ASSESSING THE INFLUENCE OF REVERSED ITEMS AND FORCE-CHOICE ON THE WORK AND MEANING INVENTORY

EVALUACIÓN DE LA INFLUENCIA DE LOS ÍTEMS INVERTIDOS Y DE ELECCIÓN FORZOSA EN EL INVENTARIO DE TRABAJO SIGNIFICATIVO

MARIA DA GLÓRIA LIMA LEONARDO[1],
MICHELLE MORELO PEREIRA[1,2], FELIPE VALENTINI[3],
CLARISSA PINTO PIZARRO DE FREITAS[4], AND
MICHAEL F. STEGER[5,6]

Abstract

Response biases are issues in inventories in positive organizational psychology. The study aims to control the response bias in the assessment of meaning of work through two methods: reversed key items and forced-choice format. The sample consisted of 351 professionals; women constituted 60.0 % of the sample. The participants answered two versions of the instrument for meaning of work: Likert-type items and forced-choice. For both versions, the unifactorial model was the most appropriate for the data available. The results indicate that the random intercepts model fit the Likert data (CFI = .92), as well as

Correspondence address [Dirección para correspondencia]: Maria da Glória Lima Leonardo. Universidade Salgado de Oliveira (UNIVERSO), Brasil.
Email: bgdgloria@gmail.com

ORCID: Maria da Glória Lima Leonardo (https://orcid.org/0000-0003-4966-4169), Michelle Morelo Pereira (https://orcid.org/0000-0003-2437-2071), Felipe Valentini (https://orcid.org/0000-0002-0198-0958), Clarissa Pinto Pizarro De Freitas (https://orcid.org/0000-0002-2274-8728), and

[1] Universidade Salgado de Oliveira (UNIVERSO), Brasil.
[2] Universidade do Estado de Minas Gerais (UEMG), Brasil.
[3] Universidade São Francisco (USF), Brasil.
[4] Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Brasil.
[5] Center for Meaning and Purpose, Colorado State University, USA.
[6] North-West Univeristy, South Africa.

the forced-choice model (CFI = .97). Besides, the latent dimension of the forced-choice version did not correlate with acquiescence index ($r < .08$; $p > .05$), and approximately 20 % of the variance of the items might be due to the method (Likert or forced-choice). The present study illustrates the importance of response bias control in self-report instruments.

Palabras clave: work; meaning at work; acquiescence; forced-choice items.

Resumen

Los sesgos de respuesta son problemas en los inventarios de la psicología organizacional positiva. El estudio tiene como objetivo controlar el sesgo de respuesta en la evaluación del trabajo significativo a través de dos métodos: ítems clave invertidos y formato de elección forzosa. La muestra estuvo formada por 351 profesionales; las mujeres constituyeron el 60.0 % de la muestra. Los participantes respondieron dos versiones del instrumento de significado del trabajo: ítems tipo Likert y elección forzosa. Para ambas versiones, el modelo unifactorial fue el más apropiado para los datos disponibles. Los resultados indican que el modelo de intersecciones aleatorias se ajusta a los datos Likert (CFI = .92), así como al modelo de elección forzada (CFI = .97). Además, la dimensión latente de la versión de elección forzada no se correlacionó con el índice de aquiescencia ($r < .08$; $p > .05$), y aproximadamente el 20 % de la varianza de los ítems podría deberse al método (Likert o forzado). elección). El presente estudio ilustra la importancia del control del sesgo de respuesta en los instrumentos de autoinforme.

Keywords: trabajo; trabajo significativo; aquiescencia; ítems de elección forzosa.

Introduction

Although investigations on the meaning and purpose people perceive in their work are multiplying rapidly, research is still incipient and insufficient to clarify how to maximize the potential of this construct, in terms of its ap-

parent benefits for the individual, as well as for the organization (Jena et al., 2019; Steger et al., 2012). For the individuals, meaningful work contributes because it generates well-being and psychological adjustment. Professionals who see their work to be meaningful also perceive greater meaning and significance in their lives as a consequence of self-understanding of themselves and the world, enabling their personal growth (Allan, 2017; Steger et al., 2012). In the organization, meaningful work is associated with higher involvement, better relations in workplace, productivity, and performance (Jena et al., 2019).

Meaningful work commonly is assessed through the Work and Meaning Inventory (WAMI; Steger et al., 2012). The WAMI was developed in the United States of America (Steger et al., 2012) and has already been adapted for use in Turkey (Akin et al., 2013), South Africa (Finch, 2014) and Brazil (Leonardo et al., 2019).

Although the Brazilian study of WAMI found evidence of validity and reliability of the scale, the results also indicated that biases due to the method (e.g., acquiescence) could distort the representativeness and precision of the instrument to investigate meaningful work (Leonardo et al., 2019). The use of forced-choice items may shed light on the degree to which acquiescence due to homogeneous response biases influences measurement of meaningful work by the WAMI. Thus, the main objective of this study is to investigate if the presence of acquiescence bias in the assessment of meaning of work may be controlled through forced-choice items and reversed Likert Items.

It is also noteworthy the WAMI utilizes a preponderance of very positive items (e.g., "I found a job that is fulfilling"), which may be fitting for a construct whose descriptive content is generally positive. However, traditional Likert-type scales do not partial out the positive content of the construct from the general endorsement trend, leaving unaddressed the possibility of uncontrolled response biases. The deployment of forced-choice response options can eliminate the need for additional control of response biases (Brown & Maydeu-Olivares, 2012, 2018). For this reason, the present study is innovative in proposing an evaluation of meaningful work that compares forced-choice items and the assessment of the construct by the original version of WAMI (Likert type) both

with and without the additional control of acquiescence bias. The aim of this study is to further the development of reliable scales to assess meaningful work, which will more effectively contribute to producing evidence on research and interventions to promote well-being at work.

Work and Meaning Inventory (WAMI)

The WAMI was initially developed to measure three dimensions (Steger et al., 2012). The first dimension, labeled "positive meaning," refers to the degree to which people find meaning and purpose in their work. It reflects a subjective experience that the work has meaning, and it is relevant. The second dimension, labeled "meaning-making through work", implies scope, involvement in work, and harmonious impact of work on the personal lives of individuals (Steger et al., in press). The third dimension, labeled "greater good motivations", concerns the expectation that the work will contribute positively to the greater collective good (Steger et al., 2012).

Earlier research sought to validate the WAMI in a Brazilian sample of 667 professionals (74 % women, mean age 35.7, SD = 10.5). The results indicated that the Brazilian version of the inventory, unlike the original version which proposed three facotrs, presented better fit indices in the unifactorial structure ($\chi^2(28) = 144.76$; CFI = .99; TLI = .99; RMSEA (CI 90%) = .08 (.07 - .09), [internal consistency index of .94] (Leonardo et al., 2019). The factorial loads varied from .65 to .95 (VME = .70). When the three originally-proposed were calculated and analyzed, a pattern of correlations was observed that was consistent with patterns reported in other countries and language versions of the scale. The sense of purpose at work was positively associated with occupational self-efficacy [$r = .55$; $p < .05$], intrinsic motivation at work [$r = .77$; $p < .05$] and engagement at work [$r = .81$; $p < .05$] (Leonardo et al., 2019).

The failure to recover the three dimensions originally proposed, as well as the high convergent correlation coefficients observed with other work-related variables might have been influenced by the larger number of positively keyed items in the WAMI (Leonardo et al., 2019). Thus, there is a need for research with greater control over the influence of acquiescence bias in the responses of individuals.

Single Stimulus Inventories and Forced-Choice Inventories

In organizational psychology, where the assessments of psychological constructs depend on self-reported measures, scholars who use the quantitative approach often opt for employing the Likert scales. When using them, the respondent evaluates one item at a time and grades it separately. This response mode is called the single stimulus format (Brown & Maydeu-Olivares, 2012). As it is easy to answer, this survey method is the most popular; however, there are concerns about response biases within this method. One of the most frequently observed problems is the acquiescence bias, which refers to the tendency of endorsing positive Likert categories, despite the item content and the key, positive or reversed (Valentini, 2017).

In a previous study of the WAMI in a Brazilian sample, an analysis of the response pattern suggested that the poor discrimination among factors might have been due to the way the items are phrased and rated (Leonardo et al., 2019). The fact that the items are mostly positive keyed, answered using a five-point Likert scale, and exerted low or no control over the bias of the individuals' response style, likely reduced the representation of the items along the construct continuum.

The use of forced-choice responses has been suggested as a strategy to overcome the limitations of the single stimulus inventories precisely because forced-choice questions are designed to reduce, or even eliminate, response biases (Brown & Maydeu-Olivares, 2012). Unlike a single stimulus inventory, an inventory constructed to use forced-choice techniques is presented in blocks with two or three items designed to measure different attributes. Respondents classifies the items, ordering from the least to the most characteristic of their selves. This process has become innovative in item response theory approaches as a means to prevent response bias. Further, there is some evidence that it improves the fit of measurement models for assessment data (Brown et al., 2018; Valentini, 2017).

In this sense, if the response bias is homogeneous between items, it tends to be canceled in any forced-choice comparisons. For example, a participant, forced to choose between item A and item B, will respond comparing how much items A and B represent his/her characteristics (Representation of Item A - Item B). In this example, if any biases were the same or very similar between items A and B, they will tend to cancel each other (Item A + Bias - Item B - Bias = Item A - Item B) (Brown & Maydeau-Olivares, 2018). This is a theoretical framework for explain how the forced-choice items eliminates homogeneous bias, despite the coding method for the items.

Acquiescence

As noted earlier, forced-choice item formats are often used to control response bias. Acquiescence is one of the most frequently observed response biases. It may be understood as the individuals' behavior to always respond positively to the questionnaire, regardless of the descriptive content of the item (Billiet & McClendon, 2000).

The lack of control over acquiescence can jeopardize the interpretation of the scores and yield bias correlations between variables (Billiet & McClendon, 2000; Valentini, 2017). Some methods are available to control acquiescence, such as ipsatization and the random intercept. Ipsatization, also called intra-subject standardization, involves the change of the raw scores based on the mean and standard deviation (SD) of the positive and negative items of each subject. The Random intercepts uses statistical modeling of a general factor not correlated with the descriptive content of the items themselves (Maydeu-Olivares & Coffman, 2006; Valentini, 2017). Both ipsatization and random intercepts approaches assume the scale is composed of positive and negative keyed items.

The present study aims to evaluate different forms of measurement of meaningful work using a Brazilian language version of the WAMI. Specifically, the objectives are: to test models with one and three factor for assessing the meaningful work; to test if the acquiescence control yields better fit model and items representativeness; to test if a forced-choice scale is valid; and to correlate content trait with the method factor for acquiescence, estimated by random intercepts and classical scoring (ipsatization).

Method

Participants

The sample was of convenience and it was composed by 369 respondents, however 18 (5 %) questionnaires were excluded because they were not completed correctly. Most workers were single (49 %), 43% were married or in a stable relationship and 8 % were divorced. The inclusion criteria for this sample were: having been working for at least 12 months; and be over 18 years old. The final sample of the study was composed by 351, of which 60.9 % were women. Participants age ranged from 18 to 74 years, the majority was 25 to 34 years (33 %), followed by 18 to 24 years (27 %), 35 to 44 years (20 %), 55 to 64 years (9 %), 45 to 54 years (8 %) and 65 to 74 years (3 %). As for education, most individuals held a bachelor's degree (51 %), 18% held a high school degree, 16 % held a specialization degree and 15% held a master or a PhD degree. The average length of service in the current job is 6.2 years (SD = 7.1 years, ranging from 1 to 43 years), and the total work time is 12.2 years (SD = 11.44 years, ranging from 1 to years 61). The majority of participants was withe-collar workers (72 %) and 28 % were blue-collars workers.

Instruments

Workers answered a general questionnaire with self-report questions and sociodemographic data. In this study, two strategies were also used to control the acquiescence and response style of the traditional WAMI. For such purpose, two adaptations of the original version of the Brazilian version of WAMI were made (Leonardo et al., 2019), an 18-item version of the WAMI and a forced-choice version of the WAMI (Appendix 1).

The original WAMI has ten items (Steger et al., 2012). Items are answered on a five-point Likert scale, ranging from 1 (totally false) to 5 (totally true). The original version of WAMI (Steger et al., 2012) showed excellent psychometric proprieties ($\chi2(gl) = (30)$ 64.2, CFI = .96, NFI = .95, RMSEA (90% C.I.) = .09 (0.06 – 0.11); a = .93). The same was observed on the Brazilian version of WAMI (Leonardo et al., 2019), that had great fit index ($\chi2(gl) = (28)$ 144.8, CFI = .99, NFI = .99, RMSEA (90 %

C.I.) = .08 (0.07 – 0.09) and excellent internal consistency (a = .94). Example of an item: "I have found a meaningful career".

We add eight reverse-scored items to the original WAMI for controlling response bias. These statements were evaluated by experts so that the items had content opposite and equivalent to the original WAMI items. All questions were answered on a five-point Likert scale, ranging from 1 (totally false) to 5 (totally true). The internal consistency of the18 items (10 original items of the WAMI and 8 new reverse-scored items) in this study was satisfactory (a = .93). Example of an item: "My work makes me indifferent to others."

We also used a forced-choice format for assessing the meaning of work. The 18 items (10 original items of the WAMI and 8 reversed) were set into six blocks composed of three items, for which the respondent should mark the item that more describes him and the one that least describes his work. Each block was composed of items with the most distinct factorial loadings possible, and negative items. This is a case of Most and Least format (MOLE), and for triplets, it is equivalent to a full ranking, as the sentence which was not marked is assumed to be classified between the least and the most. We provide further details about coding the blocks in the section Data Analysis. The Block example:

Data Collection Procedure

The project was approved on the Research Ethics Committee of the first author's Institution. The participants were recruited using a convenience sampling technique. The participants were contacted in different ways, such as social and professional media networks (e.g., Facebook, LinkedIn), the HR departments of institutions and organizations. All participants answered the questionnaire after agreeing with the Consent Form. Overall, 55% answered the instruments in the paper-and-pencil form, whereas the remaining 45% of the total sample responded the questionnaires on a web-based platform (i.e., Qualtrics). The paper-and-pencil data collection form was applied to the academic staff of a public university, during working hours or breaks by a member of the research team.

Data Analysis

The adequacy of two different models for the instruments was investigated: The one-dimensional model identified by Leonardo et al. (2019), and the model proposed by Steger et al. (2012) for the original WAMI, consisting of three first-order factors. The adequacy of the one-dimensional model and of three first-order oblique factors was investigated both without controlling for the effects of acquiescence and with their control.

For the control of acquiescence, we used random intercept modeling (Maydeu-Olivares & Coffman, 2006), considering 18 balanced items (9 positive and 9 reversed). We model the response bias as a latent orthogonal variable to the content factors. The factor loadings of the bias factor (or random intercept) were fixed at +1, for both positive and negative items, thus capturing the tendency to agree indiscriminately with all items. In addition to the random intercepts, we calculated the classic indicator of acquiescence by averaging all items (positive and negative, without inverting them). Thus, if the participant endorsed both positive and negative items, despite of the item key, the average of the items would be positive, indicating acquiescence.

For the forced-choice blocks, we applied the Thurstonian-IRT model, T-IRT (Brown & Maydeu-Olivares, 2018). First, we code the answers within each block into binary comparisons of items: {item A x item B}, {item A x item C}, and {item B x item C}. For each binary, a code of 1 is applied if the participant preferred the first item over the second; and a code of 0 is applied for the opposite preference. Table 1 shows an example of coding for one block of items. The coding procedure and the model (Brown & Maydeu-Olivares, 2018) solve the classic issue of ipsative scores, which makes factor analysis feasible for this type of data. The model assumes the preferences are due the differences in the item utility (T parameter), defined as the value that the subject attributes to the item sentence. Participant will prefer item A over B if the utility for the item A is higher than the

Table 1.

Example of coding the forced-choice items.

Items (within a block)	Participant 1		Ranking	Code
	Least	Most		
A. I found a fulfilling job.		X	(A, C, B)	{A,B} = 1
B. My work is irrelevant to the world.	X	⇒	⇒	{A,C} = 1
C. My work helps me to understand myself better				{B,C} = 0

utility for B (i.e., {A, B} = 1, if TA > TB), and so forth. Utility is predicted by the item parameters and the latent trait (i.e., T = intercept + loading*trait + error). In other words, utility (T) works as first order factor between the observed binary comparison and the content latent factor, which is modeled as a second order dimension. Consult Brown and Maydeu-Olivares (2018) for further details.

We use the ULSMV estimator (Muthén & Muthén, 2010) as it is the standard for forced-choice (Brown & Maydeu-Olivares, 2018), and all items were declared as categorical, including those in the Likert format. As the estimator used the information of item frequency, we did not test the items' distribution. Some models we tested for forced-choice yielded improper solution (for instance, negative variance). In order to solve this problem, we constrained the problematic variances of the forced-choice items as equal to the variance of the Likert format. This strategy was utilized by Guenole, Brown and Cooper (2018) as well.

The goodness of fit indices used established that for the model to be considered adequate, the comparative fit index (CFI) and Tucker-Lewis index (TLI) should have values higher than .95, the root mean square error of approximation (RMSEA) values less than .08, with a confidence interval of 90% lower than .10 (Brown, 2015). The thresholds (transition from one category to another) of the Likert items were investigated both with and without the control of acquiescence. This analysis aimed to examine whether the thresholds would show more significant variability after controlling for the items' acquiescence effects.

Results

Confirmatory Factor Analysis

Four confirmatory factor analyses were performed to investigate which structure constitutes the best solution.

Table 2.

Goodness of fit of the Confirmatory Factor Analysis of Unifactorial and Three Factors of First Order, with and without Control of Acquiescence

Models	χ^2 (*df*)	CFI	TLI	RMSEA (CI 90%)	Correlations between Factors			
					M3		M4	
1fat	**874.25* (135)**	**.91**	**.90**	**.12 (.11 – .13)**				
1fat + acq	442.98* (134)	.96	.96	.08 (.07 – .09)	PMxMM	.90*	PMxMM	.92*
3fat	783.71* (132)	.92	.91	.12 (.11 – .13)	PMxGM	.98*	PMxGM	.97*
3fat + acq	359.05* (131)	.97	.97	.07 (.06 – .08)	MMxGM	.88*	MMxGM	.87*

Note: * = *p* < 0.001; *χ²*= Chi-square; *df* = Degrees of freedom; TLI = Tucker-Lewis Index*;* CFI = Comparative Fit Index*;* RMSEA = Root Mean Square Error of Approximation; 1fat = Unifactorial Structure Without Acquiescence Control; 1fat + acq = Unifactorial Structure Controlled by Acquiescence, estimate by Random Intercept; 3fat = First Order Oblique Factors Structure Without Acquiescence Control; 3fat + acq = Structure of Three First Order Factors Controlled by Acquiescence, estimate by Random Intercept; PM = Positive meaning; MM = Meaning making through work; GM = Greater good motivation.

Table 3.

Analysis of Item Parameter with and without Acquiescence Control.

Items	M1 – Without acquiescence control					M2 – Control of acquiescence (random intercepts)				
	Ld	*Thresholds*				Ld	*Thresholds*			
		δ_{j1}	δ_{j1}	δ_{j1}	δ_{j1}		δ_{j1}	δ_{j1}	δ_{j1}	δ_{j1}
1	.77	-2,03	-1,58	-0,21	0,77	.75	-2.20	-1.70	-23	.83
2	.71	-2,59	-1,81	-1,16	-0,04	.71	-2.76	-1.92	-1.24	-.04
3	-.77	0,59	1,11	1,74	2,13	-.71	.67	1.26	1.99	2.43
4	.86	-2,22	-1,51	-0,55	0,58	.69	-2.38	-1.61	-.58	.62
5	.61	-2,91	-2,29	-1,30	-0,02	.75	-3.12	-2.46	-1.40	-.02
6	.83	-2,67	-1,89	-0,82	0,24	.75	-2.88	-2.04	-.88	.26
7	.74	-1,90	-1,30	-0,49	0,72	.60	-2.05	-1.40	-.53	.77
8	.62	-2,89	-2,24	-1,06	0,64	.84	-2.90	-2.24	-1.07	.65
9	.74	-2,31	-1,98	-0,70	0,68	.72	-2.48	-2.13	-.75	.74
10	.65	-2,42	-1,73	-0,72	0,48	.76	-2.63	-1.88	-.79	.52
11	-.83	0,75	1,29	1,73	2,03	-.62	.85	1.47	1.97	2.32
12	-.88	1,67	2,25	3,22	3,51	-.83	1.89	2.55	3.64	3.97
13	-.80	0,88	1,41	1,99	2,64	-.65	.98	1.58	2.23	2.95
14	-.70	1,08	1,79	2,67	3,64	-.82	1.15	1.91	2.84	3.88
15	-.76	1,10	2,00	2,97	3,80	-.86	1.17	2.13	3.15	4.03
16	-.77	1,10	1,80	2,58	3,16	-.78	1.17	1.91	2.75	3.37
17	-.83	1,32	1,99	2,60	3,23	-.82	1.45	2.18	2.86	3.55
18	-.70	-0,43	0,10	0,84	1,42	-.72	-.51	.13	1.00	1.70

Note: all parameter were significant (*p* < 0.001); *Model 1* = Unifactorial Structure Without Acquiescence Control; Model 2 = Unifactorial Structure Controlled by Acquiescence; Ld = Factorial Loading.

The tested models were unifactorial and with three factors, as well as controlling and without the control for acquiescence.

It was observed that for all solutions, the items presented adequate loadings. The models controlling for acquiescence outperformed those without controlling. The First Order Three-Factor Model presented fit indexes superior to the Unifactorial Model. However, the correlations between the dimensions of the Three-Factor Model where higher than the average of the factor loadings, indicating low discrimination between the factors (Farell, 2010; Va-

lentini & Damásio, 2016). These results support the use of both models. Considering the goodness of fit and the theoretical framework, we recommend to use the three factor model, when the discrimination between factors is not an issue; and we suggest use the one factor model, when the multicollinearity poses a threat to the analysis. In both models, we suggest controlling for acquiescence bias (Table 2).

As anticipated, analysis of item parameters showed that the control of the acquiescence contributed to improving the variability of the items' threshold parameters (Ta-

ble 3). It was observed that the acquiescence control allowed an increase in the representativeness of the items. That is, without the acquiescence control, the thresholds ranged from -2.91 to 3.80; and controlling for the response bias, the thresholds ranged from -3.12 to 4.03.

It is important to note that the increase in the items' representativeness can be observed both for negative items, as well as for positive items. The results indicated that was possible to differentiate extreme low scores to extremes high scores. By including the control of the effects of acquiescence, the instrument became sensitive to discriminate participants with small, medium, and high indices of the construct meaning of work (Table 3).

Por el contrario, ni la inducción de emociones $(F(7,15) = 1.42, p = .26, \eta^2 = .39)$ ni la interacción (momento de medida*inducción emocional) $(F(7,15) = .75, p = .63, \eta^2 = .26)$, tuvieron una influencia significativa en el rendimiento en las tareas emocionales. Por lo tanto, las diferencias significativas entre la primera y la segunda aplicación no se pueden explicar por la inducción de emociones.

En cuanto a la inducción o no de emociones neutras, el efecto del factor intersujetos no es significativo para ninguna de las tareas del estudio: la tarea de reconocimiento $(F(1,21) = .01, p = .92, \eta^2 = .00)$, la tarea de identificación $(F(1,21) = .47, p = .49, \eta^2 = .02)$, la tarea de discrimina-

Table 4. Analysis of the Forced-choice Version

	Item	One-dimensional	Loading (Standard Error)		
			Three Factors		
			F1	F2	F3
Block1					
	1	.71 (fixed)	.89 (.03)		
	7	.29 (.19)		.52 (.08)	
	17	-.97 (.03)			-.63 (.09)
Block2					
	4	.71 (fixed)	.75 (.06)		
	16	-.95 (.06)		-.96 (.05)	
	6	.69 (.07)			.69 (.07)
Block3					
	5	-.71 (fixed)	-.71 (fixa)		
	9	.42 (.09)		.66 (.09)	
	10	.55 (.07)			.56 (.07)
Block4					
	5	.71 (fixed)	.86 (.06)		
	13	-.71 (.17)		-.11 (.17)	
	3	-.91 (.09)			-.53 (.23)
Block5					
	11	-.71 (fixed)	-.92 (.04)		
	2	.75 (.09)		-.03 (.33)	
	18	-.35 (.18)			-.81 (.07)
Block6					
	14	-.71 (fixed)		.65 (.31)	
	15	-.70 (.12)		.40 (.35)	
	8	.87 (.06)	.93 (.03)		

		Correlations between Factors	
		F2	F3
	F1	.70	.99 *
	F2		.40

Note. * = Correlation when freely estimated implied a non-positive model matrix. Therefore, the three-dimensional model is not a plausible solution. One loading per block is fixed for identification, in case of unidimensional model. Table presents loadings from the item utility (T) to the content latent factor.

Table 5.

Correlations among content factors and acquiescence.

	Model with random intercept			Model without random intercept		
	1	2	3	1	2	3
1. Factor (Forced-Choice)						
2. Factor (Likert)	.88*			.88*		
3. Random Intercept	.02 (n.s.)	0 (fixed)		-	-	
4. Acquiescence (CTT)	.08 (n.s.)	.13*	.62*	.09 (n.s.)	.14*	-

*Note: * (p < 0.001); Random Intercept = method factor with all loadings fixed in 1 (including those for negative keyed items); Acquiescence (CTT) = calculate by averaging positive and negative items (without reversing the negative keyed items), it is also nominated as ipsatization.*

ción facial emocional ($F(1,21) = .18$, $p = .67$, $\eta^2 = .00$), la tarea de discriminación facial según la edad ($F(1,21) = .51$; $p = .48$, $\eta^2 = .02$), la tarea de identidad de emparejamiento a la muestra ($F(1,21) = .05$, $p = .81$, $\eta^2 = .00$) y la tarea emocional de emparejamiento a la muestra de la Batería de Emociones de Rojahn ($F(1,21) = 1.42$, $p = .24$, $\eta^2 = .06$). Además, el tamaño del efecto está por debajo de .10 (η^2) en todas las medidas.

.

Forced -Choice Items

In addition to the traditional Likert version, participants responded to the forced-choice version. First, a unifactorial model was tested, which fit well to the data [$c^2(df) = (129)\ 200.1$; RMSEA = .04; CFI=.96; TLI=.95]. This modeling is complex due to the number of estimated parameters, and it was necessary to impose four additional restrictions for the identification of an appropriate solution. We use the parameters from the Likert format to constraint the estimation in the forced-choice model.

Considering that a utility (T) is estimated for each item, the relationship between the utilities and the content factor (second-order) can be interpreted as the traditional factor loadings. Table 4 shows the factorial loads of the utilities, which represent the items.

Note that, in general, the loadings were high and support the composition of the blocks. It should be noted that these values are block-dependent and may vary if the item is relocated to another block. In this context, it is suggested

that, in future versions, items 8 and 9 could be reallocated to different blocks.

Considering that the original version of WAMI postulates three first-order dimensions, an attempt was made to test model for forced-choice items based on the three-factor structure. The model adjusted to the data [$c^2(df)= 174.8$ (121); RMSEA = .04; CFI= .97; TLI= .96], however correlations between the latent factors were high, and between dimensions 1 and 3 it was higher than 1. Consequently, the model matrix was not positive. To overcome the problem, we constrained the correlation to a value lower than .99. Furthermore, two utilities presented positive loadings when they should be negative for the construct of work meaning (items 14 and 15). Again, the models with one and three factors as plausible for the forced-choice instrument.

Finally, we sought to examine correlations among estimated scores using the forced-choice scale and the Likert scale both with and without acquiescence control through random intercepts. The factors of both models were correlated with the classic indicator of acquiescence (TCT- an average of positive and negative items, without reversing them). The latent correlations are shown in Table 5.

The correlation between the factors estimated by the forced-choice items and Likert-type items was strong (.88) even without the control of random intercepts. Thus, at least in this sample, both versions share most part of the variance. However, the method was responsible for explaining a significant part of the variance. Nevertheless, it is not necessarily due the acquiescence bias.

Regarding the response bias, a correlation close to 1 was expected between the random intercept and the classic indicator of acquiescence, as they represent only different methods to estimate acquiescence. However, such association was only moderate (.62).

The relationship between the response bias and the factor estimated through forced-choice was not statistically significant, confirming that this format is not susceptible to problems of response idiosyncrasies (indicating simultaneously, positive and negative aspects of the meaning of the work). It is noteworthy the correlations between the classic indicator of acquiescence and the content factors estimated by forced-choice and Likert scale did not decrease after controlling for the random intercept. The result indicates the acquiescence is more related to items (like in the random intercept) than the latent construct, as methodological expected.

Discussion

In this study, we evaluated the factorial structures of WAMI, containing positive and negative items, and we modeled an acquiescence factor. Also, it was investigated if the use of a forced-choice rating system improved resulted in an accurate and valid instrument to measure meaningful work.

Concerning WAMI-18 for Likert-type items, although the three-dimensional model showed the best fit to the data, correlations between the factors were high. The goodness of fit indices of the unifactorial model were slightly below expectations; however, the factorial loads were high.. These results confirm a previous study carried out in Brazil (Leonardo et al., 2019), in which a one-dimensional structure for the original scale with ten items was also suggested. Of course, one possibility is that the one-dimensional structure emerges from the influence of the response bias. However, we also encourage researchers to use the three-dimensional model due its theoretical framework (Steager et al., 2012), as well as the fit to the data, whenever the multicollinearity between factors is not a threat.

This investigation deepened the discussion on the role of response bias in the internal structure of WAMI. We increased the number of items, with opposite pairs

Models showed that a significant part of the items' variance was explained by acquiescence. This result confirms the theoretical studies that indicate that the lack of control of acquiescence can result in bias on the items parameters (loadings and thresholds) and fit, jeopardizing the interpretability of the instrument's internal structure (Billiet & McClendon, 2000; Maydeu-Olivares & Coffman, 2006; Valentini, 2017).

The acquiescence control improved the interpretability of the items' difficulty. The thresholds were estimated in a more diversified manner, after controlling for the response bias. These findings suggest that after counting for acquiescence, the scale's representativeness is expanded to investigate low, medium, and high levels of meaning at work. On the other hand, when acquiescence is not controlled, there is a flattening of the participants' scores at the extremes of low scores and high scores, because there is a low differentiation of the parameters of difficulty.

For the forced-choice version, the models with one and three factors were plausible. The forced-choice WAMI presented adequate fit indices for the single-factor model, even though the three-factors showed better fit to the data. However, the three factors were strong correlated, and two loadings were inverted (i.e., they should be negative instead of positive).

One of the most important and original aspects of the present study refers to the comparison of the WAMI scores in the forced-choice and Likert versions. The correlation of the factor scores was high (.88), indicating that they are, in fact, the same latent construct. However, approximately 20 % of the variance (1 - r2) might be due to the method adopted in the versions. It is noteworthy that the latent correlations were estimated by SEM; therefore, the method effect (20 %) discounts the measurement error yet.

Part of this unshared variance, attributed to the method, might be due to the response bias. The factorial score of the Likert version exhibited correlations with the classic

indicator of acquiescence. On the other hand, in the forced-choice version, we did not find significant correlations between the factorial score and the acquiescence (neither of the classical estimate, nor the random intercepts). Such results also indicated that WAMI in the forced-choice version is not susceptible to acquiescence, as theoretically predicted (Brown & Maydeu-Olivares, 2018), and is a great choice to avoid response bias in self-report surveys. It is noteworthy, however, that items of forced-choice may be susceptible to other biases still lacking appropriate investigation.

Conclusions

In the present study, we made available a forced-choice version of the WAMI inventory and presented evidence supporting the control of response biases on Likert-type scales. Our findings support that both versions might be used as unidimensional or as a three dimensional structure; and acquiescence bias must be removed from the data.

The main contribution of the present study was to provide two reliable and valid scales to evaluate meaning of work with control of acquiescence bias. Such methodological tool can increase the usefulness of these instruments in scientific research and in the practice of job assessments. Furthermore, the possibility to discriminate the levels of meaning of work may support studies that aims to produce evidence to interventions focused on promoting well-being at work.

The strengths of the study include the robustness of the data analysis procedures. Furthermore, all analyses were performed with corrections for the characteristics of ordinal and nonscalar variables. However, the results of the study should be reviewed with caution, as it has some limitations, like the cross-sectional design and the use of a nonrepresentative sample.

The data collection at one point only is the first limitation of this study. The main limitation of the cross-sectional design is that it makes impossible to evaluate the variance of the construct over time. Since, it is not possible to compare the perceive meaning of the work at various time points in the professional career and in other organizations he/she has been.

A second limitation is the use of a nonrepresentative sample. Although the interviews involved professionals of different age groups and working time, it was observed that most of them had a high level of education. Future studies should investigate the performance of the WAMI-18 and WAMI with forced-choice items among workers with lower education.

It should be emphasized that we do not control cognitive biases, such as intelligence and working memory. Due to the complexity of the response, the forced-choice version may be less understood by less cognitively skilled participants, which could skew the scores.

Concerning a future investigation plan, a reduced version of the scale is suggested, since one of the items on the original scale did not obtain good factor loads. Besides, items that self-cancel should be replaced in different blocks. Furthermore, future investigations should address other types of potential bias specific for the forced-choice format, among which, the influence of intelligence.

Even so, it can be stated that WAMI-18 and WAMI with forced-response items are theoretical and practical relevance. Future use of the WAMI inventory, designed to measure the meaning of work, in the style of forced-response or WAMI-18 is suggested.

References

Akin, A., Hamedoglu, M. A., & Kaya, C. (2013). Turkish Version of the Work and Meaning Inventory (WAMI): Validity and Reliability Study. *Journal of European Education, 3*(2), 11–16.

Allan, B. A. (2017). Task Significance and Meaningful Work: A longitudinal study. *Journal of Vocational Behavior, 102,* 74–182. https://doi.org/10.1016/j.jvb.2017.07.011.

Billiet, J. B. & McClendon, M. J. (2000). Modeling Acquiescence in Measurement Models for Two Bal-

anced Sets of Items. *Structural Equation Modeling, 74,* 608–628. https://doi.org/10.1207/S15328007SEM0704.

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). The Guilford Press.

Brown, A. & Maydeu-Olivares, A. (2012). How IRT Can Solve Problems of Ipsative Data in Forced-Choice Questionnaires. 2013, *American Psychological Association,18*(1) 36–52. https://doi.org/10.1037/a0030641.

Brown, A. & Maydeu-Olivares, A. (2018). Modeling Forced-Choice Response Formats. In P. Irwing, T. Booth, & D. Hughes (Eds.), *The Wiley Handbook of Psychometric Testing* (pp. 523–570). Wiley-Blackwell.

Farrell, A. M. (2010). Insufficient Discriminant Validity: A Comment on Bove, Pervan, Beatty, and Shiu (2009). *Journal of Business Research, 63*(3), 324–327. https://doi.org/10.1016/j.jbusres.2009.05.003

Finch, J. D. (2014). *The Dimensionality of the Work and Meaning Inventory* [Doctoral Thesis, University of Johannesburg, South Afric]. UJContent. http://hdl.handle.net/10210/12007

Guenole, N., Brown, A., & Cooper, A. J. (2018). Forced-Choice Assessment of Work-Related Maladaptive Personality Traits: Preliminary Evidence From an Application of Thurstonian Item Response Modeling. *Assessment, 25*(4), 513–526. https://doi.org/10.1177/1073191116641181

Jena, L. K., Bhattacharyya, P., & Pradhan, S. (2019). Am I Empowered through Meaningful Work? The Moderating Role of Perceived Flexibility in Connecting Meaningful Work and Psychological Empowerment. *IIMB Management Review, 31*(3), 298–308. https://doi.org/10.1016/j.iimb.2019.03.010

Leonardo, M. G. L., Pereira, M. M., Valentini, F., Freitas, C., & Damásio, B. (2019). Adaptação do Inventário de Sentido do Trabalho (WAMI) para o contexto brasileiro [Adaptation of Work and Meaning Inventory (WAMI) to the Brazilian context]. *Revista Brasileira de Orientação Profissional, 20*(1), 79–89. http://dx.doi.org/10.26707/1984-7270/2019v20n1p79

Maydeu-Olivares, A. & Coffman, D. L. (2006). Random Intercept Item Factor Analysis. *Psychological Methods, 11,* 344–362. https://doi.org/10.1037/1082-989X.11.4.344.

Muthén, L. K. & Muthén, B. O. (2010). *Mplus: Statistical Analysis with Latent Variable. User's Guide.* Muthén & Muthén.

Rose, N. & Steger, M. F. (2017). *Führung, die Sinn macht. Organisationsentwicklung–Zeitschrift für Unternehmensentwicklung und Change Management* (04), 41-45.

Steger, M. F., Dik, B. J., & Shim, Y. (in press). Assessing Meaning and Satisfaction at Work. In S. J. Lopez (Ed.), *The Oxford Handbook of Positive Psychology Assessment* (2nd Ed.). Oxford University Press.

Steger, M. F., Dik, B. J., & Duffy, R. D. (2012) Measuring Meaningful Work: The Work and Meaning Inventory (WAMI). *Jornal of Career Assessment, 20*(3) 322–337. https://doi.org/10.1177/1069072711436160.

Valentini, F. (2017). Influência e controle da aquiescência na análise fatorial. *Avaliação Psicológica, 16* (2), 120-251. https://doi.org/10.15689/ap.2017.1602.ed.

Valentini, F. & Damásio, B. F. (2016). Variância Média Extraída e Confiabilidade Composta: Indicadores de Precisão [Average Variance Extracted and Composite Reliability: Reliability Coefficients]. *Psicologia: Teoria e Pesquisa, 32*(2). https://doi.org/10.1590/0102-3772e322225

Appendix

Work and Meaning Inventory with Control of Aquiescence

(1) Totalmente Falsa [Absolutely Untrue]	(2) Geralmente Falsa [Mostly Untrue]	(3) Nem falsa nem verdadeira [Neither True nor Untrue]	(4) Geralmente verdadeira [Mostly True]	(5) Totalmente verdadeira [Absolutely True]

1. Encontrei um trabalho realizador [a]   (1) (2) (3) (4) (5)
[I have found a meaningful career]
2. Meu trabalho contribui para o meu desenvolvimento pessoal [a] (1) (2) (3) (4) (5)
[I view my work as contributing to my personal growth]
3. Meu trabalho não faz nenhuma diferença para o mundo [a]   (1) (2) (3) (4) (5)
[My work really makes no difference to the world]
4. Eu percebo como o meu trabalho contribui para o sentido da minha   (1) (2) (3) (4) (5)
vida [a]
[I understand how my work contributes to my life's meaning]
5. Eu tenho uma clara noção do que faz meu trabalho ser significativo [a] (1) (2) (3) (4) (5)
[I have a good sense of what makes my job meaningful]
6. Eu sei que o meu trabalho faz uma diferença positiva no mundo [a] (1) (2) (3) (4) (5)
[I know my work makes a positive difference in the world]
7. Meu trabalho me ajuda a me entender melhor [a]   (1) (2) (3) (4) (5)
[My work helps me better understand myself]
8. Eu descobri um trabalho que tem um propósito satisfatório [a] (1) (2) (3) (4) (5)
[I have discovered work that has a satisfying purpose]
9. Meu trabalho me ajuda a compreender o mundo ao meu redor [a]   (1) (2) (3) (4) (5)
[My work helps me make sense of the world around me]
10. Meu trabalho tem um propósito maior [a]   (1) (2) (3) (4) (5)
[The work Ido serves a greater purpose]
11. O meu trabalho poderia ser substituído por uma máquina [b]   (1) (2) (3) (4) (5)
[My work can be replaced by a machine]
12. Meu trabalho é desnecessário [b]   (1) (2) (3) (4) (5)
[My work is unnecessary]
13. Meu trabalho me torna indiferente em relação aos outros [b]   (1) (2) (3) (4) (5)
[My work makes me indifferent towards others]
14. Meu trabalho prejudica o meu autoconhecimento [b]   (1) (2) (3) (4) (5)
[My harmful work or my self-knowledge]
15. Meu trabalho limita a minha visão de mundo [b]   (1) (2) (3) (4) (5)
[My work limited to my worldview]
16. Meu trabalho me torna superficial [b]   (1) (2) (3) (4) (5)
[My work makes me superficial]
17. Meu trabalho é irrelevante para o mundo [b]   (1) (2) (3) (4) (5)
[My work is irrelevant to the world]
18. Estou neste trabalho apenas por questões financeiras [b]   (1) (2) (3) (4) (5)
[I'm in this job just for financial reasons]

Note: a – Items of the WAMI (Steger, et al., 2012); b – Items developed on the present study to control the acquiescence.