on the factor analyses, the item loadings and the tucker's phi coefficients we can support structural equivalence among the different language versions of the NADINE Hard Skill Tests. In some tests the tucker's phi was below 0.90 but above 0.80 which indicate similarity. Twenty-four items in total showed poor psychometric properties and were omitted to ensure better results (numeracy test: 2 items, accuracy test: 12 items, mechanical test: 2 items, phonetic spelling: 1 item, spatial ability test: 7 items).

5. Discussion

The present paper aimed at presenting both the NADINE Hard Skill Tests for refugees, migrants and asylum seekers, and how we ensured the linguistic and cultural adaptation for all different language versions of the Tests. We investigated the adaptation procedures, and the structural equivalence of the ten tests. Prior to translating and checking the tests on a sample of each of the target groups, the generated items were subjected to a process of review from expert groups, so as to ensure the content validity of the tests. Multiple methods were used during the translation/adaptation process such as back-translation and the committee approach. Moreover, three focus groups were formed with Farsi-speaking, French-speaking and Arabic-speaking migrants, asylum seekers and refugees. Each scale's items were further reviewed by the three focus groups.

This process aimed at assessing the tests' content and at examining whether the items, the response scale and the instructions are comprehensible and culturally relevant for the target population. The review of the tests resulted in minor changes, while participants agreed that the items were culturally relevant and clear. Thus, the different languages' versions of the NADINE Tests are comprehensible and culturally relevant for the different categories of migrants, asylum seekers and refugees (Arabic, Farsi, Spanish and French speaking).

The psychometric properties of the tests were further investigated in two subsequent studies with 120 and 427 participants respectively. Our findings supported both the existence of one dimension in each test and the cross-cultural equivalency of all Tests' versions. We examined the test items' difficulty level, the reliability and the structure of each test and cross-cultural equivalency. The last study showed that 39 items had poor psychometric properties and were therefore eliminated, resulting in a shorter and more coherent test. Structural equivalence among different language versions of the tests was examined by calculating Tucker's congruence coefficient for each one of the Tests on a sample of 427 subjects. There was strong evidence for cross-cultural equivalency, while all reliability coefficients were above 0.70. The shorter tests are easier and less time-consuming for the target population, while there is no information loss.

Therefore, the different language versions of the NADINE Hard Skill tests are linguistically and culturally equivalent, and constitute valid and reliable instruments to measure migrants', refugees' and asylum seekers' skills.

Work integration is a key ingredient for social inclusion of migrants and refugees. A comprehensive career guidance system is essential to assist them not only in accessing the labour market, but more importantly in acquiring self-knowledge and constructing their careers in their new societies. Understanding their skills and their career interests is essential in order for them to design their future careers and make long-term career plans, considering the initial work placement as the first step in their new career. Without effective career guidance services, public policy measures to integrate migrants at individual or societal level are likely to fail. The NADINE tests are user-friendly tools for employability and career guidance that provide an accurate assessment of migrants' and refugees' skills. This allows guidance professionals to offer a more personalised and contextualized service to migrants based on real needs and the opportunities available to them. The NADINE hard skill tests can have multiple functions: 1) they are a useful tool for career practitioners to assess their clients' skills; 2) they can facilitate the selection of skillful employees by employers and organizations and 3) contribute in designing trainings for upskilling that correspond to the needs of migrants/refugees/asylum seekers. It can be used as a complimentary tool to other EU tools, such as the Labour-Int (http://www.labour-int.eu/) that has developed tools to assess the qualifications and the skills that asylum-seekers' and refugees' have acquired via their job experience.

Acknowledgements

This article is part of the NADINE project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 822601.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Alfonso, V. C., Flanagan, D. P., & Radwan, S. (2005). The Impact of the Cattell-Horn-Carroll Theory on Test Development and Interpretation of Cognitive and Academic Abilities. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (2nd ed., pp. 185-202). New York: Guilford Press.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *The Standards for Educational and Psychological Testing*. Washington DC: AERA Publications.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511571312

Carroll, J. B. (1997). The Three-Stratum Theory of Cognitive Abilities. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories*, *Tests*, *and Issues* (pp. 122-130). New York: Guilford Press.

Carroll, J. B. (2005). The Three-Stratum Theory of Cognitive Abilities. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories*, *Tests*, *and Issues* (2nd ed., pp. 69-76). New York: Guilford Press.

Cattell, R. B. (1957). *Personality and Motivation Structure and Measurement*. New York: World Book.

Cattell, R. B. (1971). *Abilities: Their Structure*, *Growth*, *and Action*. Boston, MA: Houghton Mifflin.

CEDEFOP's Skills Panorama Glossary (2020). https://skillspanorama.cedefop.europa.eu/en/glossary

Clayton, J. (2015). *UNHCR Chief Issues Key Guidelines for Dealing with Europe's Refugee Crisis*. https://www.unhcr.org/55e9793b6.html

Cosgrave, J., Hargrave, K., Foresti, M., & Massa, M. (2016). *Europe's Refugees and Migrants: Hidden Flows, Tightened Borders and Spiralling Costs*. Overseas Development Institute Report. https://euagenda.eu/upload/publications/untitled-67413-ea.pdf

DeMaio, T., & Landreth, A. (2004). Do Different Cognitive Interview Techniques Produce Different Results? In S. Presser et al. (Eds.), *Methods for Testing and Evaluating Survey Questionnaires* (pp. 891-08). Hoboken, NJ: John Wiley and Sons.

Desiderio, M. V. (2016). *Integrating Refugees into Host Country Labor Markets: Challenges and Policy Options*. Washington DC: Migration Policy Institute.

DeVellis, R. F. (2017). *Scale Development: Theory and Applications* (4th ed.). Thousand Oaks, CA: Sage.

EC (European Commission) (2016). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the*

Committee of Regions: Action Plan on the Integration of Third Country Nationals. https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-we-do/policies/european-agenda-migration/proposal-implementation-package/docs/20160607/communication_action_plan_integration_third-country_nationals_en.pdf

EC (European Commission) (2020). *Integrating Migrants and Refugees into the Labour Market: Commission and Social and Economic Partners Relaunch Cooperation*. https://ec.europa.eu/cyprus/news/20200907_1_en

Flanagan, D. P., McGrew, K. S., & Ortiz, S. O. (2000). *The Wechsler Intelligence Scales and CHC Theory: A Contemporary Approach to Interpretation*. Boston, MA: Allyn & Bacon.

Gopinath, G. (2020). *The Great Lockdown: Worst Economic Downturn since the Great Depression*. IMFBlog. https://blogs.imf.org/2020/04/14/the-great-lockdown-worst-economic-downturn-since-the-great-depression/

Hambleton, R. K. (2005). Issues, Designs, and Technical Guidelines for Adapting Tests into Multiple Languages and Cultures. In K. Hambleton, P. Merenda, & C. Spielberger (Eds.), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment* (pp. 3-38). Mahwah, NJ: Lawrence Erlbaum. https://doi.org/10.4324/9781410611758

Harkness, J., Pennell, B. E., & Schoua-Glusberg, A. (2004). Survey Questionnaire Translation and Assessment. In *Methods for Testing and Evaluating Survey Questionnaires* (Vol. 546, pp. 453-473). New York: Wiley. https://doi.org/10.1002/0471654728.ch22

Horn, J. L. (1991). Measurement of Intellectual Capabilities: A Review of Theory. In K. S. McGrew, J. K. Werder, & R. W. Woodcock (Eds.), *Woodcock-Johnson Technical Manual* (pp. 197-232). Chicago, IL: Riverside.

Horn, J. L., & Blankson, N. (2005). Foundations for Better Understanding of Cognitive Abilities. In D. P. Flanagan, & P. L. Harrison (Eds.*), Contemporary Intellectual Assessment: Theories, Tests, and Issues (*2nd ed., pp. 41-68). New York: Guilford.

ILO (International Labour Organization) (2008). Conclusions on Skills for Improved Productivity, Employment Growth and Development. In *International Labour Conference, 2008*. Geneva: ILO. https://www.ilo.org/wcmsp5/groups/public/---ed_emp/---ifp_skills/documents/publication/wcms_103457.pdf

ILO (International Labour Organization) (2015). *Labour Migration and Skills Matching. Labour Migration Highlights, No. 6*. Geneva: Labour Migration Branch and Skills and Employability Branch International Labour Office.

ILO (International Labour Organization) (2016). *Guiding Principles on the Access of Refugees and Other Forcibly Displaced Persons to the Labour Market*. https://www.ilo.org/global/topics/labour-migration/publications/WCMS_536440/lang--en/index.htm

ILO (International Labour Organization) (2017). *How to Facilitate the Recognition of Skills of Migrant Workers—Guide for Employment Service Providers*. Geneva: ILO.

ILO International Labour Organization (2018). *Skills for Migration and Employment*. Policy Brief, Geneva: International Labour Organization 2018, Skills and Employability Branch in Collaboration with Labour Migration Branch.

Kankaraš, M., & Moors, G. (2010). Researching Measurement Equivalence in Cross-Cultural Studies. *Psihologija, 43,* 121-135. https://doi.org/10.2298/PSI1002121K

Kushnir, I., Kilkey, M., & Strumia, F. (2020). EU Integration in the (Post)-Migrant-Crisis Context: Learning New Integration Modes? *European Review, 28,* 306-324.

https://doi.org/10.1017/S1062798719000425

McGorry, S. (2000). Measurement in a Cross-Cultural Environment: Survey Translation Issues. *Qualitative Market Research: An International Journal, 3,* 74-81. https://doi.org/10.1108/13522750010322070

McGrew, K. S. (1997). Analysis of the Major Intelligence Batteries According to a Proposed Comprehensive Gf-Gc Framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 151-180). New York: Guilford Press.

McGrew, K. S. (2005). The Cattell-Horn-Carroll Theory of Cognitive Abilities: Past, Present, and Future. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (2nd ed., pp. 136-182). New York: Guilford Press.

McGrew, K. S., & Flanagan, D. P. (1998). *The Intelligence Test Desk Reference (ITDR): Gf-Gc Cross-Battery Assessment*. Boston, MA: Allyn & Bacon.

Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.

OECD (2005). *Definition and Selection of Key Competencies—Executive Summary*. https://www.oecd.org/pisa/35070367.pdf

OECD (2014). *OECD—Learn, Perform, Succeed, Competency Framework*. https://www.oecd.org/careers/competency_framework_en.pdf

OECD (2018a). *Skills for Jobs*. https://www.oecdskillsforjobsdatabase.org

OECD (2018b). *Skills on the Move: Migrants in the Survey of Adult Skills*. OECD Skills Studies, Paris: OECD Publishing.

Oremus, M., Cosby, J. L., & Wolfson, C. (2005). A Hybrid Qualitative Method for Pretesting Questionnaires: The Example of a Questionnaire to Caregivers of Alzheimer Disease Patients. *Research in Nursing & Health, 28,* 419-430. https://doi.org/10.1002/nur.20095

Polit, D. F., & Beck, C. T. (2006). The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recommendations. *Research in Nursing & Health, 29,* 489-497. https://doi.org/10.1002/nur.20147

Presser, *S.*, & Blair, J. *(*1994*).* Survey Pretesting: Do Different Methods Produce Different Results? *Sociological Methodology, 24,* 73-104. https://doi.org/10.2307/270979

Rankin, J. (2019). *EU Declares Migration Crisis over as It Hits Out at "Fake News"*. https://www.theguardian.com/world/2019/mar/06/eu-declares-migration-crisis-over-hits-out-fake-news-european-commission

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll Model of Intelligence. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (3rd ed., pp. 99-144). New York: Guilford Press.

Sireci, S. G., Yang, Y., Harter, J., & Ehrlich, E. J. (2006). Evaluating Guidelines for Test Adaptations: A Methodological Analysis of Translation Quality. *Journal of Cross-Cultural Psychology, 37,* 557-567. https://doi.org/10.1177/0022022106290478

Society for Industrial and Organizational Psychology (2003). *Principles for the Validation and Use of Personnel Selection Procedures*. Bowling Green, OH: Author, Google Scholar.

Statista (2021). *Unemployment Rate in Selected European Countries as of April 2021*. https://www.statista.com/statistics/1115276/unemployment-in-europe-by-country

Tanzer, N. K. (2005). Developing Tests for Use in Multiple Languages and Cultures: A Plea for Simultaneous Development. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*

(pp. 235-264). Mahwah, NJ: Erlbaum.

The Adecco Group (2017a). *"The Labour Market Integration of Refugees" White Paper: A Focus on Europe*. Heidelberg: The Reallabor Asyl, Heidelberg University of Education. https://www.adeccogroup.com/power-of-work/labour-market-integration-of-refugeesfocus-europe

The Adecco Group (2017b). *The Soft Skills Imperative*. White Paper 1/2017.

Tountopoulou, M., Karaminas, N., Drosos, N., & Vlachaki, F. (2020). The NADINE Soft Skills Tests for Migrants, Refugees and Asylum Seekers: Cultural Adaptation and Preliminary Testing of Their Psychometric Properties. *Psychology, 11,* 1910-1924. https://doi.org/10.4236/psych.2020.1112121

Udayar, S., Fedrigo, L., Durante, F., Clot-Siegrist, E., & Masdonati, J. (2021). Labour Market Integration of Young Refugees and Asylum Seekers: A Look at Perceived Barriers and Resources. *British Journal of Guidance & Counselling, 49,* 287-303. https://doi.org/10.1080/03069885.2020.1858023

UN (United Nations) (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*. https://sdgs.un.org/2030agenda

UN (United Nations) (2018). *Report of the United Nations High Commissioner for Refugees Part II: Global Compact on Refugees*. New York: United Nations. https://www.unhcr.org/gcr/GCR_English.pdf

UN (United Nations) (2020). *Shared Responsibility, Global Solidarity: Responding to the Socio-Economic Impacts of COVID-19*. https://www.un.org/sites/un2.un.org/files/sg_report_socio-economic_impact_of_covid19.pdf

Van de Vijver, F. J. R., & Leung, K. (1997). *Cross-Cultural Psychology Series, Vol. 1. Methods and Data Analysis for Cross-Cultural Research*. London: Sage Publications, Inc.

Willis, G., Schechter, S., & Whitaker, K. (2000). A Comparison of Cognitive Reviewing, Expert Review and Behavior Coding: What Do They Tell Us? In *Proceedings of the Section on Survey Methods, American Statistical Association* (pp. 28-37). Alexandria, VA: American Statistical Association.