

Review

# Psychometric Properties of Heavy Work Investment Measures: A Systematic Review

Julio César Acosta-Prado <sup>1,2,\*</sup> , Arnold Alejandro Tafur-Mendoza <sup>3</sup> , Rodrigo Arturo Zárate-Torres <sup>4</sup>   
and Duván Emilio Ramírez-Ospina <sup>2</sup>

<sup>1</sup> School of Business Science, Universidad del Pacífico, Lima 15072, Peru

<sup>2</sup> School of Accounting, Economic, and Business Sciences, Universidad de Manizales, Manizales 170001, Colombia; merca2@umanizales.edu.co

<sup>3</sup> Research Center (CIUP), Universidad del Pacífico, Lima 15072, Peru; aa.tafurm@up.edu.pe

<sup>4</sup> Colegio de Estudios Superiores de Administración, Bogotá 111071, Colombia; rodrigo.zarate@cesa.edu.co

\* Correspondence: jc.acostap@up.edu.pe; Tel.: +511-219-0100 (ext. 2660)

**Abstract:** In recent years, the study of heavy work investment (HWI) has been diversifying greatly in the various fields of application in the organizational field, for example, occupational health, human resources, quality at work among others. However, to date, no systematic review has been carried out to examine the methodological quality of the instruments designed to measure HWI. Therefore, the present systematic review examines the psychometric properties of three main measures of HWI: Workaholism Battery (WorkBAT), Work Addiction Risk Test (WART), and Dutch Work Addiction Scale (DUWAS). Five electronic databases were systematically searched, selecting psychometric articles. Of the 2621 articles identified, 35 articles met all inclusion criteria published between 1992 and 2019. The findings indicated that most of the articles were focused on reviewing psychometric properties, analyses were conducted from classical test theory, collected validity evidence based on internal structure and relationship with other variables, and reliability of scores was obtained through the internal consistency method. Of the instruments reviewed, the DUWAS is the one with the highest methodological quality. Recommendations are made for future research to address the psychometric study of these instruments based on recent advances in the field of organizational measurement.

**Keywords:** heavy work investment; workaholism; work addiction; psychometric properties; systematic review



**Citation:** Acosta-Prado, J.C.; Tafur-Mendoza, A.A.; Zárate-Torres, R.A.; Ramírez-Ospina, D.E. Psychometric Properties of Heavy Work Investment Measures: A Systematic Review. *Sustainability* **2021**, *13*, 12539. <https://doi.org/10.3390/su132212539>

Academic Editors: Gabriele Giorgi, Antonio Ariza-Montes, Horacio Molina-Sánchez and Dante Castillo Guajardo

Received: 15 October 2021

Accepted: 8 November 2021

Published: 12 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Heavy work investment (HWI) is a concept developed by Snir and Harpaz [1] who define it as the time and effort invested in work as the two central axes of HWI. Based on Weiner's attributional framework they propose three main categories of possible predictors of HWI: background (e.g., gender, parenthood, education), internal (e.g., work addiction, workaholism), and external variables (e.g., financial needs, employer demands) [2]. Thus, the present review focuses on the possible internal predictors, i.e., work addiction and workaholism. Current theoretical proposals of HWI use the job demands–resources (JD–R) model, considering this construct as a continuum consisting of three main parts: antecedents, dimensions, and outcomes [3].

HWI is a construct that underlies many others such as work addiction, work engagement, passion for work, workaholism, among others [4]. That is, HWI is seen as a higher-level umbrella construct and that work addiction, work engagement, passion for work, and workaholism would be at a lower-level, subsumed by HWI [3]. The time invested in work consists of the use of a large number of hours, and work intensity is the effort involved in the various tasks that a person performs at the physical and intellectual level. Regarding internal predictors, work addiction presents compulsive behaviors due to the pressure received by workers [5]. Therefore, within work addiction, heavy investment

of time is marked. It is important to note that HWI is not in itself a negative or positive construct, but that this will depend on the context in which it develops. Thus, a positive implication of this variable is work commitment, where the worker puts dedication and strength into his work activities [6].

Several studies have explored the relationship between workaholism and work engagement, showing significant associations [7], although with varying results, both positively and negatively [2]. In this sense, HWI has been related to greater job security, better career development opportunities, and higher salary. In parallel, HWI may also be negatively related to individual affective and behavioral outcomes that are to be avoided [8,9]. For example, workaholics tend to be associated with poorer mental and physical health, emotional and cognitive exhaustion, poor sleep habits, cardiovascular problems, poor social relationships, and work–life conflicts [10–12]. Thus, it is important to have tools for a correct assessment of the internal factors of HWI (workaholism or work addiction). Workaholism or workaholism includes one of the components of HWI, the heavy time investment in work, however, it does not necessarily include the other component of HWI, the intensity of work during that period [1]. Thus, workaholism or workaholism is related to one aspect of HWI.

In this regard, this review will focus on assessing the quality of three of the most widely used HWI measures reported in the literature: Workaholism Battery (WorkBAT), Work Addiction Risk Test (WART), and Dutch Work Addiction Scale (DUWAS) [13]. In addition, the definitions of standards for educational and psychological tests will be used to ensure consistency with current psychometric guidelines for instrumental studies. To date, there has not been a published study that systematically reviews research on the psychometric properties of HWI measures. This is a significant gap given the increased interest in research beyond the organizational setting. The validity of research findings is contingent on the use of tools with appropriate validity and reliability evidence that measure the constructs of interest. Therefore, the evaluation of psychometrically robust instruments through a systematic review is warranted.

The concept of validity has undergone a long evolutionary process up to the present day. In the 1985 Standards for Educational and Psychological Testing and Manuals, it is maintained that validity is a unitary concept and that the validity of a test is construct validity [14]. In the 2014 Standards, it is insisted that the term test validity is inapplicable, and therefore that of types of validity; rather, it is necessary to provide information from various sources about the purpose for which the test or measurement instrument was developed. It is reiterated that the test is not validated, but rather the inferences made from the scores of the subjects for a given purpose [15].

From this perspective, the test creator is not only responsible for the validity of the test but also the test user. Furthermore, the validity of a test is not established once and for all but is a continuous process of gathering evidence. From a scientific point of view, the only admissible validity is construct validity [16]. Therefore, the logic that underlies it and the methods used to determine it are those that correspond in general to the scientific method. Its evidence comes from various sources, which presuppose a clear definition of the construct and its dimensions or facets if these are necessary.

It is necessary to keep in mind that no study tests or validates a complete theory only concerning some of the inferences that can be derived from it. For constructs, negative results can be interpreted in three directions: the test may not measure the construct; the theoretical framework may be flawed, making it possible for incorrect inferences to be made; or the study design would not allow for proper testing of the hypotheses. These interpretations communicate poor psychometric and research theoretical training and practice, which can lead to ambiguous interpretation of negative results. Finally, it should always be kept in mind that unexpected relationships, as well as predicted ones, are part of the nomological network of the construct and provide arguments for the meaning of the scores.

Conversely, reliability is the degree to which test scores for a group of subjects are consistent in repeated applications of a measurement procedure and hence the reliability and consistency of a subject's score is inferred [17]. It is implicit in this definition that test scores are obtained on different occasions under the same conditions as administration and scoring; it also follows that reliability refers to the precision of the measurement, regardless of what the test measures, and that the reliability of the scores is relative since it is subject to the characteristics of the group of subjects in which it is estimated [18].

## 2. Objectives

The objectives of the present review are to: (1) to systematically identify three of the main measures of HWI in the literature, (2) to evaluate the psychometric properties presented in the reviewed studies, and (3) to determine whether there is a gold standard measure HWI.

## 3. Materials and Methods

### 3.1. Study Design

The present review followed the official PRISMA guidelines [19] on data identification, collection, and analysis. In this sense, the article is a systematic review of instrumental studies on three of the most commonly used measures of HWI: Workaholism Battery (Work-BAT), Work Addiction Risk Test (WART), and Dutch Work Addiction Scale (DUWAS) [13].

### 3.2. Literature Search

First, five selection criteria were determined: (1) instrumental or psychometric studies [20], where the main objective of the article was the creation, adaptation, or psychometric analysis of an instrument measuring HWI or any of its components (Workaholism or Work addiction), specifically, the WorkBAT, WART, and DUWAS; (2) conducted in any country; (3) published up to 2020; (4) peer-reviewed articles; and (5) published in the Spanish or English language.

Three groups of keywords and phrases were used to search for potential articles. The first group is related to psychometric analysis (psychometric, validation, validity, reliability, adaptation, and dimensionality). The second group is related to HWI (heavy work investment, workaholism, work addiction, passion to work, job demands, work craving, work engagement, addiction to work, passion towards work passion for work, and heavy-work investment). The third group is related to self-report measures (questionnaire, measure, assessment, tool, instrument, scale, inventory, and battery). The final search was performed in all databases on 17 July 2021. All articles published up to 2020 were selected.

The search strategy generated a total of 2621 studies with 770 studies from Scopus, 1034 studies from Web of Science (Web of Science databases, including Science Citation Index Expanded [SCI-EXPANDED], Social Sciences Citation Index [SSCI], Arts & Humanities Citation Index [A&HCI], and Emerging Sources Citation Index [ESCI]), 654 studies from PsycNET (PsycInfo and PsycArticles), 110 studies from MEDLINE via Ovid, and 53 studies from Psychology and Behavioral Sciences Collection via EBSCO. Table 1 shows the search strategy followed in the five electronic databases.

**Table 1.** Search strategy in the five electronic databases.

Database	Search Strategy
Scopus	TITLE-ABS-KEY (psychometric* OR validation OR validity OR reliability OR adaptation OR dimensionality) AND TITLE-ABS-KEY ("heavy work investment" OR workaholism OR "work addiction" OR "passion to work" OR "job demands" OR "work craving" OR "work engagement" OR "addiction to work" OR "passion towards work" OR "passion for work" OR "heavy-work investment") AND TITLE-ABS-KEY (questionnaire OR measure * OR assessment OR tool OR instrument OR scale OR inventory OR battery) AND PUBYEAR < 2021 AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (LANGUAGE, "English") OR LIMIT-TO (LANGUAGE, "Spanish"))

Table 1. Cont.

Database	Search Strategy
Web of Science	TS = (psychometric* OR validation OR validity OR reliability OR adaptation OR dimensionality) AND TS = ("heavy work investment" OR workaholism OR "work addiction" OR "passion to work" OR "job demands" OR "work craving" OR "work engagement" OR "addiction to work" OR "passion towards work" OR "passion for work" OR "heavy-work investment") AND TS = (questionnaire OR measure* OR assessment OR tool OR instrument OR scale OR inventory OR battery) Refined By: NOT Publication Years: 2021; Document Types: Articles; Languages: English or Spanish
PsycNET	(Any Field: psychometric* OR Any Field: validation OR Any Field: validity OR Any Field: reliability OR Any Field: adaptation OR Any Field: dimensionality) AND (Any Field: "heavy work investment" OR Any Field: workaholism OR Any Field: "work addiction" OR Any Field: "passion to work" OR Any Field: "job demands" OR Any Field: "work craving" OR Any Field: "work engagement" OR Any Field: "addiction to work" OR Any Field: "passion towards work" OR Any Field: "passion for work" OR Any Field: "heavy-work investment") AND (Any Field: questionnaire OR Any Field: measure* OR Any Field: assessment OR Any Field: tool OR Any Field: instrument OR Any Field: scale OR Any Field: inventory OR Any Field: battery) AND Document Type: Journal Article AND Year: 0 To 2020
Psychology and Behavioral Sciences Collection via EBSCO	(psychometric* OR validation OR validity OR reliability OR adaptation OR dimensionality) AND ("heavy work investment" OR workaholism OR "work addiction" OR "passion to work" OR "job demands" OR "work craving" OR "work engagement" OR "addiction to work" OR "passion towards work" OR "passion for work" OR "heavy-work investment") AND (questionnaire OR measure* OR assessment OR tool OR instrument OR scale OR inventory OR battery) limit year 2020
MEDLINE via Ovid	(psychometric* OR validation OR validity OR reliability OR adaptation OR dimensionality) AND ("heavy work investment" OR workaholism OR "work addiction" OR "passion to work" OR "job demands" OR "work craving" OR "work engagement" OR "addiction to work" OR "passion towards work" OR "passion for work" OR "heavy-work investment") AND (questionnaire OR measure* OR assessment OR tool OR instrument OR scale OR inventory OR battery) limit 1 to year = "1860–2020"

### 3.3. Selection Process

The search results were initially screened by title and abstract to exclude research that did not meet the selection criteria. Subsequently, the remaining studies were retrieved from the different databases and the full articles were evaluated according to their relevance to meet the stipulated selection criteria.

After excluding 795 duplicate searches and 1773 studies that did not meet the selection criteria (because they were qualitative studies, empirical studies, were in a language other than Spanish or English, or were psychometric but worked with other instruments), 53 articles were selected from the review of the title and abstract, and subsequently acquired the 50 articles that were reviewed in their entirety, obtaining a final sample of 35 articles containing 37 studies (one of the articles analyzed the three instruments of interest independently). The flow diagram of the selection of the preceding studies is shown in Figure 1.

### 3.4. Coding Process

The coding was carried out considering the following information from the items analyzed: (a) instrument name, authors, and year; (b) study design; (c) number of items in the instrument; (d) dimensions it measures; (e) number of response options; (f) sample size; (g) participant characteristics; (h) mean and standard deviation of participant age; (i) country of origin of the sample; (j) psychometric theory used for the analyses; (k) validity evidence collected; and (l) reliability evidence collected.

### 3.5. Literature Analysis

After selection and coding of the articles, 35 articles met all the requirements. However, one of the articles had three studies from each of the objective instruments (WorkBAT, WART and DUWAS). Therefore, 37 studies were compiled in the database for further analysis.

sis. The characteristics of each study were organized in Table S1 (Supplementary Materials). Based on the structured information, the results were elaborated to methodologically describe the three main measures of HWI.

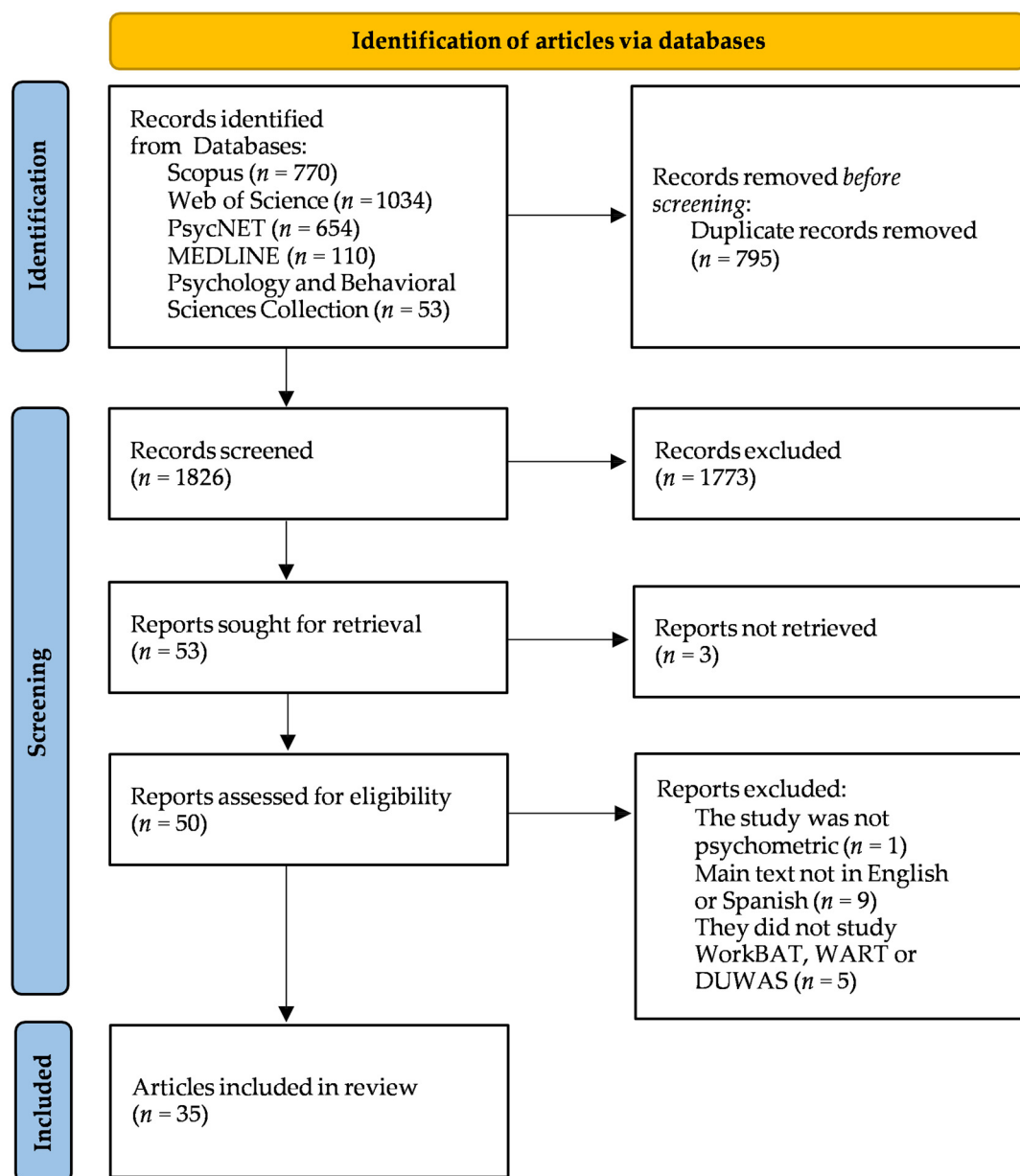


Figure 1. Flow chart of selection of preceding articles.

## 4. Results

### 4.1. Characteristics of the Studies

Of the 37 studies analyzed (Table 2), 13 belong to WorkBAT, 12 correspond to WART and 12 worked with DUWAS. The studies were published between 1992 and 2019. Regarding the study design of the articles reviewed, most of them ( $n = 22$ ) focused on reviewing the psychometric properties (validity and reliability) of the WorkBAT, WART, and DUWAS. Likewise, a significant number of adaptations were reported ( $n = 12$ ), where they sought to adjust the original tests to different cultural contexts. Finally, three studies were found where the tests that are the main focus of this review were developed or constructed [21–23].

**Table 2.** Characteristics of the studies for each instrument.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
1	Workaholism Battery (WorkBAT) (Spence and Robbins, 1992) [21]	Development	Workaholism (Work involvement, Work enjoyment, and Drive).	CTT	<p>Relations to other variables</p> <p>All three factors correlated with Time commitment and Work involvement. In addition, Work involvement and Drive were correlated with Job stress, Perfectionism, and Non-delegation. Finally, Drive and Work enjoyment correlated with health complaints.</p>	<p>Internal consistence</p> <p>Factors: Work involvement (<math>\alpha = 0.69</math> and <math>0.67</math>), Drive (<math>\alpha = 0.81</math> and <math>0.67</math>), and Work enjoyment (<math>\alpha = 0.86</math>).</p>
2	Workaholism Battery (WorkBAT) (Kanai, Wakabayashi, and Fling, 1996) [24]	Adaptation	Workaholism (Work enjoyment and Drive).	CTT	<p>Internal structure</p> <p>Orthogonal two-factor model (EFA), factor loadings greater than 0.30 and variance explained 32%.</p> <p>Relations to other variables</p> <p>Both factors correlated with Time commitment, Job involvement and Perfectionism. Drive was also correlated with Job stress, Non-delegation and Health complaints.</p>	<p>Internal consistence</p> <p>Factors: Drive (<math>\alpha = 0.70</math>) and Work enjoyment (<math>\alpha = 0.85</math>).</p>
3	Workaholism Battery (WorkBAT) (Burke, 1999) [25]	Psychometric properties	Workaholism (Work involvement, Work enjoyment, and Drive).	CTT	<p>Relations to other variables</p> <p>The three factors correlated with Job involvement, Time on the job and Hours worked. Drive correlated with Perfectionism (<math>r = 0.42</math>), Overtime worked (<math>r = 0.26</math>) and Non-delegation (<math>r = 0.20</math>). Work enjoyment correlated with Overtime worked (<math>r = -0.26</math>).</p>	<p>Internal consistence</p> <p>Factors: Work involvement (<math>\alpha = 0.67</math>), Drive (<math>\alpha = 0.80</math>), and Work enjoyment (<math>\alpha = 0.88</math>).</p>
4	Workaholism Battery (WorkBAT) (Burke, 2001) [26]	Psychometric properties	Workaholism (Work involvement, Work enjoyment, and Drive).	CTT	Not reported.	<p>Internal consistence</p> <p>Factors: Work involvement (<math>\alpha = 0.66</math> and <math>0.67</math>), Drive (<math>\alpha = 0.73</math> and <math>0.71</math>), and Work enjoyment (<math>\alpha = 0.88</math>).</p> <p>Test-retest</p> <p>Three months. Work involvement (<math>r = 0.52</math>), Drive (<math>r = 0.59</math>), and Work enjoyment (<math>r = 0.76</math>).</p>
5	Workaholism Battery (WorkBAT) (Burke and Koksai, 2002) [27]	Adaptation	Workaholism (Work involvement, Work enjoyment, and Drive).	CTT	<p>Relations to other variables</p> <p>The three factors were correlated with work behaviors, job satisfaction, emotional well-being, beliefs and fears, and balance values.</p>	<p>Internal consistence</p> <p>Factors: Work involvement (<math>\alpha = 0.56</math>), Drive (<math>\alpha = 0.46</math>), and Work enjoyment (<math>\alpha = 0.79</math>).</p>



Table 2. Cont.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
6	Workaholism Battery (WorkBAT) (Burke, Richardson, and Martinussen, 2002) [28]	Psychometric properties	Workaholism (Work involvement, Work enjoyment, and Drive).	CTT	Internal structure Three-factor orthogonal model (PCA) and variance explained 39%.	Internal consistence Factors: Work involvement ( $\alpha = 0.45$ and $0.64$ ), Drive ( $\alpha = 0.81$ and $0.87$ ), and Work enjoyment ( $\alpha = 0.85$ and $0.84$ ). Test-retest Six months. Work involvement ( $r = 0.49$ ), Drive ( $r = 0.45$ ), and Work enjoyment ( $r = 0.56$ ).
7	Workaholism Battery Revised (WorkBAT-R), revised version of the Workaholism Battery (WorkBAT) (McMillan et al., 2002) [29]	Psychometric properties	Workaholism (Work enjoyment and Drive).	CTT	Internal structure Orthogonal two-factor model (EFA), factor loadings greater than 0.40, variance explained 41.43%, correlation between factors 0.22. Relations to other variables Enjoyment and Drive were correlated with job satisfaction (except for Drive), work involvement, an alternative measure of work addiction (SNAP-Work), and intrinsic work motivation. Likewise, the number of hours worked correlated 0.16 with Enjoyment and 0.22 with Drive.	Internal consistence Factors: Enjoyment ( $\alpha = 0.85$ ) and Drive ( $\alpha = 0.75$ ).
8	Workaholism Battery (WorkBAT) (Ersoy-Kart, 2005) [30]	Adaptation	Workaholism (Work involvement and Drive).	CTT	Internal structure Two-factor oblique model (EFA), factor loadings greater than 0.40, variance explained 29.60%, correlation between factors was 0.47. Relations to other variables Work involvement ( $r = 0.22$ ) and drive ( $r = 0.24$ ) were correlated with type A behavior.	Internal consistence Workaholism ( $\alpha = 0.83$ , split-half reliability coefficient = 0.69). Factors: Work involvement ( $\alpha = 0.81$ ) and Drive ( $\alpha = 0.81$ ).
9	Workaholism Battery (WorkBAT) (Andreassen, Ursin, and Eriksen, 2007) [31]	Psychometric properties	Workaholism (Work involvement, Work enjoyment, and Drive).	CTT	Internal structure The two-factor (no work involvement) and three-factor (CFA) models had an acceptable fit. Relations to other variables Drive and Work enjoyment were correlated with work stress, burnout and subjective health complaints. Work enjoyment was also correlated with work engagement components.	Internal consistence Factors: Work involvement ( $\alpha = 0.49$ ), Drive ( $\alpha = 0.80$ ), and Work enjoyment ( $\alpha = 0.79$ ).

Table 2. Cont.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
10	Workaholism Battery (WorkBAT) (Huang, Hu, and Wu, 2010) [32]	Adaptation	Workaholism (Work enjoyment, Work involvement-enjoyment, Drive-work involvement, Drive, and Work involvement).	CTT	<p>Internal structure</p> <p>Five-factor model (EFA), factor loadings greater than 0.30, variance explained 54.65%, correlation between factors ranged from <math>-0.24</math> to <math>0.31</math>.</p> <p>Relations to other variables</p> <p>The five WorkBAT factors correlated with the Work Addiction Risk Test (WART) and most of its factors, as well as with career commitment, job involvement, emotional exhaustion, and job satisfaction.</p>	<p>Internal consistence</p> <p>Factors: Enjoyment (<math>\alpha = 0.88</math>), Work involvement-enjoyment (<math>\alpha = 0.69</math>), Drive-work involvement (<math>\alpha = 0.58</math>), Drive (<math>\alpha = 0.73</math>), and Work involvement (<math>\alpha = 0.60</math>).</p>
11	Workaholism Battery (WorkBAT) (Andreassen, Hetland, and Pallesen, 2013) [13]	Psychometric properties	Workaholism (Work involvement, Work enjoyment, and Drive).	CTT	<p>Internal structure</p> <p>Four-factor oblique model (PCA), factor loadings greater than 0.40. The three-factor model (CFA) had a poor fit.</p> <p>Relations to other variables</p> <p>The Drive factor had the best correlations with the WART and DUWAS. Work involvement and Work enjoyment showed variable results.</p>	<p>Internal consistence</p> <p>Factors: Involvement (<math>\alpha = 0.63</math>), Drive (<math>\alpha = 0.82</math>), and Work enjoyment (<math>\alpha = 0.84</math>).</p> <p>Test-retest</p> <p>24–30 months. Work involvement (ICC = 0.65), Drive (ICC = 0.64), and Work enjoyment (ICC = 0.61).</p>
12	Workaholism Battery (WorkBAT) (Boada-Grau et al., 2013) [33]	Adaptation	Workaholism (Work enjoyment and Drive).	CTT	<p>Internal structure</p> <p>Two-factor oblique model (EFA), factor loadings greater than 0.30, variance explained 42.64%, correlation between factors 0.19.</p> <p>Two-factor related and good fit model (CFA).</p> <p>Relations to other variables</p> <p>The WorkBAT correlated with irritation, burnout and obsessive beliefs.</p>	<p>Internal consistence</p> <p>Factors: Work enjoyment (<math>\alpha = 0.82</math>, CI 0.80–0.84) and Drive (<math>\alpha = 0.80</math>, CI 0.78–0.83).</p>
13	Workaholism Battery (WorkBAT) (Santos et al., 2018) [34]	Adaptation	Workaholism (Work involvement, Work enjoyment, and Drive).	CTT	<p>Internal structure</p> <p>Two- and three-factor related (CFA) model, poor fit in all goodness-of-fit indices.</p>	<p>Internal consistence</p> <p>Workaholism (<math>\alpha = 0.81</math>). Factors: Work involvement (<math>\alpha = 0.56</math>), Drive (<math>\alpha = 0.82</math>), and Work enjoyment (<math>\alpha = 0.76</math>).</p>



Table 2. Cont.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
14	Work Addiction Risk Test Revised (WART-R), revised version of the Work Addiction Risk Test (WART) (Urbán et al., 2019) [35]	Adaptation	Work addiction (Overcommitment, Impatience, Hard-working, and Salience).	N/A	<p>Internal structure</p> <p>Four-factor oblique model (EFA), factor loadings greater than 0.40 and correlations between factors from 0.16 to 0.50.</p> <p>Related four-factor model (CFA), good fit, factor loadings between 0.39 and 0.73 and correlations between factors from 0.47 to 0.74.</p> <p>Relations to other variables</p> <p>The amount of Time spent working, Mental health symptoms, and Hostility were significantly associated with four factors.</p>	Not reported.
15	Work Addiction Risk Test (WART) (Robinson, Post, and Khakee, 1992) [36]	Psychometric properties	Work addiction	CTT	Not reported.	<p>Internal consistence</p> <p>Work addiction (<math>\alpha = 0.85</math>).</p> <p>Test-retest</p> <p>Two weeks. Work addiction (<math>r = 0.83</math>).</p>
16	Work Addiction Risk Test (WART) (Robinson and Post, 1994) [37]	Psychometric properties	Work addiction	N/A	<p>Test content</p> <p>Subjects were presented with five of the major symptoms of work addiction around which the 25-item measure was constructed: Overdoing, Self-worth, Control-Perfectionism, Intimacy, and Mental Preoccupation-Future Reference. Subjects were asked to select the symptom that matched each of the 25 items. The percentage of correct categorizations ranged from 40% to 96%.</p>	Not reported.
17	Work Addiction Risk Test (WART) (Robinson and Phillips, 1995) [38]	Psychometric properties	Work addiction	N/A	<p>Test content</p> <p>Ten statements unrelated to work addiction were nested throughout the 25-item test. Subjects were asked to identify the 25 items from a list of 35 statements they believed to be symptoms of work addiction. The mean percentage score of correctly identified symptoms was 89.4%. The percentages ranged from 65% to 100%.</p>	Not reported.

Table 2. Cont.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
18	Work Addiction Risk Test (WART) (Robinson and Post, 1995) [39]	Psychometric properties	Work addiction	CTT	Not reported.	Internal consistence Work addiction (Spearman-Brown split-half correlation coefficient = 0.85).
19	Work Addiction Risk Test (WART) (Robinson, 1996) [40]	Psychometric properties	Work addiction	N/A	Relations to other variables Workaholism correlated with generalized anxiety ( $r = 0.40$ ), type A behavior patterns ( $r = 0.37$ ), and with scores on the four scales of the Jenkins Activity Survey with 0.50 on the Type A scale, 0.50 on the Speed and Impatience scale, 0.39 on the Hard-driving and Competitive scale, and 0.20 on the Job Involvement scale.	Not reported.
20	Work Addiction Risk Test (WART) (Robinson, 1999) [22]	Development	Work addiction	CTT	Relations to other variables Work addiction was correlated with anxiety ( $r = 0.40$ ), Type A behaviors ( $r = 0.37$ ), Type A ( $r = 0.50$ ), Speed and impatience ( $r = 0.49$ ), Hard-driving and competitive ( $r = 0.38$ ), and Job involvement ( $r = 0.20$ ).	Internal consistence Work addiction ( $\alpha = 0.88$ ).
21	Work Addiction Risk Test (WART) (Flowers and Robinson, 2002) [41]	Psychometric properties	Work addiction (Compulsive tendencies, Control, Impaired communication/Self-absorption, Inability to delegate, and Self-Worth).	CTT	Internal structure Five-factor oblique model (PCA), factor loadings greater than 0.40 and variance explained 52%. Relations to other variables Four discriminant analyses were conducted to examine the correct classification rate of scores on the WART and explore which of the factors and items accounted for the differences in the average score profiles of the WG and CW. The correct classification rate for the CG remained consistent, ranging from 93.8 to 95.3. The correct classification rate for the WG varied between 57.3 and 70.	Internal consistence Work addiction ( $\alpha = 0.90$ ).

Table 2. Cont.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
22	Work Addiction Risk Test (WART) (Taris, Schaufeli, and Verhoeven, 2005) [42]	Adaptation	Work addiction (Compulsive tendencies, Control, Impaired Communication/Self-Absorption, Inability to delegate, and Self-worth).	CTT	<p>Internal structure</p> <p>Model with five first-order and one second-order factor (CFA), good fit and factor loadings between 0.30 and 0.80 (first order) and 0.85 and 0.96 (second order).</p> <p>Relations to other variables</p> <p>The WART correlated highly with the Compulsive tendencies factor (<math>r = 0.89</math> and <math>0.93</math>), which is proposed as a short version of the WART. Both versions correlated with job stress (job demands and overtime), job strain (work–nonwork conflict, exhaustion, and cynism), and mental health.</p>	<p>Internal consistence</p> <p>Work addiction (<math>\alpha = 0.93</math>). Factors: Compulsive tendencies (<math>\alpha = 0.90</math>), Control (<math>\alpha = 0.82</math>), Impaired Communication/Self-Absorption (<math>\alpha = 0.62</math>), and Self-worth (<math>\alpha = 0.56</math>).</p>
23	Work Addiction Risk Test (WART) (Andreassen, Hetland, and Pallesen, 2013) [13]	Psychometric properties	Work addiction (Compulsive tendencies (CT), Control (CL), Impaired communication (IC), Self-worth (SW), and Inability to delegate (ID)).	CTT	<p>Internal structure</p> <p>Four-factor oblique model (PCA), factor loadings greater than 0.40. The five-factor model (CFA) had a poor fit.</p> <p>Relations to other variables</p> <p>The five WART factors (and the total) correlated with the DUWAS factors and with the WorkBAT Drive factor.</p>	<p>Internal consistence</p> <p>Work addiction (<math>\alpha = 0.86</math>). Factors: CT (<math>\alpha = 0.77</math>), CL (<math>\alpha = 0.75</math>), IC (<math>\alpha = 0.59</math>), and SW (<math>\alpha = 0.36</math>).</p> <p>Test-retest</p> <p>24–30 months. Work addiction (ICC = 0.70), CT (ICC = 0.63), CL (ICC = 0.69), IC (ICC = 0.56), ID (ICC = 0.32), and SW (ICC = 0.56).</p>
24	Work Addiction Risk Test Portuguese of Brazil Version (WART15-PBV), adapted version of the Work Addiction Risk Test (WART) (Romeo et al., 2014) [43]	Adaptation	Work addiction (Compulsive tendencies, Control, and Impaired Communication/Self-Absorption).	CTT	<p>Internal structure</p> <p>Three-factor model (CFA), good fit and factor loadings between 0.36 and 0.95.</p> <p>Relations to other variables</p> <p>WART15-PBV correlated with the DUWAS (<math>r = 0.90</math>) and the correlations between the factors of both tests were greater than 0.50. WART15-PBV also correlated with general health perception (<math>r = 0.29</math>).</p>	<p>Internal consistence</p> <p>Work addiction (<math>\alpha = 0.83</math>). Factors: Compulsive tendencies (<math>\alpha = 0.79</math>), Control (<math>\alpha = 0.54</math>), and Impaired communication/Self-absorption (<math>\alpha = 0.68</math>).</p>

Table 2. Cont.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
25	Work Addiction Risk Test (WART) (Ravoux et al., 2018) [44]	Adaptation	Work addiction (Compulsive tendencies, Control, Impaired communication and self-absorption, Self-worth, and Inability to delegate).	CTT	<p>Internal structure</p> <p>Four-component model (PCA). Relations to other variables</p> <p>Work addiction and the factors Compulsive tendencies, Control, and Impaired communication and self-absorption were correlated with the visual analog scale of stress at work (<math>r = 0.43</math>), stress at home (<math>r = 0.41</math>), and the visual analog scale of well-being (<math>r = -0.40</math>).</p>	<p>Internal consistence</p> <p>Work addiction (<math>\alpha = 0.90</math>). Factors: <math>\alpha</math> between 0.57 and 0.85.</p> <p>Test-retest</p> <p>One week. The Lin concordance coefficient indicated a value of 0.90 for the total test and values between 0.66 and 0.86 for the factors.</p>
26	Dutch Work Addiction Scale (DUWAS) (Schaufeli, Shimazu, and Taris, 2009) [23]	Development	Workaholism (Working excessively (WE) and Working compulsively (WC)).	CTT	<p>Internal structure</p> <p>Two-component model (PCA), factor loadings between 0.57 and 0.82 (TN) and 0.52 and 0.74 (JP). Explained variance of 52.5% (TN) and 46.5% (JP). Related two-factor model (CFA), good fit and relationship between factors high in TN (<math>r = 0.50</math>) and JP (<math>r = 0.59</math>).</p> <p>Factor invariance: MG-CFA tested the equivalence of the CFA (configural) model in the Dutch and Japanese samples.</p> <p>Relations to other variables</p> <p>All correlations between workaholism (WE and WC) and excess working time (overtime percentage and overwork) are positive.</p> <p>Workaholism was not related to Engagement (<math>r = -0.19</math>, NT; <math>r = -0.05</math>, JP) and was related to Burnout (<math>r = 0.53</math>, NT; <math>r = 0.64</math>, JP).</p>	<p>Internal consistence</p> <p>Factors (TN): WE (<math>\alpha = 0.78</math>) and WC (<math>\alpha = 0.78</math>). Factors (JP): WE (<math>\alpha = 0.73</math>) and WC (<math>\alpha = 0.68</math>).</p>

Table 2. Cont.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
27	Dutch Work Addiction Scale (DUWAS) (del Líbano et al., 2010) [45]	Psychometric properties	Workaholism (Working excessively (WE) and Working compulsively (WC)).	CTT	<p>Relations to other variables (DUWAS-10 and DUWAS-17): Intercorrelations between the original and shortened factors ranged between 0.92 and 0.94 in the TN and SP samples. The intercorrelations of WE and WC with perceived health and happiness were negative in both samples.</p> <p>Internal structure (DUWAS-10): Two-factor related (CFA) model and good fit in the TN and SP samples. Factor invariance (DUWAS-10): MG-CFA tested the equivalence of the CFA model (factor structure, covariance, and factor loadings) between the TN and SP samples.</p>	<p>Internal consistence DUWAS-17 Factors. TN: WE (<math>\alpha = 0.82</math>) and WC (<math>\alpha = 0.84</math>). SP: WE (<math>\alpha = 0.85</math>) and WC (<math>\alpha = 0.79</math>). DUWAS-10 Factors. TN: WE (<math>\alpha = 0.75</math>) and WC (<math>\alpha = 0.81</math>). SP: WE (<math>\alpha = 0.78</math>) and WC (<math>\alpha = 0.79</math>).</p>
28	Dutch Work Addiction Scale -Observer Rating (DUWAS-OR) (Falco et al. 2012) [46]	Psychometric properties	Workaholism (Working excessively (WE) and Working compulsively (WC)).	CTT	<p>Internal structure Related two-factor model (CFA), acceptable fit, factor loadings between 0.61 and 0.86, in addition, the correlation between factors was 0.69.</p> <p>Relations to other variables AVE was 0.50 for WE and 0.65 for WC. In addition, the two factors correlated with the UWES (self-report), WE (<math>r = 0.49</math>) and WC (<math>r = 0.43</math>). WE and WC correlated with Workload and WE with Work-family conflict.</p> <p>Discriminant evidence: Evidence was provided through the Fornell and Larcker criterion for the two factors.</p>	<p>Internal consistence Factors: WE (<math>\alpha = 0.86</math>) and WC (<math>\alpha = 0.89</math>).</p>
29	Dutch Work Addiction Scale (DUWAS) (Molino, Ghislieri, and Colombo, 2012) [47]	Psychometric properties	Working excessively	CTT	<p>Internal structure One-factor model (EFA), factor loadings between 0.41 and 0.76 and explained a variance of 38.62%. Unifactorial model (CFA), good fit, factor loadings between 0.38 and 0.84.</p> <p>Relations to other variables Relationship between Working excessively and Work-family conflict (<math>r = 0.49</math>).</p>	<p>Internal consistence Working excessively (<math>\alpha = 0.74</math>).</p>

Table 2. Cont.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
30	Dutch Work Addiction Scale (DUWAS) (Andreassen, Hetland, and Pallesen, 2013) [13]	Psychometric properties	Workaholism (Working excessively (WE) and Working compulsively (WC)).	CTT	<p>Internal structure</p> <p>Two-factor oblique model (PCA), factor loadings greater than 0.40. The two-factor model (CFA) had a poor fit.</p> <p>Relations to other variables</p> <p>The two DUWAS factors correlated with the WART factors and the WorkBAT factors.</p>	<p>Internal consistence</p> <p>Factors: WE (<math>\alpha = 0.69</math>) and WC (<math>\alpha = 0.63</math>).</p> <p>Test-retest</p> <p>24–30 months. WE (ICC = 0.61) and WC (ICC = 0.65).</p>
31	Dutch Work Addiction Scale (DUWAS) (Sharma and Sharma, 2013) [48]	Psychometric properties	Workaholism (Working excessively (WE), Working compulsively (WC), and Overwork(O)).	CTT	<p>Relations to other variables</p> <p>Convergent evidence: AVE greater than 0.50 for the three factors.</p> <p>Discriminant evidence: Evidence was provided through the Fornell and Larcker criterion for the three factors.</p> <p>Internal structure</p> <p>Three-factor hierarchical model (CFA). Poor fit. Factor loadings between 0.50 and 0.77 (first order), and between 0.51 and 0.89 (second order).</p>	<p>Internal consistence</p> <p>Workaholism (<math>\alpha = 0.83</math>). Factors: WE (<math>\alpha = 0.88</math>), WC (<math>\alpha = 0.83</math>), and O (<math>\alpha = 0.60</math>).</p>
32	Dutch Work Addiction Scale (DUWAS) (Littman-Ovadia, Balducci, and Ben-Moshe, 2014) [49]	Adaptation	Workaholism (Working excessively (WE) and Working compulsively (WC)).	CTT	<p>Internal structure</p> <p>Related two-factor model (CFA), acceptable fit, factor loadings between 0.38 and 0.77, in addition, the correlation between factors was 0.76.</p> <p>Relations to other variables</p> <p>Self-reports and peer-reports of workaholism (UWES answered by spouse, boyfriend, girlfriend, friend, or colleague) correlated; Workaholism (<math>r = 0.52</math>), WE (<math>r = 0.50</math>), and WC (<math>r = 0.43</math>).</p> <p>Workaholism and its scales showed positive correlations with overcommitment, the actual number of hours worked per week, burnout (emotional exhaustion), work engagement (Absorption), and intrinsic aspects of the job.</p>	<p>Internal consistence</p> <p>Workaholism (<math>\alpha = 0.78</math>). Factors: WE (<math>\alpha = 0.61</math>) and WC (<math>\alpha = 0.70</math>).</p> <p>Test-retest</p> <p>Two or three months. Workaholism (<math>r = 0.79</math>). Factors: WE (<math>r = 0.77</math>) and WC (<math>r = 0.71</math>).</p>



Table 2. Cont.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
33	Dutch Work Addiction Scale (DUWAS) (Rantanen et al., 2015) [50]	Psychometric properties	Workaholism (Working excessively (WE) and Working compulsively (WC)).	CTT	<p>Internal structure</p> <p>Related two-factor model (CFA), an acceptable fit in both samples. However, a second-order model showed a better fit (WE: working frantically and working long hours; WC: obsessive work drive and unease if not working) with loadings greater than 0.50. Likewise, a four-factor related model (with the first-order factors of the previous model) also indicated a good fit.</p> <p>Factorial invariance: Second-order factor structure showed reasonable measurement invariance and stability of factor structure across the two samples and time in the Finnish sub-sample of managers with two measurement points two years apart.</p>	<p>Internal consistence</p> <p>Workaholism (<math>\alpha = 0.82</math>, TN; <math>\alpha = 0.86</math>, FL). Factors (TN): WE (<math>\alpha = 0.72</math>) and WC (<math>\alpha = 0.80</math>). Factors (FL): WE (<math>\alpha = 0.80</math>) and WC (<math>\alpha = 0.80</math>). Subfactors <math>\alpha</math> between 0.51 and 0.80.</p> <p>Test-retest</p> <p>Two-year. Factors: WE (<math>r = 0.66</math>) and WC (<math>r = 0.76</math>). Subfactors' <math>r</math> between 0.60 and 0.71.</p>
34	Dutch Work Addiction Scale (DUWAS) (Mir, Kamal, and Masood, 2016) [51]	Adaptation	Workaholism (Working excessively (WE) and Working compulsively (WC)).	CTT	<p>Internal structure</p> <p>Two-factor model (CFA), good fit, factor loadings between 0.46 and 0.74, 5 items for WE and 3 items for WC.</p>	<p>Internal consistence</p> <p>Workaholism (<math>\alpha = 0.71</math>). Factors: WE (<math>\alpha = 0.66</math>) and WC (<math>\alpha = 0.64</math>).</p>
35	Dutch Work Addiction Scale (DUWAS) (Azevedo and Mathias, 2017) [52]	Psychometric properties	Addiction to work	CTT	<p>Internal structure</p> <p>One-component model (PCA), factor loadings between 0.60 and 0.70, and explained variance of 44.5%.</p> <p>Relations to other variables</p> <p>Addiction to work was positively correlated with the number of shifts worked (<math>r = 0.20</math>), and it showed a negative correlation with age (<math>r = -0.20</math>).</p>	<p>Internal consistence</p> <p>Addiction to work (<math>\alpha = 0.86</math>).</p>

Table 2. Cont.

No	Instrument Name and Authors	Study Design	Dimension Measured	Psychometric Properties		
				Theory	Validity	Reliability
36	Dutch Work Addiction Scale (DUWAS) (Balducci et al., 2017) [53]	Psychometric properties	Workaholism (Working excessively (WE) and Working compulsively (WC)).	CTT	<p>Internal structure</p> <p>Related two-factor model (CFA), good fit and factor loadings between 0.44 and 0.72.</p> <p>Factor invariance: MG-CFA tested the equivalence of the CFA model (configural, metric, factor variance and factor covariance) between the Italian sample and another Dutch sample (<math>n = 7523</math>).</p> <p>Relations to other variables</p> <p>DUWAS and its factors are positively related to the number of hours worked in a week, job demands (workload and work-to-family conflict), high and low arousal job-related negative affect, and psychological strain symptoms.</p>	<p>Internal consistence</p> <p>Workaholism (<math>\alpha = 0.82</math>). Factors: WE (<math>\alpha = 0.74</math>, CI = 0.72–0.77; <math>\omega = 0.75</math>) and WC (<math>\alpha = 0.74</math>, CI = 0.72–0.77; <math>\omega = 0.75</math>).</p> <p>Test-retest</p> <p>One-year. Workaholism (<math>r = 0.57</math>). Factors: WE (<math>r = 0.62</math>) and WC (<math>r = 0.54</math>).</p>
37	Dutch Work Addiction Scale (DUWAS) (Nonnis et al., 2017) [54]	Psychometric properties	Workaholism (Working excessively (WE) and Working compulsively (WC)).	RMT	<p>Internal structure</p> <p>The parallel analysis of the residuals shows that the work engagement has two significant components. However, the eigenvalues are below the cut-off point (2) for both factors.</p>	<p>Internal consistence</p> <p>Person separation reliability R, which scored 0.49 in the WE and 0.56 in the WC.</p>

Note. USA = United States, PCA = Principal Component Analysis, EFA = Exploratory Factor Analysis, CFA = Confirmatory Factor Analysis, CTT = Classical Test Theory, RMT = Rasch Measurement Theory,  $\alpha$  = Alpha coefficient,  $\omega$  = Omega coefficient,  $r$  = Correlation coefficient, AVE = Average Variance Extracted, Mdn = Median, CI = Confidence interval, CCI = Intraclass Correlation Coefficient, N/A = Not Applicable.

Regarding the number of items, it was observed that the psychometric tests analyzed varied in number due to different factors: (1) they were short versions; (2) items were eliminated due to methodological problems; or (3) they only studied some factors of the instrument. Conversely, the dimensions measured by the tests also changed from one study to another, since the structures were not fully replicated in different samples. Nevertheless, the DUWAS was the one that most of the time showed a two-factor structure (nine studies). In terms of the number of response options, the DUWAS (from “almost never” to “almost always”) and the WART (from “never true” to “always true”) presented almost the same four-point Likert scale across studies. The WorkBAT was usually answered on a five-point Likert scale (from “strongly agree” to “strongly disagree”).

As for the sample size, in most cases it did not exceed 500 participants, which is necessary to obtain more robust and stable results, especially considering the diversity of structures presented by these tests. Furthermore, in some studies, the sample was larger than 1000 people. The participants in the studies presented diverse characteristics, mainly workers from different economic sectors (e.g., manufacturing, retail, service, education, or medical). Moreover, the age of the participants was generally between 30 and 50 years old. Finally, among the countries where these instruments have been most tested are the United States, Canada, Norway, the Netherlands, Italy, and Brazil, as well as others located in Asia and Oceania. No studies were found in Latin America or Africa.

#### 4.2. Measurement Theory

The measurement theory that predominated in the studies was the Classical Test Theory (CTT) and only one study used the Rasch Measurement Theory (RMT) [54]. No study used the Item Response Theory (IRT). In this regard, it is important to note that some studies did not use a specific theory, because their objective was to collect evidence of validity and they did not focus on the analysis of items or the reliability of the scores [37,38,40]. CTT proposes a linear relationship between the observed score, the true score, and the measurement error, with the main limitation being the dependence of the persons and the test [55]. Conversely, the IRT seeks to explain the relationship between a person's ability and the probability of answering an item correctly, considering various characteristics such as difficulty, discrimination, or pseudo-guessing [56]. The IRT is considered a descriptive model since it tries to explain as much of the variance as possible [57]. Finally, the RMT is a prescriptive model, in which it is of interest to know whether the data fit the measurement model and, like the IRT, it focuses on the individual analysis of the items based on the interaction between an item and a person [58].

#### 4.3. Validity

Regarding the collection of validity evidence, the studies analyzed focused mainly on two aspects: the internal structure and the relationship with other variables. However, two studies collected evidence based on test content. This source of validity evidence refers to the degree to which the content (domain definition, domain representation, domain relevance, and appropriateness of construction processes) of a test is congruent with the measurement objectives [59].

##### 4.3.1. Validity Evidence Based on Test Content

Two studies collected this source of evidence based on the presentation of the items and their classification by a group of students or experts in the field, a method different from the one usually used to collect this evidence of validity, which consists of a group of judges giving their assessment of the items according to certain features (representativeness, relevance or clarity). However, the method used in these studies has the characteristic of analyzing the representativeness of the items [37,38]. Both studies belonged to the WART.

#### 4.3.2. Validity Evidence Based on Internal Structure

The three tests studied are multidimensional, so this source of evidence is one of the most widely used to corroborate how the instruments are structured according to their items. The techniques used for this purpose were mainly three: principal components analysis (PCA), exploratory factor analysis (EFA), and confirmatory factor analysis (CFA). CFA was usually employed in studies reviewing psychometric properties as the first option, but, in cases where the fit was poor, a less restrictive method, such as PCA or EFA, was chosen. Of the three psychometric instruments, the DUWAS replicated its two-factor structure the most often. Thus, in nine studies, it was observed that the DUWAS was made up of the factors Working excessively and Working compulsively. Regarding the WorkBAT, a three-factor structure (Work involvement, Work enjoyment, and Drive) was observed in eight studies, although a two-factor structure (Work enjoyment and Drive) was also found in four studies. Finally, the WART showed a unidimensional structure in six studies, while four studies reported a five-factor structure (Compulsive tendencies, Control, Impaired communication, Self-worth, and Inability to delegate).

Within this validity, evidence is also the factorial invariance analysis that was employed by some studies to compare whether the structure of the instrument was invariant concerning the country of origin of the participants [23,53]. Factor invariance was performed through multi-group confirmatory factor analysis (MG-CFA), although in some cases the description of the procedure and results was not clear.

#### 4.3.3. Validity Evidence Based on Relations to Other Variables

Since HWI is a relatively recent construct, it is necessary to form a nomological network to determine with which other variables it is related and with which it diverges. In this sense, this source of validity evidence made it possible to relate the total and factor scores of the WorkBAT, WART, and DUWAS to other variables such as Job stress, Job involvement, Time on the job, Overtime worked, Job satisfaction, and other variables linked to the organizational field. In most cases, simple correlation coefficients or simple linear regressions were used, with few studies analyzing the relationship between variables from a more advanced perspective such as structural equation models.

#### 4.4. Reliability

The reliability of the scores was mostly evaluated by analyzing the internal consistency of the items and using the alpha coefficient. However, in no case were the assumptions that this reliability estimator needs to be used (unidimensional, absence of correlated errors, and tau-equivalence of the measurement model) corroborated. In addition, only one study reported confidence intervals for coefficient alpha. The DUWAS was the instrument that most often presented good levels of reliability (greater than 0.70). However, in some studies, it was found to be below this criterion. The WorkBAT and the WART had greater reliability problems in some of their dimensions, probably due to the number of factors (usually four) and the number of items that composed them.

In contrast, some studies evaluated reliability based on the temporal stability or test-retest method, which is necessary to determine whether test scores are constant over time. In most cases, the relationship between scores at time 1 and time 2 was low, less than 0.70 [28]. However, this could be influenced by the use of the simple correlation coefficient used to assess the agreement between the two measures. Only two studies used appropriate statistics for this purpose, such as the intraclass correlation coefficient [13] and Lin's coefficient of concordance [44].

### 5. Discussion

The present review aimed to (1) systematically identify three of the main measures of HWI in the literature, (2) evaluate the psychometric properties presented in the reviewed studies, and (3) determine whether there is a gold standard measure of HWI. Regarding the first objective, three measurement instruments were analyzed, the Workaholism Battery

(WorkBAT), Work Addiction Risk Test (WART), and Dutch Work Addiction Scale (DUWAS), which are the main measures with which various studies approach the phenomenon of HWI. Regarding the second objective, the empirical evidence of the instruments was systematized concerning the validity and reliability of their scores. Finally, regarding the third objective, although the DUWAS is the test that shows the best performance, because its two-factor factor structure was replicated and presented better levels of reliability in most studies, it is slightly superior to the WorkBAT and the WART, so it cannot be concluded that it is the most optimal measure for the measurement of HWI. Thus, further studies with these instruments, based on more robust methodologies, are needed.

About the measurement theory used, the tests have not been tested through the Item Response Theory (IRT) or different models of the Rasch Measurement Theory (RMT), so the performance of the items and the test in people with different levels of HWI is not known. That is, through these models, it is possible to identify at what level of trait of the person, the instruments are more or less reliable since from the CTT reliability is only estimated globally. Moreover, from these two perspectives, it is possible to construct computerized adaptive tests, where workers do not always have to answer the same items and the same number of items, but where the items are adapted to their responses and their trait levels.

Regarding validity, it is necessary to provide evidence of the content of the items. The two studies that were found, and that collected this evidence are more than 20 years old, so the context in which they were evaluated is different from the current one. Likewise, two additional sources of validity have not been studied to date. The evidence of validity is based on response processes, which would allow us to know to what emotions, feelings, thoughts, memories, and behaviors of those evaluated associate when they read the items of the instruments. In this way, we could approximate an explanation that the items are indeed assessing aspects of HWI. In addition, the last source of evidence that has not been considered is that based on the consequences of the application of the test and that involves aspects that go beyond the assessment, such as the actions that could be taken after learning the results in a particular group.

Regarding the reliability of the measures, the results were very variable, finding in many cases values below the minimum acceptable (0.70). However, it is important to note that in no case were the alpha coefficient assumptions corroborated, so there is a possibility that they are underestimated or distorted. Likewise, as in other statistical techniques, it is necessary to consider the nature of the items being worked on, since the WorkBAT, WART, and DUWAS have a Likert-type scale, whose level of measurement is ordinal. Therefore, reliability estimators should be calculated based on a polychoric matrix, which allows a coherent evaluation of the internal consistency of the scores. Conversely, temporal stability has been used in several studies, although methodologically, the best coefficient was not used to see if the scores changed substantially from one assessment to another. In almost all cases, a correlation coefficient was used, which is responsible for seeing whether the scores would be found to be related. However, in these cases, it is suggested to use a coefficient that assesses the concordance (coincidence) of the scores.

Currently, there are systematic reviews of measures of different variables relevant to the organizational field, for example, stakeholder engagement [60] or human resource management systems [61]. However, as mentioned at the beginning of the review, this is the first work that systematically evaluates the psychometric properties of HWI measures. Nevertheless, the study by Andreassen et al. (2013) [13] empirically studied the psychometric properties of the WorkBAT, WART, and DUWAS, through a cross-sectional study where they analyzed the temporal stability and factorial structure of the aforementioned instruments. The study was conducted with 368 Norwegian workers and found that the correlations between the scales were low, showing low convergent evidence, weak temporal stability of scores, and an internal structure for the WorkBAT of four factors, as well as for the WART, while the DUWAS showed a two-factor factorial solution. The authors concluded that the three measures could be used interchangeably. This coincides with the

findings of the present review since no outstanding performance of any of the measures was observed over the others.

The present review has some limitations. First, this review only synthesized those articles written in English or Spanish, and articles written in other languages were not involved. Further, three instruments measuring one side of HWI were analyzed, which, although they are not the main ones, are not the only ones, since other tests could also be observed in the selection of the final articles, although in a smaller proportion (for example, the Workaholism Analysis Questionnaire—WAQ or Bergen Work Addiction Scale—BWAS). Moreover, only articles published up to 2020 were analyzed. However, in 2021, one psychometric study has been published that presents tools specific to the HWI considering Time Commitment and Work Intensity as dimensions [62]. Thus, the aforementioned limitations should be taken into consideration for future studies, where other instruments are analyzed, that can help to understand this construct that is in full development and boom in terms of research to have a clearer approximation of its functioning.

This systematic review has theoretical and practical implications related to HWI. At a theoretical level, the review of the three instruments allows us to know the different aspects of HWI, specifically workaholism and work addiction, exploring how they are theoretically related to other variables, such as job satisfaction and engagement. In this way, it allows the construction of a nomological network for HWI, where different organizational and mental health constructs are explored. However, the practical implications lie in knowing the strengths and weaknesses of the three measures analyzed to know which versions have the most empirical support and which are still to be studied.

Likewise, the article contributes to the HWI research line, especially in the current context of a health emergency, where teleworking conditions have made the line between personal and professional life increasingly thinner, or in some cases, disappearing. Therefore, it is necessary to have good measurement tools that can capture relevant details of HWI in workers who have high levels of time and effort investment. In this sense, the findings obtained suggest that the DUWAS, the WorkBAT, and the WART can be used to measure HWI. However, the DUWAS presents a slightly better psychometric performance than the other two scales, so it should be considered as a first option although the decision should be linked to other factors such as the objective of the measurement, the components to be explored, and the context of the application.

These findings suggest that HWI has consequences in the organizational field and applied settings (e.g., occupational health or human resources). Moreover, integral to theory development is the ability to differentiate a construct from its antecedents and outcomes. Therefore, developing a thorough understanding of the nature of the HWI phenomenon and its consequences requires the use of sound and appropriate psychometric tools to measure the construct.

## 6. Conclusions

The present systematic review constitutes the first effort to summarize the psychometric properties of three popular HWI measures (WorkBAT, WART, and DUWAS). In this way, the review provides empirical evidence of the performance of the three instruments so that researchers, psychologists, managers, and other interested parties can choose the measure that best suits their needs and objectives of use. According to the findings obtained, the three HWI measurement instruments have similar performance, so it is advisable to conduct additional psychometric studies focused on other sources of validity evidence, analysis of bias due to certain sociodemographic variables, or with more robust psychometric models (for example, from the IRT or RMT). Finally, based on this review, it is possible to propose new lines of research focused on covering under-explored or unexplored psychometric aspects of the WorkBAT, WART, and DUWAS.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/su132212539/s1>, Table S1: Characteristics of the studies for each instrument.



**Author Contributions:** Conceptualization, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; methodology, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; software, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; validation, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; formal analysis, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; investigation, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; resources, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; writing—original draft preparation, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; writing—review and editing, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; visualization, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; supervision, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O.; project administration, J.C.A.-P., A.A.T.-M., R.A.Z.-T. and D.E.R.-O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Snir, R.; Harpaz, I. Beyond workaholism: Towards a general model of heavy work investment. *Hum. Resour. Manag. Rev.* **2012**, *22*, 232–243. [\[CrossRef\]](#)
2. Di Stefano, G.; Gaudiino, M. Workaholism and work engagement: How are they similar? How are they different? A systematic review and meta-analysis. *Eur. J. Work Organ. Psychol.* **2019**, *28*, 329–347. [\[CrossRef\]](#)
3. Tabak, F.; Tziner, A.; Shkoler, O.; Rabenu, E. The complexity of Heavy Work Investment (HWI): A conceptual integration and review of antecedents, dimensions, and outcomes. *Sustainability* **2021**, *13*, 7803. [\[CrossRef\]](#)
4. Tziner, A.; Buzea, C.; Rabenu, E.; Shkoler, O.; Truța, C. Understanding the relationship between antecedents of Heavy Work Investment (HWI) and burnout. *Amfiteatru Econ.* **2019**, *21*, 153–176. [\[CrossRef\]](#)
5. Griffiths, M.D.; Demetrovics, Z.; Atroszko, P.A. Ten myths about work addiction. *J. Behav. Addict.* **2018**, *7*, 845–857. [\[CrossRef\]](#)
6. Bakker, A.B.; Shimazu, A.; Demerouti, E.; Shimada, K.; Kawakami, N. Work engagement versus workaholism: A test of the spillover-crossover model. *J. Manag. Psychol.* **2014**, *29*, 63–80. [\[CrossRef\]](#)
7. Tóth-Király, I.; Morin, A.J.S.; Salmela-Aro, K. A longitudinal perspective on the associations between work engagement and workaholism. *Work Stress* **2021**, *35*, 27–56. [\[CrossRef\]](#)
8. Acosta-Prado, J.C.; Sandoval-Reyes, J.G.; Sanchis-Pedregosa, C. Job demands and recovery experience: The mediation role of heavy work investment. *Amfiteatru Econ.* **2020**, *22*, 1206–1226. [\[CrossRef\]](#)
9. Mazzetti, G.; Guglielmi, D.; Schaufeli, W.B. Same involvement, different reasons: How personality factors and organizations contribute to heavy work investment. *Int. J. Environ. Res. Public Health* **2020**, *17*, 8550. [\[CrossRef\]](#)
10. Van Beek, I.; Taris, T.W.; Schaufeli, W.B.; Brenninkmeijer, V. Heavy work investment: Its motivational make-up and outcomes. *J. Manag. Psychol.* **2014**, *29*, 46–62. [\[CrossRef\]](#)
11. Gaudiino, M.; Di Stefano, G. To detach or not to detach? The role of psychological detachment on the relationship between heavy work investment and well-being: A latent profile analysis. *Curr. Psychol.* **2021**. [\[CrossRef\]](#)
12. Balducci, C.; Alessandri, G.; Zaniboni, S.; Avanzi, L.; Borgogni, L.; Fraccaroli, F. The impact of workaholism on day-level workload and emotional exhaustion, and on longer-term job performance. *Work Stress* **2021**, *35*, 6–26. [\[CrossRef\]](#)
13. Andreassen, C.S.; Hetland, J.; Pallesen, S. Psychometric assessment of workaholism measures. *J. Manag. Psychol.* **2013**, *29*, 7–24. [\[CrossRef\]](#)
14. Izquierdo, I.; Olea, J.; Abad, F.J. Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema* **2014**, *26*, 395–400. [\[CrossRef\]](#) [\[PubMed\]](#)
15. American Educational Research Association; American Psychological Association; National Council on Measurement in Education. *Standards for Educational and Psychological Testing*; American Educational Research Association: Washington, DC, USA, 2014; ISBN 978-0935-302-356.
16. Nunnally, J.; Bernstein, I. *Psychometric Theory*, 3rd ed.; McGraw-Hill: New York, NY, USA, 1994; ISBN 9780070478497.
17. Raykov, T. Estimation of composite reliability for congeneric measures. *Appl. Psychol. Meas.* **1997**, *21*, 173–184. [\[CrossRef\]](#)
18. Flora, D.B. Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Adv. Methods Pract. Psychol. Sci.* **2020**, *3*, 484–501. [\[CrossRef\]](#)
19. Page, M.J.; Moher, D.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ* **2021**, *372*. [\[CrossRef\]](#)
20. Ato, M.; López, J.J.; Benavente, A. A classification system for research designs in psychology. *An. Psicol.* **2013**, *29*, 1038–1059. [\[CrossRef\]](#)

21. Spence, J.T.; Robbins, A.S. Workaholism: Definition, measurement, and preliminary results. *J. Pers. Assess.* **1992**, *58*, 160–178. [\[CrossRef\]](#)
22. Robinson, B.E. The Work Addiction Risk Test: Development of a tentative measure of workaholism. *Percept. Mot. Skills* **1999**, *88*, 199–210. [\[CrossRef\]](#)
23. Schaufeli, W.B.; Shimazu, A.; Taris, T.W. Being driven to work excessively hard: The evaluation of a two-factor measure of workaholism in the Netherlands and Japan. *Cross-Cult. Res.* **2009**, *43*, 320–348. [\[CrossRef\]](#)
24. Kanai, A.; Wakabayashi, M.; Fling, S. Workaholism among employees in Japanese corporations: An examination based on the Japanese version of the Workaholism Scales. *Jpn. Psychol. Res.* **1996**, *38*, 192–203. [\[CrossRef\]](#)
25. Burke, R.J. Workaholism in organizations: Measurement validation and replication. *Int. J. Stress Manag.* **1999**, *6*, 45–55. [\[CrossRef\]](#)
26. Burke, R.J. Spence and Robbins' measures of workaholism components: Test-retest stability. *Psychol. Rep.* **2001**, *88*, 882–888. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Burke, R.J.; Koksall, H. Workaholism among a sample of Turkish managers and professionals: An exploratory study. *Psychol. Rep.* **2002**, *91*, 60–68. [\[CrossRef\]](#)
28. Burke, R.J.; Richardsen, A.M.; Martinussen, M. Psychometric properties of Spence and Robbins' measures of workaholism components. *Psychol. Rep.* **2002**, *91*, 1098–1104. [\[CrossRef\]](#) [\[PubMed\]](#)
29. McMillan, L.H.W.; Brady, E.C.; O'Driscoll, M.P.; Marsh, N.V. A multifaceted validation study of Spence and Robbins' (1992) Workaholism Battery. *J. Occup. Organ. Psychol.* **2002**, *75*, 357–368. [\[CrossRef\]](#)
30. Ersoy-Kart, M. Reliability and validity of the Workaholism Battery (Work-BAT): Turkish Form. *Soc. Behav. Pers.* **2005**, *33*, 609–618. [\[CrossRef\]](#)
31. Andreassen, C.S.; Ursin, H.; Eriksen, H.R. The relationship between strong motivation to work, "workaholism", and health. *Psychol. Health* **2007**, *22*, 615–629. [\[CrossRef\]](#)
32. Huang, J.C.; Hu, C.; Wu, T.C. Psychometric properties of the Chinese version of the workaholism battery. *J. Psychol.* **2010**, *144*, 163–183. [\[CrossRef\]](#)
33. Boada-Grau, J.; Prizmic-Kuzmica, A.-J.; Serrano-Fernández, M.-J.; Vigil-Colet, A. Estructura factorial, fiabilidad y validez de la escala de adicción al trabajo (WorkBAT): Versión española. *An. Psicol.* **2013**, *29*, 923–933. [\[CrossRef\]](#)
34. Santos, J.; Sousa, C.; Sousa, A.; Figueiredo, L.; Gonçalves, G. Psychometric evidences of the Workaholism Battery in a Portuguese sample. *J. Spat. Organ. Dyn.* **2018**, *6*, 40–51.
35. Urbán, R.; Kun, B.; Mózes, T.; Soltész, P.; Paksi, B.; Farkas, J.; Kökönyei, G.; Orosz, G.; Maráz, A.; Felvinczi, K.; et al. A four-factor model of work addiction: The development of the Work Addiction Risk Test Revised. *Eur. Addict. Res.* **2019**, *25*, 145–159. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Robinson, B.E.; Post, P.; Khakee, J.F. Test-retest reliability of the Work Addiction Risk Test. *Percept. Mot. Skills* **1992**, *74*, 926. [\[CrossRef\]](#)
37. Robinson, B.E.; Post, P. Validity of the Work Addiction Risk Test. *Percept. Mot. Skills* **1994**, *78*, 337–338. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Robinson, B.E.; Phillips, B. Measuring workaholism: Content validity of the Work Addiction Risk Test. *Psychol. Rep.* **1995**, *77*, 657–658. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Robinson, B.E.; Post, P. Split-half reliability of the Work Addiction Risk Test: Development of a measure of workaholism. *Psychol. Rep.* **1995**, *76*, 1226. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Robinson, B.E. Concurrent validity of the Work Addiction Risk Test as a measure of workaholism. *Psychol. Rep.* **1996**, *79*, 1313–1314. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Flowers, C.P.; Robinson, B. A structural and discriminant analysis of the Work Addiction Risk Test. *Educ. Psychol. Meas.* **2002**, *62*, 517–526. [\[CrossRef\]](#)
42. Taris, T.W.; Schaufeli, W.B.; Verhoeven, L.C. Workaholism in the Netherlands: Measurement and implications for job strain and work-nonwork conflict. *Appl. Psychol.* **2005**, *54*, 37–60. [\[CrossRef\]](#)
43. Romeo, M.; Yepes-Baldó, M.; Berger, R.; da Costa, F.F.N. Workaholism in Brazil: Measurement and individual differences. *Adicciones* **2014**, *26*, 312–320. [\[CrossRef\]](#)
44. Ravoux, H.; Pereira, B.; Brousse, G.; Dewavrin, S.; Cornet, T.; Mermillod, M.; Mondillon, L.; Vallet, G.; Moustafa, F.; Dutheil, F. Work Addiction Test questionnaire to assess workaholism: Validation of French version. *JMIR Ment. Health* **2018**, *5*, e12. [\[CrossRef\]](#)
45. del Líbano, M.; Llorens, S.; Salanova, M.; Schaufeli, W. Validity of a brief workaholism scale. *Psicothema* **2010**, *22*, 143–150.
46. Falco, A.; Kravina, L.; Girardi, D.; Dal Corso, L.; Di Sipio, A.; De Carlo, N.A. The convergence between self and observer ratings of workaholism: A comparison between couples. *Test. Psychom. Methodol. Appl. Psychol.* **2012**, *19*, 311–324. [\[CrossRef\]](#)
47. Molino, M.; Ghislieri, C.; Colombo, L. Working excessively: Theoretical and methodological considerations. *G. Ital. Med. Lav. Ergon.* **2012**, *34*, A5–A10. [\[PubMed\]](#)
48. Sharma, P.; Sharma, J. A confirmatory factor analysis of Dutch Work Addiction Scale (DUWAS). *Glob. Bus. Rev.* **2013**, *14*, 211–223. [\[CrossRef\]](#)
49. Littman-Ovadia, H.; Balducci, C.; Ben-Moshe, T. Psychometric properties of the Hebrew version of the Dutch Work Addiction Scale (DUWAS-10). *J. Psychol.* **2014**, *148*, 327–346. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Rantanen, J.; Feldt, T.; Hakanen, J.J.; Kokko, K.; Huhtala, M.; Pulkkinen, L.; Schaufeli, W. Cross-national and longitudinal investigation of a short measure of workaholism. *Ind. Health* **2015**, *53*, 113–123. [\[CrossRef\]](#)

51. Mir, I.; Kamal, A.; Masood, S. Translation and validation of Dutch Workaholism Scale. *Pakistan J. Psychol. Res.* **2016**, *31*, 331–346.
52. Azevedo, W.F.; Mathias, L.A. da S.T. Addiction to work and factors relating to this: A cross-sectional study on doctors in the state of Paraíba. *Sao Paulo Med. J.* **2017**, *135*, 511–517. [[CrossRef](#)]
53. Balducci, C.; Avanzi, L.; Consiglio, C.; Fraccaroli, F.; Schaufeli, W. A cross-national study on the psychometric quality of the Italian version of the Dutch Work Addiction Scale (DUWAS). *Eur. J. Psychol. Assess.* **2017**, *33*, 422–428. [[CrossRef](#)]
54. Nonnis, M.; Cuccu, S.; Cortese, C.G.; Massidda, D. The Italian version of the Dutch Workaholism Scale (DUWAS): A study on a group of nurses. *BPA Appl. Psychol. Bull.* **2017**, *65*, 47–57.
55. Traub, R.E. Classical Test Theory in historical perspective. *Educ. Meas. Issues Pract.* **1997**, *16*, 8–14. [[CrossRef](#)]
56. DeMars, C. *Item Response Theory*; Oxford University Press: New York, NY, USA, 2010.
57. Sharkness, J. Item Response Theory: Overview, applications, and promise for institutional research. *New Dir. Institut. Res.* **2014**, *2014*, 41–58. [[CrossRef](#)]
58. Bond, T.G.; Yan, Z.; Heene, M. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, 4th ed.; Routledge: New York, NY, USA, 2021.
59. Sireci, S.; Faulkner-Bond, M. Validity evidence based on test content. *Psicothema* **2014**, *26*, 100–107. [[CrossRef](#)] [[PubMed](#)]
60. Bowen, D.J.; Hyams, T.; Goodman, M.; West, K.M.; Harris-Wai, J.; Yu, J.-H. Systematic review of quantitative measures of stakeholder engagement. *Clin. Transl. Sci.* **2017**, *10*, 314–336. [[CrossRef](#)] [[PubMed](#)]
61. Boon, C.; Den, D.N.; Lepak, D.P. A systematic review of human resource management systems and their measurement. *J. Manag.* **2019**, *45*, 2498–2537. [[CrossRef](#)]
62. Shkoler, O.; Rabenu, E.; Iqbal, M.Z.; Ferrari, F.; Hatipoglu, B.; Roazzi, A.; Kimura, T.; Tabak, F.; Moasa, H.; Vasiliu, C.; et al. Heavy-Work Investment: Its dimensionality, invariance across 9 countries and levels before and during the COVID-19's pandemic. *J. Work Organ. Psychol.* **2021**, *37*, 67–83. [[CrossRef](#)]