

The Development and Initial Validation of the Decent Work Scale

Ryan D. Duffy
University of Florida

Blake A. Allan
Purdue University

Jessica W. England
University of Florida

David L. Blustein
Boston College

Kelsey L. Autin and Richard P. Douglass
University of Florida

Joaquim Ferreira and Eduardo J. R. Santos
University of Coimbra

Decent work is positioned as the centerpiece of the recently developed Psychology of Working Theory (PWT; Duffy, Blustein, Diemer, & Autin, 2016). However, to date, no instrument exists which assesses all 5 components of decent work from a psychological perspective. In the current study, we developed the Decent Work Scale (DWS) and demonstrated several aspects of validity with 2 samples of working adults. In Study 1 ($N = 275$), a large pool of items were developed and exploratory factor analysis was conducted resulting in a final 15-item scale with 5 factors/subscales corresponding to the 5 components of decent work: (a) physically and interpersonally safe working conditions, (b) access to health care, (c) adequate compensation, (d) hours that allow for free time and rest, and (e) organizational values that complement family and social values. In Study 2 ($N = 589$), confirmatory factor analysis demonstrated that a 5-factor, bifactor model offered the strongest and most parsimonious fit to the data. Configural, metric, and scalar invariance models were tested demonstrating that the structure of the instrument did not differ across gender, income, social class, and majority/minority racial/ethnic groups. Finally, the overall scale score and 5 subscale scores correlated in the expected directions with similar constructs supporting convergent and discriminant evidence of validity, and subscale scores evidenced predictive validity in the prediction of job satisfaction, work meaning, and withdrawal intentions. The development of this scale provides a useful tool for researchers and practitioners seeking to assess the attainment of decent work among employed adults.

Public Significance Statement

This study describes the development of a self-report measure of decent work, entitled the Decent Work Scale (DWS), which was designed based on guidelines offered by the International Labor Organization (ILO). This measure will be useful for researchers, practitioners, and policy analysts who are interested in assessing the quality of work that people experience.

Keywords: decent work, instrument development, psychology of working

The recently developed Psychology of Working Theory (PWT; Duffy, Blustein, Diemer, & Autin, 2016) aims to capture the career development process of adults across the range of

social, cultural, and economic backgrounds. The theory specifically attempts to document the effects of economic constraints and experiences of marginalization on the career development process and places the securement of decent work as the central variable within the model (Blustein, Olle, Connors-Kellgren, & Diamonti, 2016; Duffy et al., 2016; International Labor Organization [ILO], 2008). Duffy et al. (2016) defined and conceptualized decent work on the basis of standards developed by the ILO (ILO, 2008, 2012). Specifically, the authors note that decent work consists of: “(a) physically and interpersonally safe working conditions (e.g., absent of physical, mental, or emotional abuse), (b) hours that allow for free time and adequate rest, (c) organizational values that complement family and social values, (d) adequate compensation, and (e) access to adequate health care” (p. 130). The theory assumes decent work is attained when all of these components exist.

This article was published Online First February 6, 2017.

Ryan D. Duffy, Department of Psychology, University of Florida; Blake A. Allan, Department of Educational Studies, Purdue University; Jessica W. England, Department of Psychology, University of Florida; David L. Blustein, Counseling, Developmental, and Educational Psychology Department, Boston College; Kelsey L. Autin and Richard P. Douglass, Department of Psychology, University of Florida; Joaquim Ferreira and Eduardo J. R. Santos, Department of Educational Psychology, University of Coimbra.

Correspondence concerning this article should be addressed to Ryan D. Duffy, Department of Psychology, University of Florida, PO Box 112250, Gainesville, FL 32611. E-mail: rduf@ufl.edu

Duffy et al. (2016) proposed a number of predictors and outcomes of securing decent work. However, all of these propositions were considered speculative, as to date, no empirically supported instrument exists which captures an individual's experiences of all five components of decent work. In order to realize the full potential of the PWT, a psychometrically sound measure of decent work that assesses the internal, psychological experience of the quality of one's work life, in accordance with the criteria developed by the ILO (2008), is needed. The lack of a measure of decent work also represents a major limitation to practitioners who may seek to integrate the theory into their work with clients and may be interested in a free, efficient, user-friendly assessment. Moreover, the need for a self-report measure of decent work is consistent with recent critiques of this construct, which have observed that the majority of research and theory stemming from the Decent Work Agenda (articulated by the ILO [2008]), is based on macrolevel economic indices, thereby neglecting the important dimension of how people construct meaning from their experiences at work (Blustein et al., 2016; Deranty & MacMillan, 2012; Ribeiro, Silva, & Figueiredo, 2016).

The goal of the current study is to address this notable gap in the literature by developing a decent work measure with scores supporting evidence of reliability and validity. In the following sections, we briefly review our guiding theoretical framework, the PWT, discuss the conceptualization of decent work from an individual or psychological perspective, and articulate the research design and analytical plan for the current studies guided by best practice recommendations in scale development (Heppner, Wampold, Owen, Wang, & Thompson, 2015; Kahn, 2006).

Theoretical Framework

The PWT was developed, in part, as a reaction to the emphasis placed on person-environment fit and fulfillment of personal interests and values that prevail in many of the dominant career theories (Blustein, 2006; Blustein, Kenna, Gill, & DeVoy, 2008; Duffy et al., 2016). Although PWT theorists do not discount the importance of these factors, they assert that previous theories operate under the assumption that people have relative access to choose work that fits their value systems and skillsets. However, in reality, most people face significant barriers in securing work that will be a good "fit" with their dreams and aspirations.

In the PWT model, decent work is placed as a central mediating variable between contextual factors and work and well-being outcomes (Duffy et al., 2016). Economic constraints and marginalization experiences are hypothesized to directly relate to an individual's ability to secure decent work. Additionally, a person's perceived freedom of work choice despite constraints (referred to as *work volition*; Duffy, Diemer, et al., 2012) and a person's ability to use resources to cope with work-related stressors (referred to as *career adaptability*) are proposed to partially mediate these relations. According to the PWT, increased work volition and career adaptability function to facilitate conditions in which people are able to pursue work of their choosing, rather than being constrained to work choices that do not encompass the five components of decent work (Duffy et al., 2016).

The theory also proposes that when people are able to choose work that provides safe conditions, time for leisure and rest, support for family and social values, adequate compensation, and

access to adequate health care, they are more likely to fulfill three basic human needs: those of survival and power, social connection, and self-determination (Allan, Autin, & Duffy, 2016; Blustein, 2006; Duffy et al., 2016). Survival and power needs refer to both biological needs (e.g., food, shelter, health care) and social capital that are needed to sustain access to opportunity structures (e.g., education systems). Social connection reflects the basic human need to belong and to be connected to others through social relationships. Self-determination reflects the need to engage in tasks where one experiences autonomy, competence, and relatedness in their work. Together, the fulfillment of these basic needs is proposed to ultimately lead to overall satisfaction at work and general well-being outside of work (Duffy et al., 2016).

Conceptualizing Decent Work

As the centerpiece of the PWT, it is important to understand how the construct of decent work is conceptualized at an individual level to best build an assessment tool to capture its unique components. Although decent work was originally constructed to address its availability at a societal level (see Blustein et al., 2016 for a review), one of the emerging critiques about the decent work construct is the relative lack of focus on individuals' experiences of their work lives (Deranty & MacMillan, 2012). Psychological research on how people experience the quality of their work has been conducted extensively within I/O psychology (Hammer & Zimmerman, 2011; Parker, 2014), and to a lesser extent, within vocational and counseling psychology (Lent & Brown, 2013). A close examination of the applied psychology literature reveals a deep interest in how people experience their work lives, with perhaps the most notable research occurring within the realm of job satisfaction (Saari & Judge, 2004). However, the job satisfaction literature has not been integrated with conceptions of decent work, reflecting a divide between the academic literature and the lived experience of people, who consider both individual and subjective factors along with objective indices in their appraisals of the adequacy of their work lives (Blustein et al., 2016; Blustein, Kozan, & Connors-Kellgren, 2013; Sharone, 2013).

Within the past few years, a number of psychologists have sought to explore decent work from a perspective that seeks to integrate social and economic factors with psychological factors. Guichard (2013) articulated a compelling case for considering decent work within vocational psychology, noting that it might be used to infuse a more explicitly social justice-oriented perspective into vocational psychology research and practice. In the Blustein et al. (2016) psychologically informed analysis of the decent work construct, the importance of soliciting the direct input from individuals about their experience of how well their work lives are meeting the criteria outlined by the ILO was highlighted, affirming similar recommendations articulated by Ribeiro et al. (2016) and Deranty and MacMillan (2012), among others. Given the emerging interest in understanding the internal, psychological dimensions of decent work and in weaving this construct into new theory development efforts (as reflected in the PWT), we believe that it is timely to develop a self-report measure of decent work that demonstrates evidence of reliability and validity.

The Present Studies

The goal of the present studies is to develop an efficient measure of decent work with scores supporting evidence of reliability and validity that captures all five components of the construct. To accomplish this goal, we followed a series of steps based on expert recommendations about psychometric scale construction (Heppner et al., 2015; Kahn, 2006) and gathered data from working adults in the United States through online surveys. Our overall strategy for this scale development project is based on a deductive approach wherein the conceptual definition of decent work, as articulated by the ILO, is used to govern the development of items, the decisions about statistical analyses, and the initial set of construct validity studies. In Study 1, we developed a large set of items, attained expert review of these items leading to addition, subtraction, and adaptation of items, and distributed the instrument to a large and diverse sample of working adults. Exploratory factor analysis (EFA) using a deductive approach seeking to find five factors was conducted, and items were pared down based on best practice recommendations (Kahn, 2006), resulting in a final, five-factor version of the instrument demonstrating reliability. In Study 2, we gathered data from a new, larger sample of working adults to examine validity evidence. Specifically, we used confirmatory factor analysis (CFA) to examine the factor structure (exploring several possible models), tested invariance models according to race/ethnicity, gender, income, and subjective social class, compared the total scale score and five subscale scores to existing instruments that assess similar content to demonstrate convergent evidence of validity, and examined divergent and predictive evidence of validity of the DWS subscale scores.

Study 1

Method

Participants. The sample for this study contained 275 adults with a mean age of 33.84 years ($SD = 10.32$). Participants self-identified as female ($n = 120$, 44.1%), male ($n = 149$, 54.8%), transgender ($n = 2$, 0.7%), and other ($n = 1$, 0.4%). According to the U.S. Census Bureau (2015) estimates, the population is made up of approximately 50.8% women and 49.2% men. Thus, our sample has more men than women compared with the national average. Regarding ethnicity, of those who responded, the sample self-identified as White/European American ($n = 202$, 68.5%), Black/African American ($n = 31$, 10.5%), Hispanic/Latina/o-American ($n = 21$, 7.1%), Asian/Asian American ($n = 18$, 6.1%), Native American ($n = 9$, 3.1%), Asian Indian ($n = 1$, 0.3%), Arab American ($n = 1$, 0.3%), Pacific Islander ($n = 2$, 0.7%), and other ($n = 1$, 0.3%). (Several participants selected multiple ethnicity categories.) The U.S. Census Bureau reports the population's race in categories of Hispanic origin (i.e., Hispanic individuals and non-Hispanic individuals); thus, the following race estimations include both Hispanic and Non Hispanic individuals: 77.1% white, 13.3% black/African American, 5.6% Asian, 1.2% American Indian and Alaska Native, 0.2% Native Hawaiian and other Pacific Islander, and 2.6% identifying with two or more races. In comparison, our sample has more white participants, but is similar to national averages in minority populations, with some of our estimates being slightly greater than national estimates. The sample

was largely employed full-time ($n = 206$, 77.7%) with the remaining participants employed part-time ($n = 38$, 14.3%), self-employed part-time ($n = 15$, 5.7%), self-employed full-time ($n = 4$, 1.5%), and those ($n = 2$, 0.8%) who chose not to answer. Education levels consisted of less than high school ($n = 1$, 0.4%), some high school ($n = 3$, 1.1%), high school graduate ($n = 29$, 10.5%), trade/vocational school ($n = 9$, 3.2%), some college ($n = 68$, 24.7%), undergraduate degree ($n = 122$, 44.4%), professional degree ($n = 41$, 14.9%), and those ($n = 2$, 0.8%) who chose not to answer. Current participant social class levels were as follows: lower class ($n = 14$, 5.1%), working class ($n = 93$, 34.1%), middle class ($n = 146$, 53.5%), upper-middle class ($n = 15$, 5.5%), and upper class ($n = 5$, 1.8%).

Instrument. A large pool of items were written to capture the five components of decent work: (a) physically and interpersonally safe working conditions, (b) access to health care, (c) adequate compensation, (d) hours that allow for free time and rest, and (e) organizational values that complement family and social values. These items were written by a five-member research team consisting of an associate professor in counseling psychology, an assistant professor in counseling psychology, and three doctoral students in counseling psychology. All members of the team had extensive knowledge of the psychology of working perspective, the PWT, and had conducted numerous research projects using these theoretical frameworks. After achieving group consensus, a group of 47 items were then sent for review to two full and one associate professor in counseling psychology who were also experts in the psychology of working. Specifically, two of these professors were from international universities, given a preliminary goal to make the wording and conceptual framework of the instrument potentially applicable to non-U.S. samples and non-English speakers. The experts were instructed to inspect the items regarding their adherence to one of the five components of decent work and their overall quality. This resulted in additions or adaptation of seven items. After this process was complete, the scale consisted of 53 items addressing all five aspects of decent work. In the current study, participants answered each of these items on a 7-point Likert scale ranging from *strongly disagree* to *strongly agree*. Several of the items were negatively valenced, and once these were recoded, higher scores on each subscale and the total scale were indicative of higher levels of decent work.

Procedure. First, IRB approval was received was the first author's institution to conduct this study. Next, participants were recruited through Mechanical Turk (MTurk), which is an online data collection service where individuals can complete research tasks for payment. Collecting data through this method has become extremely popular within psychology (see Buhrmester, Kwang, & Gosling, 2011; Crump et al., 2013; Mason & Suri, 2012 for reviews). For the current study, a link to this survey was posted on MTurk for individuals who were adults and currently employed, and participants were instructed the survey would take 10–15 min. Participants were paid \$.40 for completing the survey. An initial sample of 300 participants started the survey. Within the survey, two items were inserted which required participants to answer a specific way in order to ensure they were paying attention. Twenty-five participants who incorrectly answered these items were removed from the dataset. This resulted in a final total of 275 participants, all of whom had complete data. Because we used

online data collection, any duplicate responses were searched for and removed from the dataset.

Results

EFA was used to narrow down the initial pool of 53 items. Specifically, consistent with the deductive approach detailed earlier, our goal was to extract five factors corresponding to the five components of decent work, and as noted above, the items were developed to address one of these five unique components. Additionally, we sought to develop an efficient scale in terms of length to maximize research and practical utility. We followed recommended EFA practices within counseling psychology (Heppner et al., 2015; Kahn, 2006) and took the following steps. First, using SPSS, we conducted an EFA with a principal axis factoring extraction method and Promax rotation. These decisions were made given (a) our goal of identifying common factors, (b) our assumption that the factors would be correlated, and (c) our assumption that our total scale—or specific subscales—may not have multivariate normality. Principal axis factoring is an extraction method that seeks to identify commonalities among the factors, and a Promax rotation allows for factors to be correlated and does not assume normality (Kahn, 2006).

We also tested the factor structure using parallel analysis (O'Connor, 2000). Parallel analysis is a method for determining the number of factors that should be extracted from a dataset. The technique creates a number of random data sets with the same number of variables and observations as the original dataset. The random data sets produce a correlation matrix and eigenvalues that can then be compared to the eigenvalues from the original dataset. The number of eigenvalues from the raw dataset that are larger than the random data eigenvalues indicates the number of factors to be extracted. For this analysis, we generated 1000 random data sets with 95 percentile values and used principal axis factoring (O'Connor, 2000).

After running the EFA with these parameters, we proceeded to remove items based on the following criteria, again guided by best practice recommendations within counseling psychology (Kahn, 2006). We examined the pattern matrix and required that items loaded .50 or higher on their associated factor, that there were at least three items loading .50 or higher on every factor (Costello & Osborne, 2005), and that the loading of an item on one factor was greater than .15 apart from the loading of that item on another factor. We also evaluated items for evidence concerning content validity and specificity (Worthington & Whittaker, 2006). Table 1 depicts the factor loadings of the initial EFA prior to item removal with items bolded corresponding to the factor they load on. The initial EFA produced eight factors accounting for 64.60% of variance with parallel analysis suggesting seven factors.

Using the criteria detailed above, 18 items were removed. Twelve of these items failed to reach the loading threshold, cross loaded too closely with another factor, or were not part of a factor with three or more items. An additional six items were removed based on content validity. Specifically, the adequate compensation items were split between two factors, with five items loaded onto a second factor reflecting content about earning enough to pay bills and provide for one's family. This was in contrast to a primary factor reflecting fair pay and adequate reward for one's work. We decided that the primary factor better represented the construct

within decent work. Therefore, the six items reflecting earning enough money were removed.

After removal of these items, a 35-item EFA was conducted which purposively extracted five factors accounting for 67.18% of the variance and this factor structure was confirmed by parallel analysis. However, this 35-item solution did not represent an efficient instrument (e.g., an instrument that could be quickly taken by participants, clients, etc. while still retaining strong reliability and validity) and we again examined items for evidence concerning content validity and specificity to determine additional item removal. Following best practice recommendations, we sought to extract three items per subscale, which is the recommended minimum number of items to have for reliable and valid scales (Costello & Osborne, 2005). For the Time-Rest, Health care, and Values subscales, we chose the three highest loading items from each to represent those subscales. For the Adequate Compensation subscale, we selected the 2nd, 3rd, and 5th highest items because one item referred to salary and we felt this would not apply to many individuals who work for an hourly wage and the 5th highest loading item was the only item that was positively valenced, which we decided to retain. For the Safe Working Conditions subscale, we chose the 1st, 2nd, and 6th highest loading items as we wanted to ensure we captured both emotional and physical safety (6th highest loading item) in the workplace. We ran a final EFA with these 15 items which are depicted in Table 2 and explained 84.11% of the variance. All items loading .66 or higher on their respective factor, cross loadings between items were greater than .40 apart, and all five subscales had adequate to strong internal consistency reliability, ranging from .82 to .97. Given that the items for each subscale demonstrated strong content validity and internal consistency reliability and that the five subscales accounted for a high amount of variance in the overall decent work construct, we deemed that this final version of the scale was acceptable.

Finally, we examined the correlations among the subscales as shown in Table 3. All subscales were significantly correlated with each other with the exception of the Safe Working Conditions and Free Time and Rest subscales. Given these initial indications of a strong and cohesive scale, we sought to validate the scale's scores in Study 2.

Study 2

The goal of Study 2 was to establish validity evidence for the DWS. Eight instruments with scores previously demonstrating evidence of validity were used to assess for convergent, discriminant, and predictive evidence of validity of the DWS subscale scores. Scores from each of these five subscales—corresponding to the five components of decent work—were hypothesized to converge with scores from a specific instrument designed to assess a construct it would theoretically be related to. Specifically, we hypothesized that these relations would be strong ($> .50$), but that they wouldn't reach levels to be considered completely overlapping ($< .85$; van Mierlo et al., 2009; Voorhees, Brady, Calantone, & Ramirez, 2016). First, scores on the Safe Working Conditions subscale, which measures the degree to which the work environment is physically and emotionally safe, were correlated with scores from a measure of general job safety (Hayes, Perander, Smecko, & Trask, 1998). Given the analogous aims of each scale,

Table 1
Initial Exploratory Factor Analysis – Study 1

Variable	1	2	3	4	5	6	7	8
1. I have free time during the work week	.923	-.141	.050	-.035	.012	-.056	.056	.214
2. I do not have enough time for non-work activities	.871	.022	-.010	-.015	.063	-.118	-.052	-.253
3. I have no time to rest during the work week	.827	.061	-.057	.007	.001	-.014	.027	-.310
4. My work schedule allows me to have time to rest	.816	-.107	.020	.136	-.047	.017	.080	.065
5. I can't adequately rest because of my work	.781	.022	.037	.056	.025	-.029	-.066	-.334
6. I do not have free time given my work schedule	.774	-.009	-.033	.034	.012	-.046	.115	-.280
7. My work schedule allows me to have time for non-work activities	.769	-.098	-.047	.084	-.041	.094	.057	.086
8. Getting rest is not a problem with my job	.758	-.021	.004	-.068	-.022	.053	-.061	.232
9. Work takes up a lot of my free time	.751	-.076	.051	-.054	.016	-.001	-.149	-.144
10. I feel rested during the work week	.710	.181	-.009	-.142	-.040	.066	-.106	.305
11. I receive adequate time for breaks at work	.450	-.006	.048	.112	-.017	.096	.136	.197
12. At work, I can speak up for myself without fear of retaliation	-.090	.860	-.012	-.071	-.034	-.046	.110	-.058
13. At work, I feel safe from emotional or verbal abuse of any kind	-.036	.834	-.041	-.099	-.089	-.016	.249	-.051
14. I feel emotionally safe interacting with people at work	.005	.787	-.023	-.050	-.099	.034	.258	-.102
15. I do not feel violated in any aspect in my work	-.131	.786	.072	.031	-.003	-.044	.065	-.197
16. I do not fear the future in my work place	-.092	.743	.046	.023	.002	.094	-.104	-.190
17. My work respects my physical and mental health	.102	.697	-.006	.093	.054	-.027	-.063	.048
18. My work place is hospitable	.043	.644	-.047	.068	.023	.094	.123	-.095
19. I appreciate each and every moment I spend at work	.059	.626	.004	.175	.050	-.056	-.270	.083
20. My work satisfies all my needs	.026	.602	.029	.058	.095	-.004	-.168	.190
21. I do not feel insecure in my work place	-.044	.563	-.035	.004	-.048	-.091	.030	-.161
22. At work I am fearful of emotional abuse by customers, coworkers, or supervisors	-.171	-.459	.046	.119	.094	-.051	-.028	.429
23. I have a good healthcare plan at work	.024	-.007	1.008	-.026	.018	-.073	-.034	-.034
24. I get good healthcare benefits from my job	.015	.006	1.001	-.026	-.010	-.067	-.031	.008
25. My job provides adequate access to healthcare	-.038	.011	.957	-.021	-.028	-.012	.066	-.029
26. My employer provides acceptable options for healthcare	-.037	-.002	.956	-.007	-.039	.021	.040	-.041
27. My family and I have adequate healthcare coverage through my job	.027	-.008	.937	-.019	.004	-.037	-.018	.030
28. If I get injured at work, I am confident that I will be taken care of	.139	-.021	.593	.004	.015	.111	.176	.058
29. My healthcare coverage requires minimal out-of-pocket expenses	-.028	-.030	.558	.043	-.011	.129	-.005	.067
30. I wish that my health benefits were better	-.055	.049	.296	.160	.139	.068	-.154	-.278
31. The values of my organization match my family values	.007	-.054	.012	.997	-.040	-.014	.006	.045
32. My organization's values align with my family values	.029	-.050	.006	.969	.000	-.029	.038	.027
33. The values of my organization match the values within my community	.015	-.016	-.034	.941	-.050	-.008	.017	.080
34. My organization's values align with the values within my community	.012	-.024	-.040	.894	.028	.007	.034	.022
35. I would feel comfortable recommending a family member to work in my organization	-.027	.318	.025	.490	.059	.054	.019	.053
36. I would feel comfortable recommending a community member to work in my organization	.007	.307	-.005	.467	.074	.030	-.011	.031
37. I should receive greater compensation for my work	.045	-.075	-.027	.019	.965	-.150	-.066	-.152
38. I do not feel I am paid enough based on my qualifications and experience	.004	-.121	-.045	.016	.921	-.042	.061	.041
39. I am not properly paid for my work	-.061	-.026	.026	-.019	.828	.107	.033	-.157
40. I think my salary is fair in relation to what I do at work	.002	.064	-.037	-.071	.812	.071	.083	.204
41. I am rewarded adequately for my work	-.023	.040	.071	-.012	.666	.141	.043	.148
42. I make enough money at my job to pay my bills	-.055	-.081	.038	.045	-.037	.904	.128	-.076
43. I make enough money to provide for myself and others (e.g., family members, pets)	.052	-.003	-.049	-.061	.038	.850	.049	.114
44. I make enough money at my job to save some every month	-.064	-.052	-.010	.061	-.088	.838	-.019	.032
45. The amount of money that I make at my job isn't enough to cover my monthly expenses	.016	-.081	.010	.055	.074	.794	-.002	-.523
46. I am often stressed about money given what I make at my job	.076	.076	-.051	-.123	.128	.650	-.135	-.265
47. I feel I can live the kind of life that I would like with the money I earn at my job	-.003	.053	.117	-.043	-.044	.614	-.121	.249
48. When I am at work I do not feel the time passing by	.009	.105	-.002	.064	-.111	.203	-.157	.169

Table 1 (continued)

Variable	1	2	3	4	5	6	7	8
49. At work, I feel safe from physical abuse	−.064	.267	−.014	.030	.024	−.026	.679	−.113
50. At work, I have access to basic necessities, like food, water, and shelter	.007	.150	.080	.039	.010	.044	.539	−.088
51. I feel physically safe interacting with people at work	.013	.357	.022	.033	−.006	−.013	.533	−.180
52. The tasks I perform at work do not put me at risk for physical injury or harm	.061	.286	.046	.000	.121	−.070	.388	.009
53. At work I am fearful of physical abuse by customers, coworkers, or supervisors	−.081	−.299	.014	.102	−.028	.012	−.146	.649

Note. Bolded corresponding to the factor they load on.

we hypothesized that scores would be strongly, positively correlated (*Hypothesis 1*).

Second, scores on the Access to Health care subscale, which measures the degree to which health care is accessible through one's job, were correlated with scores from a health care satisfaction measure (Hays, Davies, & Ware, 1987). Work continues to be the main source of health care access for U.S adults (Gallup, 2016) and theoretically perceiving access to health care through one's job would be related to, but distinct from, an overall sense of satisfaction with one's health care access and availability (Duffy et al., 2016). As such, we hypothesized that scores on each of these constructs would be strongly, positively correlated (*Hypothesis 2*). Third, scores on the Adequate Compensation subscale, which measures the perception that one's pay at work is acceptable, were correlated with scores from a measure of pay satisfaction (Heneman & Schwab, 1985). Individuals who feel their pay reaches a threshold to be considered acceptable are likely to endorse greater satisfaction with their pay, and as such we hypothesized that scores from these two instruments will be strongly, positively correlated (*Hypothesis 3*).

Fourth, scores on the Free Time and Rest subscale, which assesses the degree to which individuals are able to have time and

rest outside of work, were correlated with scores on a measure of occupational fatigue (Winwood, Lushington, & Winefield, 2006; Winwood, Winefield, Dawson, & Lushington, 2005). This measure assesses the degree to which individuals feel overworked, exhausted, and fatigued. Theoretically, occupational fatigue would be strongly related to the perception that one has time and rest outside of work (Winwood et al., 2005). Accordingly, we hypothesized that scores from these two instruments will be strongly, negatively correlated (*Hypothesis 4*). Fifth, the Complementary Values subscale assesses the degree to which the work values of one's organization aligns with that person's family and community values. Scores from this subscale were correlated with scores from a measure of workplace fit perceptions (Cable & DeRue, 2002), which assessed the degree to which a job aligns with an individual's personal values and needs. Theoretically, alignment of overall personal values and needs at work would correspond with meeting family and community values at work (Edwards & Rothbard, 1999). As such, we hypothesized that scores from these two instruments will be strongly, positively correlated (*Hypothesis 5*).

In addition to convergent evidence of validity, we also sought to demonstrate discriminant and predictive evidence of validity. For

Table 2
Final Item Loadings From Exploratory Factor Analysis – Study 1

Factor name and items	Loadings				
	1	2	3	4	5
Factor 1: Physically and Interpersonally Safe Working Conditions $\alpha = .82$					
1. I feel emotionally safe interacting with people at work	.87				
2. At work, I feel safe from emotional or verbal abuse of any kind	.80				
3. I feel physically safe interacting with people at work.	.66				
Factor 2: Access to Health Care $\alpha = .97$					
4. I get good healthcare benefits from my job.		.98			
5. I have a good healthcare plan at work.		.98			
6. My employer provides acceptable options for healthcare.		.90			
Factor 3: Adequate Compensation $\alpha = .85$					
7. I am not properly paid for my work. (r)			.88		
8. I do not feel I am paid enough based on my qualifications and experience. (r)			.85		
9. I am rewarded adequately for my work			.69		
Factor 4: Hours that Allow for Free Time and Rest $\alpha = .87$					
10. I do not have enough time for non-work activities. (r)				.92	
11. I have no time to rest during the work week. (r)				.82	
12. I have free time during the work week				.76	
Factor 5: Organizational Values Complement Family and Social Values $\alpha = .96$					
13. The values of my organization match my family values.					1.00
14. My organization's values align with my family values.					.95
15. The values of my organization match the values within my community.					.88

Table 3
Correlations of Decent Work Scale Subscales – Study 1
($N = 275$)

Variable	1	2	3	4	5
1. Safe working conditions	—				
2. Access to healthcare	.19	—			
3. Adequate compensation	.20	.31	—		
4. Free time and rest	.32	.02*	.25	—	
5. Complementary values	.41	.30	.44	.33	—
<i>M</i>	17.97	12.99	11.36	14.64	14.51
<i>SD</i>	3.33	6.61	4.87	4.69	4.54

* Nonsignificant at the $p < .01$ level.

discriminant validity, we hypothesized (*Hypothesis 6*) that scores on each of the subscales would evidence stronger correlations to their respective convergent instrument scores (e.g., Complementary Values and workplace fit) versus correlations with instruments hypothesized to converge with other subscales (e.g., Complementary Values and job safety). We also examined how well each of the subscales correlated with length of tenure at the participant's specific job, hypothesizing that tenure would not be significantly correlated with any of the decent work subscales (*Hypothesis 7*). We contend that experiencing decent work is unrelated to the amount of time someone is in a particular job but rather the specific experiences within that workplace.

For predictive evidence of validity, we were interested in whether the five subscale scores would significantly predict suggested outcome variables of decent work within the PWT—job satisfaction, work meaning, and withdrawal intentions. Theoretically, higher levels of decent work would predict greater satisfaction and meaning in one's job as well as fewer intentions to withdraw (Duffy et al., 2016). We conducted stepwise regression analyses to assess the degree to which each component of decent work significantly predicted these outcomes while accounting for the other components. These analyses were exploratory as no research has examined how the five components of decent work jointly predict outcomes. Based on the results of Study 2, we present a final DWS and associated subscales and provide research and practical recommendations for using the new instrument.

Method

Participants. The sample for this study contained 589 adults with a mean age of 35.29 years ($SD = 11.72$). Participants self-identified as female ($n = 322$, 54.8%), male ($n = 261$, 44.3%), transgender ($n = 2$, 0.3%), other ($n = 3$, 0.5%), and one (0.2%) chose not to answer. Considering the U.S. Census Bureau (2015) percentages, our sample includes more women than men compared to the national average. Regarding ethnicity, of those who responded, the sample self-identified as White/European American ($n = 493$, 83.7%), Black/African American ($n = 30$, 5.1%), Hispanic/Latina/o-American ($n = 37$, 6.3%), Asian/Asian American ($n = 34$, 5.8%), Native American ($n = 13$, 2.2%), Asian Indian ($n = 4$, 0.7%), Arab American ($n = 3$, 0.5%), Pacific Islander ($n = 4$, 0.7%), and other ($n = 4$, 0.7%). Note several participants selected multiple ethnicity categories. Compared to the U.S. Census Bureau, our sample has slightly more white

participants, less black/African American participants, but is comparable in other minority groups.

The sample was largely employed full-time ($n = 438$, 74.4%) with the remaining participants employed part-time ($n = 101$, 17.1%), self-employed full-time ($n = 24$, 4.1%), self-employed part-time ($n = 25$, 4.2%), and one (0.17%) who chose not to answer. Education levels consisted of less than high school ($n = 1$, 0.2%), some high school ($n = 4$, 0.7%), high school graduate ($n = 55$, 9.3%), trade/vocational school ($n = 29$, 4.9%), some college ($n = 153$, 26.0%), undergraduate degree ($n = 251$, 42.6%), and professional degree ($n = 96$, 16.3%). Current participant social class levels were as follows: lower class ($n = 43$, 7.3%), working class ($n = 213$, 36.2%), middle class ($n = 270$, 45.8%), upper-middle class ($n = 56$, 9.5%), and upper class ($n = 4$, 0.7%). The annual personal income of participants was as follows (in thousands of US\$): less than \$25 ($n = 102$, 17.3%), \$26–\$50 ($n = 184$, 31.2%), \$51–\$75 ($n = 132$, 22.4%), \$76–\$100 ($n = 89$, 15.1%), \$101–\$125 ($n = 39$, 6.6%), \$126–\$150 ($n = 19$, 3.2%), \$151–\$175 ($n = 8$, 1.4%), \$176–\$200 ($n = 3$, 0.5%), \$201+ ($n = 12$, 2.0%), and one (0.2%) chose not to answer.

Instruments.

Decent work. The 15-item Decent Work Scale (DWS) was administered. The estimated internal consistency for the five, three-item subscales and total scale was as follows: Safe Working Conditions ($\alpha = .79$), Access to Health care ($\alpha = .97$), Adequate Compensation ($\alpha = .87$), Free Time and Rest ($\alpha = .87$), Complementary Values ($\alpha = .95$), and Total Scale ($\alpha = .86$).

Health care satisfaction. The Patient Satisfaction Questionnaire III (PSQ-III; Hays et al., 1987) is a revised version developed from the PSQ-I (Ware, Snyder, & Wright, 1976a, 1976b) and the PSQ-II (Ware, Snyder, Wright, & Davies, 1983) and was used to assess individual's satisfaction with their health care access and availability. Specifically, the 12-item Access/Availability/Convenience subscale of the PSQ-III was used to measure the degree to which participants felt satisfied with access to their medical care. This subscale was used instead of the overall scale because its content captures accessibility to medical care while the overall PSQ-III assesses global satisfaction. The Access/Availability/Convenience subscale we administered used a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Sample items from this subscale include, "The office hours when I can get medical care are convenient for me," and "I am able to get medical care whenever I need it." Hays et al. (1987) found a Cronbach's alpha of .86 for this subscale. In the development of the PSQ-III, content and discriminant evidence of validity were improved through the revision and addition of items. In the present study, the estimated internal consistency reliability of scores was .88.

Pay satisfaction. Participants' satisfaction with their compensation at work was measured using the 18-item Pay Satisfaction Questionnaire (PSQ; Heneman & Schwab, 1985), which has four dimensions: Level (four items), Benefits (four items), Raise (four items), and Structure/Administration (six items). The overall PSQ was used, as we were interested in measuring participants' overall satisfaction with their pay to compare to our general measurement of compensation, and we were not as interested in specific dimensions of pay. Additionally, research support has been mixed regarding the dimensionality of the PSQ, with evidence supporting one dimension, three and four dimensions, and five dimensions, respectively (Mulvey, Miceli, & Near, 1990). In the present study,

participants were instructed to respond based on how satisfied or dissatisfied they felt about their pay, and each response used a 5-point Likert scale ranging from 1 (*very dissatisfied*) to 5 (*very satisfied*). Examples include, “My take-home pay,” “The company’s pay structure,” and “How my raises are determined.” Among white-collar employees and nurses, Heneman and Schwab (1985) found overall internal consistency reliability estimates within each dimension exceeding .80. Adequate validity has been shown through the evaluation of discriminant evidence of validity (Judge, 1993) and confirmatory analyses regarding the factor dimensionality of the scores (Judge & Welbourne, 1994). For the present study, the estimated internal consistency reliability of total scores was .96, and the estimates of the subscales were .96 (Level), .96 (Benefits), .89 (Raise), and .92 (Structure/Administration).

Job safety. The 50-item Work Safety Scale (WSS; Hayes et al., 1998), which assesses participants’ perceptions of different areas of work safety, measures five 10-item constructs: job safety, coworker’s safety, supervisor safety, management safety practices, and safety program evaluation. Only the job safety construct was used in the present study, as we were interested in comparing this scale to our measure of Safe Working Conditions instead of looking each specific area in which safety climate is enforced. Each question used a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) with higher scores associated with higher safety. Sample items from the job safety construct area include “Dangerous,” “Fear for Health,” and “Could Get Hurt Easily.” In the scale development article, the reliability estimates for the constructs were all greater than .87, and the constructs were all significantly associated with compliance with safety behaviors (CSB) scores, quantity of safety trainings, and participation in safety orientations (Hayes et al., 1998). The constructs also correlated expectedly with numerous health measures (i.e., General Health Questionnaire, Physical Complaints Questionnaire, and Sleep Complaints Questionnaire; Hayes et al., 1998). The estimated internal consistency reliability for the job safety construct was .95.

Occupational fatigue. The 15-item Occupational Fatigue Exhaustion Recovery (OFER 15) Scale (Winwood et al., 2005, 2006) was used to assess participants’ experience of fatigue and strain at work and home over the last few months prior to the administration of the survey. The OFER 15 Scale has three subscales, including Chronic Fatigue (sample item: “I often feel I’m ‘at the end of my rope’ with my work”), Acute Fatigue (sample item: “After a typical work period I have little energy left”), and Recovery/Persistent Fatigue (sample item: “I never have enough time between work periods to recover my energy completely”). A total OFER 15 score was used, as our subscale measured free time and rest overall in participants and we deemed it most beneficial to capture all components of fatigue. Each response uses a 7-point Likert scale ranging from 0 (*strongly disagree*) to 6 (*strongly agree*). Winwood et al. (2006) found high reliability estimates within the three subscales, with all estimates greater than .84, and found evidence for face, construct, and discriminant evidence of validity. For the present study, the estimated internal consistency reliability of total scores was .95. The estimates for each of the subscales are as follows: .93 (Chronic Fatigue), .92 (Acute Fatigue), and .89 (Recovery/Persistent Fatigue).

Fit. Cable and DeRue’s (2002) nine-item scale measuring participants’ perceived fit with their current job includes three,

three-item subscales: Person-Organization Fit (P-O Fit), Needs-Supplies Fit (N-S Fit), and Demands-Abilities (D-A Fit). The P-O Fit and N-S Fit subscales were used in the present study because the D-A Fit subscale would not have supported the validity of our subscale’s scores, as our subscale was more values-based instead of abilities-based. Ranging from 1 (*strongly disagree*) to 7 (*strongly agree*), participants answered these items using a 7-point Likert scale. An sample item from the P-O Fit subscale is “The things I value in life are very similar to the things that my organization values,” and a sample item from the N-S Fit subscale is “There is a good fit between what my job offers me and what I am looking for in a job.” For each subscale, Cable and DeRue (2002) found good reliability estimates ranging from .84 to .92 and found the subscales to correlate in expected directions with measures of organizational support as well as occupational commitment. The overall estimated internal consistency reliability for the present study for the total scale was .95, and the estimates for the subscales are as follows: .97 (P-O Fit) and .95 (N-S Fit).

Job satisfaction. The five-item Job Satisfaction Scale (JSS; Judge et al., 1998) was used to measure how satisfied participants were with their current jobs. Items were answered using a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Sample items include, “Most days I am enthusiastic about my work” and “I find real enjoyment in my work.” Strong internal consistency reliability has been reported, and responses have been found to correlate with well-established measures of job satisfaction (Judge et al., 1998). Responses have also correlated in expected directions with career commitment and work meaning (Duffy, Allan, Autin, & Bott, 2013; Duffy, Allan, Autin, & Douglass, 2014; Duffy, Bott, Allan, Torrey, & Dik, 2012). For the present study, the estimated internal consistency reliability of total scores was .92.

Work meaning. Participant’s level of meaning their feel in their work was assessed with the Work and Meaning Inventory (WAMI; Steger, Dik, & Duffy, 2012), which is a 10-item instrument using a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The scale contains three factors (positive meaning, meaning making through work, and greater good motivations) which Steger et al. (2012) found all loaded onto a higher order meaningful work factor. A sample item is, “I understand how my work contributes to my life’s meaning.” Higher scores represent higher levels of meaningful work. Steger et al. (2012) found scores from the total scale to have strong internal consistency reliability ($\alpha = .93$) and correlate with career commitment, life meaning and job satisfaction. In the current study, the estimated internal consistency of total scale scores was $\alpha = .95$.

Withdrawal intentions. A three-item instrument developed by Blau (1985) was used as a measure of intention to withdraw from one’s occupation. These items were: “I am thinking about leaving my current occupation,” “I am actively searching for an alternative to my occupation,” and “I intend to stay in my current occupation for some time.” Across several studies, scores from this instrument have evidenced acceptable internal consistency reliability ($\alpha = .70 - .87$) as well as test-retest reliability (.60; Blau, 1985, 1989, 2000) and scores on the instrument have correlated with aspects of career or work commitment. The estimated internal consistency of scales scores in the current study was .89.

Procedure. First, IRB approval was received was the first author's institution to conduct this study. Next, participants were recruited through the same methodology as in Study 1. A survey link was posted to MTurk, and participants were paid \$.40 for completing the survey. Again, any duplicate responses were searched for and removed from the dataset. An initial sample of 703 participants began the survey. Six validity check items were inserted into these surveys, and the seven participants who failed to answer these correctly were removed. Next, we removed seven participants who were unemployed. Finally, given our large sample size and focus on scale development, we removed 51 participants who did not have complete data on the DWS. This resulted in sample of 642 employed adults with full data on the DWS. To compare scores on the DWS to the eight other instruments, we examined the amount of missing data for each of these eight scales. We did not include participants who were completely missing scale data for one or more of the eight scales (i.e., participants who left an entire scale blank). This resulted in the removal of 42 participants, leaving 600 participants in the sample. Within our sample of 600 participants, 101 participants were missing data from at least one of the eight instruments, ranging from missing one to three total items. Missing values represented less than 1% of the total dataset. Following recommendations by Parent (2013), we used mean substitution to impute missing data for these cases where appropriate, which Parent suggests is acceptable when the level of missing data is low (e.g., less than 10%), the sample size is strong, and the scales are reliable. All of these conditions were met in the current study. Finally, as discussed below, 11 outliers were removed from the dataset for a final total of 589 participants.

Results

Preliminary analysis. All variables were inspected for outliers and normality. Complementary Values, Free Time and Rest, and Adequate Compensation all had absolute values of skewness and kurtosis under 1 and included no extreme scores. Access to Health care had skewness under one but also had some negative skew (-1.21). Safe Working Conditions was negatively skewed (-1.53) and kurtotic (2.17) but also had several extreme scores over 3.5 standard deviations from the mean. Eleven outliers were removed, which improved non-normality of this subscale (skew = -1.29 and kurtosis = 1.04) of this factor.

Model testing. To evaluate the models, we used structural equation modeling (SEM) in Mplus 7.4 with robust maximum likelihood estimation, which effectively addresses issues of non-normality (Muthén & Muthén, 2012). Our sample size of 589 was adequate to test our SEM hypotheses with a sample of at least 200 and 5–10 cases per parameter (Kline, 2016). Indices of fit that minimized the likelihood of Type I and Type II error were selected (Hu & Bentler, 1999). These were the chi-square test (χ^2), the comparative fit index (CFI), the Tucker-Lewis Index (TLI), the root mean square error of approximation (RMSEA), and the standardized root-mean-residual (SRMR). A significant χ^2 can indicate a poor fitting model, but this test is not reliable in larger samples (Tabachnick & Fidell, 2013). Criteria for the CFI, TLI, and RMSEA have ranged from less conservative (CFI $\geq .90$; TLI $\geq .90$; RMSEA $\leq .10$; SRMR $\leq .10$) to more conservative (CFI $\geq .95$; TLI $\geq .95$; RMSEA $\leq .08$; SRMR $\leq .06$; Hu & Bentler, 1999; Quintana & Maxwell, 1999; Weston & Gore, 2006).

However, researchers should be cautious when using these criteria as strict cut-offs and should consider sample size and model complexity when judging the fit of models (Weston & Gore, 2006).

To assess the factor structure of the DWS, we examined three separate models: a correlational model, a higher order model, and a bifactor model. The correlational model consisted of the five factors that comprised decent work, which were then allowed to correlate with one another. Best fit here would suggest that the sub factors are largely independent. In the higher order model, each of the five latent factors are regressed onto a higher-order decent work factor, rather than correlated. Best fit here would suggest that each of the decent work sub factors represent facets of a higher order construct. In the bifactor model, items for each scale are loaded onto their respective subscale and a general decent work factor. Interpretation for a best fitting bifactor is more complex, suggesting that each subscale has independent variance and variance that belongs to an underlying decent work factor. If a bifactor model represents the best fit to the data, researchers can calculate bifactor indices to determine, among other things, the proportion of variance accounted for by the subscales and the general factor respectively (Rodríguez, Reise, & Haviland, 2016).

Correlational model. The correlational model had each indicator load on its respective factor and allowed each factor to correlate with one another. This model had good fit to the data, $\chi^2(80) = 241.41$, $p < .001$, CFI = .97, TLI = .96, RMSEA = .06, 90% CI [.05, .07], and SRMR = .05, and all indicators loaded on their factors at values of .61 or above.

Higher order model. The higher order model regressed the five factors onto a higher order decent work factor. The model had a slight decline in fit, $\chi^2(85) = 259.76$, $p < .001$, CFI = .96, TLI = .96, RMSEA = .06, 90% CI [.05, .07], and SRMR = .07, and this change was significant, $\Delta\chi^2(5) = 12.03$, $p < .001$. However, the CFI change was not greater than 0.01 (Δ CFI = .003), indicating that the models were not practically different (Cheung & Rensvold, 2002).

Bifactor model. The bifactor model regressed each indicator on its respective factor while simultaneously regressing each indicator onto a decent work factor. This model had a better fit than the correlational model, $\chi^2(75) = 151.31$, $p < .001$, CFI = .98, TLI = .98, RMSEA = .04, 90% CI [.03, .05], and SRMR = .05, and this change was significant, $\Delta\chi^2(5) = 40.07$, $p < .001$ and Δ CFI = .02. Therefore, we retained this model as the best fit. Figure 1 depicts this model. Given that the bifactor model had the best fit, we proceeded to calculate bifactor statistical indices, including omega, omega hierarchical, and explained common variance (Rodríguez et al., 2016). Omega is a reliability estimate for factor analysis that represents the proportion of variance in the decent work total score attributable to common variance, rather than error. The omega for the total score, which included the general factors and the subscale factors, was .84, meaning that 84% of the variance in the total score is due to the factors and 16% is attributable to error. The omegas for the subscales were Safe Working Conditions ($\omega = .64$), Access to Health care ($\omega = .74$), Adequate Compensation ($\omega = .71$), Free Time and Rest ($\omega = .72$), and Complementary Values ($\omega = .73$).

Omega hierarchical estimates are the proportion of the variance in the decent work total score attributable to the general decent work factor, which in effect treats the decent work subscales as

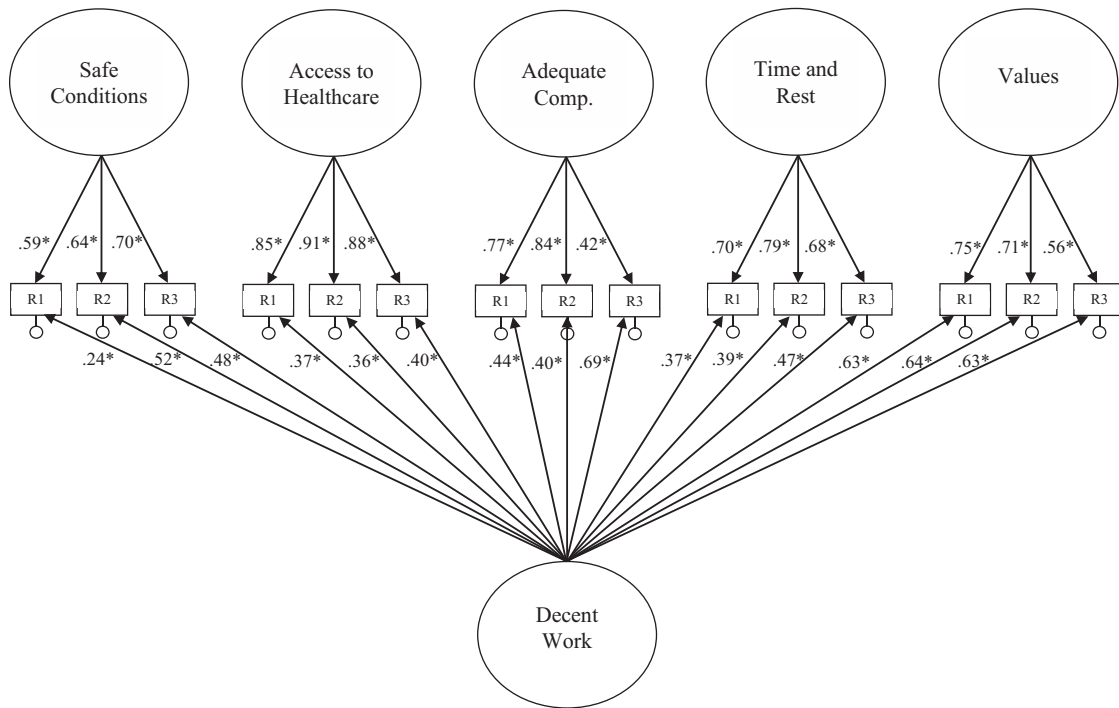


Figure 1. Final confirmatory bifactor model of decent work subscales. * $p < .01$. Fit is $\chi^2(75) = 151.31$, $p < .001$, Comparative Fit Index (CFI) = .98; Tucker-Lewis Index (TLI) = .98; Root Mean Square Error of Approximation (RMSEA) = .04; 90% CI [.03, .05]; Standardized Root-Mean-Residual (SRMR) = .05.

error. OmegaH for the decent work total was .57. Comparing omega (.84) and omegaH (.57), we can see that 68% of the reliable variance in the decent work total score ($.57/.84 = .68$) is attributable to the general factor, and 23% ($.27/.84 = .32$) is attributable to the subscale factors. The omega hierarchical for the subscales were Safe Working Conditions ($\omega_H = .19$), Access to Health care ($\omega_H = .32$), Adequate Compensation ($\omega_H = .20$), Free Time and Rest ($\omega_H = .23$), and Complementary Values ($\omega_H = .19$). Not surprisingly, after the common variance associated with the general factor is partialled out, reliability estimates decrease for the sub factors. Finally, explained common variance (ECV) describes how much of the common variance is attributable to the factor versus the five factors. In this case, the ECV was .31, suggesting that 31% of the common variance is attributable to the general factor whereas 69% is spread among the five decent work subscales.

Invariance models. We conducted invariance tests on the bifactor model, because it had the best fit to the data. To perform invariance tests for race, gender, income, and subjective social class, we first created two categories per group. For race, we compared white versus nonwhite participants, and for gender, we compared men versus women. Although grouping participants in white and nonwhite categories is potentially problematic, our sample did not have large enough subgroups to make meaningful comparisons. Therefore, we created these subgroups, in part, to evaluate the need for further, more specific invariance testing. For income and subjective social class, we split the variables at the mean to create two categorical groups each. For income, we created the two groups at greater than and less than \$50,000 in average yearly household income. For social class, we split the

groups at greater than and less than 4.94, which is the average response on the MacArthur Scale of Subjective Social Status. This scale presents participants with a picture of a ladder, and participants indicate where they fall on a scale from 1 (*worst off*) to 10 (*best off*), relative to other people in the United States (Adler, Epel, Castellazzo, & Ickovics, 2000). These analyses are exploratory in nature and designed to stimulate further research on the role of race, income, and subjective social class, if these findings indicate the utility of follow-up research.

We then successively constrained parameters across groups in three models: the configural, metric, and scalar models. In the configural model, only the model structure is held constant across groups. Poor fit here suggests that the organization of indicators is different for both groups. Next, in metric invariance, the configuration of variables and all factor loadings are constrained to be the same for each group. A reduction in fit between the metric and configural models suggests that the factor loadings vary in size between the two groups. Finally, in scalar invariance, the configuration, factor loadings, and the indicator intercepts are constrained to be the same for each group. A reduction in fit between the scalar and the metric models suggests that indicators have different intercepts in both groups (Little, 2000). Reduction in fit between the models can be tested with a chi-square test, which suggests that the models differ when significant. However, because changes in chi-square are sensitive to sample size, researchers can observe the change in CFI to examine if the change in fit is practically significant (Cheung & Rensvold, 2002). Specifically, experts assert that a change in CFI greater than 0.01 is substantial (Cheung & Rensvold, 2002). In addition to reduction in fit, the

absolute value of fit indices must be acceptable to test invariance (Little, 2000).

Race. The configural model for race had a good fit, $\chi^2(120) = 235.78$, $p < .001$, CFI = .98, TLI = .98, RMSEA = .04, 90% CI [.03, .06], and SRMR = .05. Fit was similar for the metric model, $\chi^2(174) = 260.61$, $p < .001$, CFI = .98, TLI = .98, RMSEA = .04, 90% CI [.03, .05], and SRMR = .06, and the configural and metric models did not significantly differ, $\Delta\chi^2(24) = 25.59$, $p = .37$, $\Delta\text{CFI} < .01$. Fit was also similar for the scalar model, $\chi^2(183) = 268.06$, $p < .001$, CFI = .98, TLI = .98, RMSEA = .04, 90% CI [.03, .05], and SRMR = .06, and the scalar model was not significantly different from the metric model, $\Delta\chi^2(9) = 6.40$, $p = .70$, $\Delta\text{CFI} < .01$. Therefore, factor structure and indicator intercepts were maintained across majority and minority racial groups.

Gender. The configural model for gender had a good fit, $\chi^2(150) = 260.64$, $p < .001$, CFI = .98, TLI = .97, RMSEA = .05, 90% CI [.04, .06], and SRMR = .05. The fit slightly declined for the metric model, $\chi^2(174) = 326.95$, $p < .001$, CFI = .97, RMSEA = .06, 90% CI [.05, .06], and SRMR = .08. Although the models had a significantly different chi-square, $\Delta\chi^2(24) = 65.44$, $p < .001$, the change in CFI was not substantial, $\Delta\text{CFI} < .01$. Finally, the scalar model was similar to the metric model, $\chi^2(183) = 340.12$, $p < .001$, CFI = .97, RMSEA = .06, 90% CI [.05, .06], and SRMR = .06, and did not significantly differ in fit, $\Delta\chi^2(9) = 12.37$, $p = .19$, $\Delta\text{CFI} < .01$. Therefore, factor structure and indicator intercepts were maintained across the two gender groups.

Income. The configural model for income had a good fit, $\chi^2(150) = 214.32$, $p < .001$, CFI = .99, TLI = .98, RMSEA = .04, 90% CI [.03, .05], and SRMR = .05. The fit was similar for the metric model, $\chi^2(174) = 238.78$, $p < .001$, CFI = .99, TLI = .98, RMSEA = .04, 90% CI [.02, .05], and SRMR = .06, and the fit did not significantly differ between models, $\Delta\chi^2(24) = 25.64$, $p = .37$, $\Delta\text{CFI} = .00$. Fit was also similar for the scalar model, $\chi^2(183) = 251.45$, $p < .001$, CFI = .99, TLI = .98, RMSEA = .04, 90% CI [.02, .05], and SRMR = .06, and the scalar model was

not significantly different from the metric model, $\Delta\chi^2(9) = 12.75$, $p = .17$, $\Delta\text{CFI} = .001$. Therefore, the factor structure and indicator intercepts were maintained across the two income groups.

Social class. The configural model for income had a good fit, $\chi^2(150) = 231.50$, $p < .001$, CFI = .98, TLI = .98, RMSEA = .04, 90% CI [.03, .05], and SRMR = .05. The fit was similar for the metric model, $\chi^2(174) = 243.23$, $p < .001$, CFI = .99, TLI = .98, RMSEA = .04, 90% CI [.03, .05], and SRMR = .05, and fit did not significantly differ between the models, $\Delta\chi^2(24) = 18.37$, $p = .24$, $\Delta\text{CFI} < .01$. However, fit did decline slightly for the scalar model, $\chi^2(183) = 260.99$, $p < .001$, CFI = .98, TLI = .98, RMSEA = .04, 90% CI [.03, .05], and SRMR = .06. Although the metric and scalar models had a significantly different chi-square, $\Delta\chi^2(9) = 18.70$, $p < .001$, the change in CFI was not substantial, $\Delta\text{CFI} < .01$. Therefore, the factor structure and indicator intercepts were maintained across the two social class groups.

Convergent and discriminant validity. We examined the correlations between the total DWS score and the five subscale scores with six similar instrument scores to demonstrate convergent evidence of validity (Table 4). As hypothesized, the five subscales significantly correlated with their specific convergent validity scales, but not to levels where these would be considered completely overlapping constructs: Safe Working Conditions and Job Safety ($r = .40$), Access to Healthcare and Healthcare Satisfaction ($r = .35$), Adequate Compensation and Pay Satisfaction ($r = .70$), Free Time and Rest and Occupational Fatigue ($r = -.60$), and Complimentary Values and Fit ($r = .76$). As hypothesized, three subscales correlated with their hypothesized convergent scales stronger than the other validity scales included. Specifically, the correlation values of these three subscales with their hypothesized convergent scales were compared with the correlation values of the three subscales with the strongest nonhypothesized convergent validity scale. We used Fisher's transformations to calculate a z-score representing the difference between the two correlations which can then be examined for significance ($p < .05$). These values significantly differed for the validity scale comparisons for Adequate Compensation (.70 vs. .46, $p < .05$),

Table 4

Correlations of Decent Work Scale Subscales, Total Scale, and Validity Scales – Study 2 ($N = 589$)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Total scale	—														
2. Safe working conditions	.56	—													
3. Access to healthcare	.64	.14	—												
4. Adequate compensation	.70	.27	.31	—											
5. Free time and rest	.61	.31	.08*	.29	—										
6. Complimentary values	.69	.37	.25	.34	.36	—									
7. Job safety	.27	.40	.08	.12	.21	.15	—								
8. Patient satisfaction	.49	.37	.35	.26	.31	.30	.23	—							
9. Pay satisfaction	.72	.29	.55	.70	.27	.41	.10	.37	—						
10. Occupational fatigue	-.62	-.41	-.20	-.41	-.60	-.47	-.21	-.42	-.46	—					
11. Fit	.67	.35	.28	.46	.36	.76	.11	.31	.62	-.55	—				
12. Job satisfaction	.58	.34	.21	.42	.38	.56	.08*	.31	.54	-.68	.75	—			
13. Work meaning	.48	.20	.24	.32	.26	.53	-.02*	.29	.49	-.49	.72	.80	—		
14. Withdrawal intentions	-.51	-.30	-.24	-.45	-.28	-.41	-.11	-.29	-.50	.48	-.59	-.69	-.56	—	
15. Tenure (years)	.07*	.03*	.10	.02*	.07*	-.01*	-.03*	.14	.08	-.09	.07*	.17	.17	-.15	—
<i>M</i>	71.89	17.87	13.07	11.71	14.17	14.81	43.45	56.54	53.56	42.46	27.51	22.64	46.13	8.08	5.11
<i>SD</i>	15.54	3.63	6.45	5.02	4.73	4.46	8.53	13.02	17.35	22.16	9.47	8.11	15.17	3.90	5.75

* $p > .05$.

Free Time and Rest (.60 vs. .36, $p < .05$), and Complimentary Values (.76 vs. .47, $p < .05$). However, disconfirming hypotheses, Access to Health care correlated more strongly with pay satisfaction than health care satisfaction (.55 vs. .35, $p < .05$) and Safe Working Conditions was equally correlated with job safety and occupational fatigue (.40 vs. .41, $p = .84$). We address these findings in detail in the discussion.

Additionally, the correlations of the subscales with tenure were either nonsignificant or weak (Access to Health care and tenure, $r = .10$). Finally, all subscales significantly correlated with each other except for Free Time and Rest and Access to Health care (r s .14–.37), each subscale significantly correlated with job satisfaction, work meaning, and withdrawal intentions (r s .20–.56), and the total scale strongly correlated with job satisfaction ($r = .56$), work meaning ($r = .48$), and withdrawal intention ($r = -.51$).

Predictive evidence of validity. Three hierarchical regressions were conducted to examine how well the five subscales of the DWS predicted three outcome variables: job satisfaction, work meaning, and withdrawal intentions. As seen in Table 5, the decent work subscales accounted for 41% of the variance in job satisfaction, 31% of the variance in work meaning, and 29% of the variance in withdrawal intentions. Additionally, multiple components of decent work were significant ($p < .05$) predictors of job satisfaction (safe working conditions, adequate compensation, free time and rest, complementary values), work meaning (access to health care, adequate compensation, complementary values), and withdrawal intentions (safe working conditions, adequate compensation, complementary values) while accounting for the other components of decent work. All components of decent work were found to be predictive of at least one outcome variable that was conceptually related to the content of the given subscale of the DWS.

Discussion

Building from the Psychology of Working Theory (Duffy et al., 2016), the present study developed a multidimensional, efficient measure of decent work demonstrating reliable and valid scores. Based on expert recommendations (Heppner et al., 2015; Kahn, 2006), we conducted two studies. Study 1 consisted of item development, expert review of these items, and the administration of these items to a diverse sample of working adults. After conducting an EFA in Study 1, we gathered additional data from a new sample of working adults in Study 2 to confirm the scale's factor

structure using CFA. Additionally, we tested invariance models according to race, gender, income, and subjective social class and examined the validity of the scores of the instrument as compared with similar instruments.

Exploratory Factor Analysis

Following recommended practices within counseling psychology (Heppner et al., 2015; Kahn, 2006), we conducted an EFA to narrow down an initial item pool to demonstrate a five-factor structure corresponding to the five components of decent work. After the final iteration, the scale was reduced to 15 items that consisted of three items per factor/subscale. All items on the final scale evidenced high factor loadings on their associated factors, minimal cross-loading on additional factors, and strong internal consistency reliability. Taken together, the EFA conducted in Study 1 suggested that the DWS consisted of five factors that represented distinct subscales. The results offered preliminary support for a multidimensional measure of decent work with reliable responses.

Confirmatory Factor Analysis and Scoring Instructions

To confirm the factor structure of the DWS obtained from Study 1, we conducted a CFA. We examined three separate models in Study 2: a correlational model, a higher order model, and a bifactor model. The bifactor model was the best fit to the data, suggesting that the DWS items share a common, underlying factor while also loading onto their own subfactors. The bifactor indices enabled us to assess the relative contribution of the general versus subscales to the DWS total score. Results suggested that a majority (68%) of the reliable variance in the DWS total score is attributable to the general factor. However, only one third of the common variance was attributable to the general factor, which suggests a weaker general factor. Although using a manifest total score is warranted, these results offer some complexity in how researchers should use the DWS in future research. Specifically, the manifest decent work subscales represent variance from the general factor, the subscale in question, and error. Therefore, any relation between a given decent work subscale and a criterion variable is difficult to interpret.

There are several ways to address this concern, each offering options regarding how researchers desire to use the scale. First,

Table 5

Hierarchical Regression Analysis of Decent Work Scale Subscales Predicting Job Satisfaction, Work Meaning, and Withdrawal Intentions-Study 2 (N = 589)

Subscale	Job satisfaction $R = .64$, $R^2 = .41$				Work meaning $R = .55$, $R^2 = .31$				Withdrawal intentions $R = .54$, $R^2 = .29$			
	β	<i>B</i>	<i>SE B</i>	95% CI	β	<i>B</i>	<i>SE B</i>	95% CI	β	<i>B</i>	<i>SE B</i>	95% CI
Safe working conditions	.09*	.21	.09	[-.04, .38]	-.03	-.12	.17	[-.46, .22]	-.10*	-.12	.05	[-.21, -.04]
Access to healthcare	.02	.02	.04	[-.06, .11]	.09*	.21	.09	[-.04, .38]	-.07	-.04	.02	[-.09, .00]
Adequate compensation	.22*	.35	.06	[-.23, .46]	.12*	.38	.12	[-.15, .61]	-.29*	-.23	.03	[-.29, -.17]
Free time and rest	.15*	.25	.06	[-.14, .37]	.06	.18	.12	[-.06, .42]	-.07	-.06	.03	[-.12, .00]
Complementary values	.40*	.73	.07	[-.60, .87]	.45*	1.54	.14	[1.27, 1.81]	-.22*	-.20	.04	[-.27, -.13]

* $p < .05$.

studies should ideally represent decent work as a bifactor model within a latent framework. This type of analysis is best for separating the relative common variance associated with the general factor and the subscales respectively. For researchers unfamiliar with this specific analysis, the steps to construct a bifactor model are as follows. Using structural equation modeling software (e.g., AMOS, EQS, MPlus), the three individual items of each factor should be set to load on their specific factor, and all 15 items should also be set to load on one general decent work factor (see Figure 1). The subscales are uncorrelated in a bifactor model. The general decent work factor and the decent work subscale factors could then be used as a latent constructs within structural equation models of interest to the researcher (e.g., examining how well variables such as economic constraints and work volition predict securing decent work in general as well as the individual components). Second, researchers might be interested in examining only the decent work general factor in a structural model (e.g., how well economic constraints and work volition predict decent work in general). To do this, researchers can load the manifest subscales onto a decent work general factor within a latent model. This would represent the shared variance among subscales - in other words, the general factor in the bifactor model. However, in this case, the unique variance contributed by each individual component would be treated as error.

Researchers might also be interested in operating outside of a latent framework especially if they are interested in examining individual components of decent work. Here, a third option is to estimate a bifactor model in a latent framework and save factor scores to use in SPSS or other programs. If no efforts were made to correct subscale scores, researchers could conduct traditional regression analysis if all subscales were included to remove the common variance between them. A fourth (and preferred) option for using the subscales outside of a latent framework is to use ipsative scoring. This involves calculating the mean of the total DWS (i.e., DWS total score divided by 15) and subtracting it from each subscale mean score to general values for each subscale (e.g., Tracey, 2012). This removes the common variance attributable to the general factor and allows the subscales to be used independently in analyses.

Invariance Tests

Tests for invariance across race/ethnicity, gender, income, and subjective social class were also conducted. Three models—configural, metric, and scalar—were run successively and fit did not substantially decline for any of these models across any of the groupings. This suggests that the general structure of the scale and the responses to the scale itself may be equally valid between participants within these groups. However, because of sample size limitations, the race comparison in the current study collapsed several different racial groups into a nonwhite category. Although the nonwhite category may share some characteristics, such as economic constraints and experiences of marginalization (Duffy et al., 2016), we cannot assume homogeneity among group members. Therefore, future studies should test measurement invariance between specific racial groups. Likewise, the invariance testing for income and social class groups involved arbitrarily splitting the

variables at the mean, which diminishes information from the continuous variables. Therefore, future studies with larger samples sizes may be able to test invariance with more groups than two.

Evidence of Construct Validity

Aside from the CFA detailed above, we also examined relations between the five subscale scores that comprised the DWS and similar instrument scores in order to demonstrate evidence convergent, discriminant, and predictive validity. Hypotheses were partially supported as three of the five subscales strongly correlated with scores from their associated convergent instrument and correlated more strongly with scores from this instrument than other convergent measures. However, in the case of Access to Health care and Health care Satisfaction, this DWS subscale correlated more strongly with Pay Satisfaction. This may have been attributable to getting health care coverage at work only moderately overlapping with one's experience receiving health care. Additionally, because the instrument assessing pay satisfaction also contained items concerning satisfaction with benefits, this may have been a better correlate with health care access. In the case of Safe Working Conditions, this subscale correlated equally as strong with its hypothesized validity scale (measuring job safety) and the validity scale measuring occupational fatigue. This may be attributable the measure of job safety being primarily concerned with physical safety whereas the Safe Working Conditions subscale addressed both emotional and physical safety. It may be that emotional safety is equally related to experiencing the job as safe and experiencing lower levels of occupational fatigue, which also addressed emotions around one's work life.

For all of these convergent associations, none of these correlations were large enough to suggest that the DWS subscales completely overlapped with any of the preexisting scales (van Mierlo, Vermunt, & Rutte, 2009). All five convergent instruments were also found to strongly correlate with the scores from the total DWS and regarding divergent validity, only one of the subscales (Access to Health care) was found to significantly, weakly correlate with ($r = .10$) tenure. In total, this group of correlational results suggested that the DWS scales demonstrate reasonable evidence of convergent and discriminant validity, wherein subscales are associated with analogous constructs and different from constructs such as tenure that are not theoretically implied within the PWT.

Finally, evidence existed for the predictive validity of each of the five DWS subscales. As a group, the subscales accounted for significant variance in the prediction of job satisfaction, work meaning, and withdrawal intentions, confirming propositions from the PWT. Complementary Values and Adequate Compensation were unique predictors for all three outcomes suggesting that these may be the two most important components of decent work when predicting work fulfillment. The other three subscales accounted for unique variance in at least one of these work outcomes, indicating that all components may effect work fulfillment to some degree. It will be essential that future research examine how decent work effects need satisfaction, other aspects of work fulfillment, and overall well-being, which is an additional outcome variable is the PWT.

Implications for Practice

Although further research would be useful in elaborating on the evidence of construct and criterion validity of the DWS scores, the preliminary results do offer some support for using this measure in practice. Psychologists and counselors who are working with clients in work-based interventions might find it useful for clients to assess the extent to which their current jobs meet the criteria for decent work, as advocated by the ILO. The results of this study indicate that each component of decent work is unique, and using the DWS to capture not only an overall perception of decent work but also how and to what degree the subcomponents contribute to that perception is important. Understanding how well each subcomponent of decent work is met will help counselors most effectively target specific parts of a person's work life that need improvement.

Additionally, using this assessment tool with clients may help to foster dialogue about social justice and human rights issues within the workplace. Ideally, work should be a place where clients feel safe, supported, and have their basic needs met (ILO, 2016). Using psychoeducation to help clients understand that access to decent work is a right for all individuals may be facilitative of conversations that promote critical consciousness about work and career. As evidence continues to demonstrate the important role critical consciousness plays in the career development of marginalized populations (Diemer & Rapa, 2016; Diemer, Rapa, Voight, & McWhirter, in press), developing these attitudes in counseling may be beneficial for individual clients and communities.

Limitations and Future Directions

There are a number of limitations in the present studies that should be considered. First, the DWS is a new scale that should continue to be tested, involving replications of these results and use of this scale in different studies. The creation of the DWS will help us further understand the complexities of decent work, but future research must focus on further improvement of this scale. Considering this is the first scale to adequately measure decent work from a self-report, psychological perspective, further research must also aim to advance our knowledge about decent work as a psychological variable, exploring its nature longitudinally and across different groups, as well as its relation with other vocational and well-being outcomes. Second, in considering our use of self-report, it is possible that income and socioeconomic status affected how individuals reported on their working conditions. It is possible that marginalized individuals are more likely to accept their working conditions as adequate as a result of having less power at the workplace. To effectively investigate these perspectives on decent work, future research must incorporate third-party reports to triangulate self-report findings.

Third, concerns exist with our sampling method through MTurk, where scholars have showcased how samples collected through this method may not best represent the population at large (Arditte et al., 2016; Paolacci & Chandler, 2014), and this may have affected how individuals answered the instrument items. A number of articles have been written confirming that this method of data collection is far from perfect—participants are often from lower social status groups than the general population, often skip questions, and can answer at random—but that scores tend to be just as valid as other forms of online data collection, especially when

validity checks are used (Buhrmester et al., 2011; Hauser & Schwarz, 2015). However, one recent article (Arditte, Cek, Shaw, & Timpano, 2016) highlighted how MTurk samples may have elevated mental health concerns compared to the general population. Further, although our samples were relatively close to the national averages, the majority of participants were white, middle-class individuals. As such, it is possible that our samples may not have been representative populations; subsequently, this instrument may not perform as effectively with more diverse populations, and it is essential that future research focus specifically on gathering data from more diverse samples. In particular, data are needed from population-representative samples to confirm the structure and evidence of validity of the scores. Fourth, given small sample sizes according to racial/ethnic groups, in this study we split participants into racial ethnic majority and minority groups. It will be essential that future research demonstrates the appropriate structure and evidence of validity for specific racial/ethnic minority groups.

Finally, a critical future direction in research on the DWS is testing its applicability among non-U.S. populations. From an international perspective, future studies should be contextualized in different sociocultural systems. Existing research in counseling and vocational psychology reveals that different cultures are characterized by different work values and career patterns (Guichard, 2013; Savickas & Porfeli, 2012). The centrality of work differs in different contexts, where the balance between occupation and other social roles could be quite dissimilar. Also, we should consider the specificities of different societal situations, which act as mediator variables in what concerns our relation to work, namely, status of employment (e.g., unemployed, precarious employment), type of social security system, labor laws, and family structure.

Beyond these differential characteristics, the ways in which workers are embedded in different political and economic systems (e.g., democratic vs. dictatorial; liberal vs. state control) are, in addition, complementary variables to be explored as moderator variables within the complexities of decent work. Considering the need to strengthen the invariance models, holistic measures like quality of work life or social, symbolic, and cultural capitals should be introduced to foster a more comprehensive vision. Regarding this goal, in methodological terms, the development of qualitative studies would improve our ability to critically analyze the psychometric differences between groups coming from different sociocultural systems. Combining results from different sources may yield a clearer view of decent work themes in a local and, simultaneously, global perspective, extracting transcultural patterns that will inform future considerations of decent work.

References

- Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. *Health Psychology, 19*, 586–592. <http://dx.doi.org/10.1037/0278-6133.19.6.586>
- Allan, B. A., Autin, K. L., & Duffy, R. D. (2016). Self-determination and meaningful work: Exploring socioeconomic constraints. *Frontiers in Psychology, 7*, 71.
- Arditte, K. A., Çek, D., Shaw, A. M., & Timpano, K. R. (2016). The importance of assessing clinical phenomena in Mechanical Turk re-

- search. *Psychological Assessment*, 28, 684–691. <http://dx.doi.org/10.1037/pas0000217>
- Blau, G. J. (1985). The measurement and prediction of career commitment. *Journal of Occupational Psychology*, 58, 277–288. <http://dx.doi.org/10.1111/j.2044-8325.1985.tb00201.x>
- Blau, G. (1989). Testing the generalizability of a career commitment measure and its impact on employee turnover. *Journal of Vocational Behavior*, 35, 88–103. [http://dx.doi.org/10.1016/0001-8791\(89\)90050-X](http://dx.doi.org/10.1016/0001-8791(89)90050-X)
- Blau, G. (2000). Job, organizational, and professional context antecedents as predictors of intent for interrole work transitions. *Journal of Vocational Behavior*, 56, 330–345. <http://dx.doi.org/10.1006/jvbe.1999.1718>
- Blustein, D. L. (2006). *The psychology of working: A new perspective for career development, counseling, and public policy*. New York, NY: Routledge.
- Blustein, D. L., Kenna, A. C., Gill, N., & DeVoy, J. E. (2008). The psychology of working: A new framework for counseling practice and public policy. *The Career Development Quarterly*, 56, 294–308. <http://dx.doi.org/10.1002/j.2161-0045.2008.tb00095.x>
- Blustein, D. L., Kozan, S., & Connors-Kellgren, A. (2013). Unemployment and underemployment: A narrative analysis about loss. *Journal of Vocational Behavior*, 82, 256–265. <http://dx.doi.org/10.1016/j.jvb.2013.02.005>
- Blustein, D. L., Olle, C., Connors-Kellgren, A., & Diamonti, A. J. (2016). Decent Work: A psychological perspective. *Frontiers in Psychology*, 7, 407. <http://dx.doi.org/10.3389/fpsyg.2016.00407>
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6, 3–5. <http://dx.doi.org/10.1177/1745691610393980>
- Cable, D. M., & DeRue, D. S. (2002). The convergent and discriminant validity of subjective fit perceptions. *Journal of Applied Psychology*, 87, 875–884. <http://dx.doi.org/10.1037/0021-9010.87.5.875>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10, 1–9.
- Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE*, 8, e57410. <http://dx.doi.org/10.1371/journal.pone.0057410>
- Deranty, J. P., & MacMillan, C. (2012). The ILO's decent work initiative: Suggestions for an extension of the notion of “decent work.” *Journal of Social Philosophy*, 43, 386–405. <http://dx.doi.org/10.1111/josp.12003>
- Diemer, M. A., & Rapa, L. J. (2016). Unraveling the complexity of critical consciousness, political efficacy, and political action among marginalized adolescents. *Child Development*, 87, 221–238. <http://dx.doi.org/10.1111/cdev.12446>
- Diemer, M. A., Rapa, L., Voight, A., & McWhirter, E. H. (in press). Critical consciousness: A developmental approach to addressing marginalization and oppression. *Child Development Perspectives*.
- Duffy, R. D., Allan, B. A., Autin, K. L., & Bott, E. M. (2013). Calling and life satisfaction: It's not about having it, it's about living it. *Journal of Counseling Psychology*, 60, 42–52. <http://dx.doi.org/10.1037/a0030635>
- Duffy, R. D., Allan, B. A., Autin, K. L., & Douglass, R. P. (2014). Living a calling and work well-being: A longitudinal study. *Journal of Counseling Psychology*, 61, 605–615. <http://dx.doi.org/10.1037/cou0000042>
- Duffy, R. D., Blustein, D. L., Diemer, M. A., & Autin, K. L. (2016). The Psychology of Working Theory. *Journal of Counseling Psychology*, 63, 127–148. <http://dx.doi.org/10.1037/cou0000140>
- Duffy, R. D., Bott, E. M., Allan, B. A., Torrey, C. L., & Dik, B. J. (2012). Perceiving a calling, living a calling, and job satisfaction: Testing a moderated, multiple mediator model. *Journal of Counseling Psychology*, 59, 50–59. <http://dx.doi.org/10.1037/a0026129>
- Duffy, R. D., Diemer, M. A., Perry, J. C., Laurenzi, C., & Torrey, C. L. (2012). The construction and initial validation of the Work Volition Scale. *Journal of Vocational Behavior*, 80, 400–411. <http://dx.doi.org/10.1016/j.jvb.2011.04.002>
- Edwards, J. R., & Rothbard, N. P. (1999). Work and family stress and well-being: An examination of person-environment fit in the work and family domains. *Organizational Behavior and Human Decision Processes*, 77, 85–129. <http://dx.doi.org/10.1006/obhd.1998.2813>
- Gallup. (2016). *U.S. uninsured rate at 11.0%, lowest in eight-year trend*. Retrieved from <http://www.gallup.com/poll/190484/uninsured-rate-lowest-eight-year-trend.aspx>
- Guichard, J. (2013). Career guidance, education, and dialogues for a fair and sustainable human development. In Inaugural Conference of the UNESCO Chair of Lifelong Guidance and Counselling, University of Wroclaw, Wroclaw, Poland.
- Hammer, L. B., & Zimmerman, K. L. (2011). Quality of work life. In S. Zedeck (Ed.), *American psychological association handbook of industrial and organizational psychology* (pp. 399–431). Washington, DC: American Psychological Association.
- Hauser, D. J., & Schwarz, N. (2015). It's a trap! Instructional manipulation checks prompt systematic thinking on “tricky” tasks. *SAGE OPEN*, 5, 1–6. <http://dx.doi.org/10.1177/2158244015584617>
- Hayes, B. E., Perander, J., Smecko, T., & Trask, J. (1998). Measuring perceptions of workplace safety: Development and validation of the work safety scale. *Journal of Safety Research*, 29, 145–161. [http://dx.doi.org/10.1016/S0022-4375\(98\)00011-5](http://dx.doi.org/10.1016/S0022-4375(98)00011-5)
- Hays, R. D., Davies, A. R., & Ware, J. E. (1987, August 19). *Scoring the medical outcomes study Patient Satisfaction Questionnaire: PSQ-III*. [Memorandum].
- Heneman, H. G., III, & Schwab, D. P. (1985). Pay satisfaction: Its multidimensional nature and measurement. *International Journal of Psychology*, 20, 129–141.
- Heppner, P., Wampold, B., Owen, J., Wang, K., & Thompson, M. (2015). *Research in counseling* (4th ed.). San Francisco, CA: Cengage Publishing.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. <http://dx.doi.org/10.1080/10705519909540118>
- International Labor Organization. (2008). *Work of work report 2008: Income inequalities in the age of financial globalization*. Retrieved from http://www.ilo.org/wcmsp5/groups/public/@dgreports/@dcomm/@publ/documents/publication/wcms_100354.pdf
- International Labor Organization. (2012). *Decent work indicators: Concepts and definitions*. Retrieved from http://www.ilo.org/wcmsp5/groups/public/---dgreports/---integration/documents/publication/wcms_229374.pdf
- International Labor Organization. (2016). *Decent work agenda*. Geneva, Switzerland: International Labor Office.
- Judge, T. A. (1993). Validity of the dimensions of the Pay Satisfaction Questionnaire: Evidence of differential prediction. *Personnel Psychology*, 46, 331–335. <http://dx.doi.org/10.1111/j.1744-6570.1993.tb00877.x>
- Judge, T. A., Locke, E. A., Durham, C. C., & Kluger, A. N. (1998). Dispositional effects on job and life satisfaction: The role of core evaluations. *Journal of Applied Psychology*, 83, 17–34. <http://dx.doi.org/10.1037/0021-9010.83.1.17>
- Judge, T. A., & Welbourne, T. M. (1994). A confirmatory investigation of the dimensionality of the Pay Satisfaction Questionnaire. *Journal of Applied Psychology*, 79, 461–466. <http://dx.doi.org/10.1037/0021-9010.79.3.461>

- Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice: Principles, advances, and applications. *The Counseling Psychologist*, 34, 684–718. <http://dx.doi.org/10.1177/0011000006286347>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: Guilford.
- Lent, R. W., & Brown, S. D. (2013). Social cognitive model of career self-management: Toward a unifying view of adaptive career behavior across the life span. *Journal of Counseling Psychology*, 60, 557–568. <http://dx.doi.org/10.1037/a0033446>
- Little, T. D. (2000). On the comparability of constructs in cross-cultural research. A critique of Cheung and Rensvold. *Journal of Cross-Cultural Psychology*, 31, 213–219. <http://dx.doi.org/10.1177/0022022100031002004>
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, 44, 1–23. <http://dx.doi.org/10.3758/s13428-011-0124-6>
- Mulvey, P. W., Miceli, M. P., & Near, J. P. (1990, May). The Pay Satisfaction Questionnaire: A review and extension (Ohio State University working paper, Report No. WPS 90–36). Columbus, OH: Ohio State University.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Author.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32, 396–402. <http://dx.doi.org/10.3758/BF03200807>
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, 23, 184–188. <http://dx.doi.org/10.1177/0963721414531598>
- Parent, M. C. (2013). Handling item level missing data: Simpler is just as good. *The Counseling Psychologist*, 41, 568–600. <http://dx.doi.org/10.1177/0011000012445176>
- Parker, S. K. (2014). Beyond motivation: Job and work design for development, health, ambidexterity, and more. *Annual Review of Psychology*, 65, 661–691. <http://dx.doi.org/10.1146/annurev-psych-010213-115208>
- Quintana, S. M., & Maxwell, S. E. (1999). Implications of recent developments in structural equation modeling for counseling psychology. *The Counseling Psychologist*, 27, 485–527. <http://dx.doi.org/10.1177/0011000099274002>
- Ribeiro, M. A., Silva, F. F., & Figueiredo, P. M. (2016). Discussing the notion of decent work: Senses of working for a group of Brazilian workers without college education. *Frontiers in Psychology*, 7, 207. <http://dx.doi.org/10.3389/fpsyg.2016.00207>
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, 21, 137–150. <http://dx.doi.org/10.1037/met0000045>
- Saari, L. M., & Judge, T. A. (2004). Employee attitudes and job satisfaction. *Human Resource Management*, 43, 395–407. <http://dx.doi.org/10.1002/hrm.20032>
- Savickas, M. L., & Porfeli, E. J. (2012). Career Adapt-Abilities Scale: Construction, reliability, and measurement equivalence across 13 countries. *Journal of Vocational Behavior*, 80, 661–673. <http://dx.doi.org/10.1016/j.jvb.2012.01.011>
- Sharone, O. (2013). *Flawed system/flawed self: Job searching and unemployment experiences*. Chicago, IL: University of Chicago Press. <http://dx.doi.org/10.7208/chicago/9780226073675.001.0001>
- Steger, M. F., Dik, B. J., & Duffy, R. D. (2012). Measuring meaningful work: The Work and Meaning Inventory (WAMI). *Journal of Career Assessment*, 20, 322–337. <http://dx.doi.org/10.1177/1069072711436160>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Boston, MA: Pearson Education.
- Tracey, T. J. G. (2012). Problems with single interest scales: Implications of the general factor. *Journal of Vocational Behavior*, 81, 378–384. <http://dx.doi.org/10.1016/j.jvb.2012.10.001>
- U.S. Census Bureau. (2015). Annual estimates of the resident population by sex, race, and Hispanic origin for the United States, states, and counties: April 1, 2010 to July 1, 2015. Retrieved from <http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk>
- van Mierlo, H., Vermunt, J. K., & Rutte, C. G. (2009). Composing group-level constructs from individual-level survey data. *Organizational Research Methods*, 12, 368–392. <http://dx.doi.org/10.1177/1094428107309322>
- Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: An analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science*, 44, 119–134. <http://dx.doi.org/10.1007/s11747-015-0455-4>
- Ware, J. E., Jr., Snyder, M. K., Wright, W. R., & Davies, A. R. (1983). Defining and measuring patient satisfaction with medical care. *Evaluation and Program Planning*, 6, 247–263. [http://dx.doi.org/10.1016/0149-7189\(83\)90005-8](http://dx.doi.org/10.1016/0149-7189(83)90005-8)
- Ware, J. E., Snyder, M. K., & Wright, W. R. (1976a). *Development and validation of scales to measure patient satisfaction with medical care services: Vol. I. Part A. Review of literature, overview of methods, and results regarding construction of scales* (NTIS Publication No. PB 288-329). Springfield, VA: National Technical Information Service.
- Ware, J. E., Snyder, M. K., & Wright, W. R. (1976b). *Development and validation of scales to measure patient satisfaction with medical care services: Vol. I. Part B. Results regarding scales constructed from the Patient Satisfaction Questionnaire and measures of other health care perceptions* (NTIS Publication No. PB 288-330). Springfield, VA: National Technical Information Service.
- Weston, R., & Gore, P. A. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist*, 34, 719–751. <http://dx.doi.org/10.1177/0011000006286345>
- Winwood, P. C., Lushington, K., & Winefield, A. H. (2006). Further development and validation of the Occupational Fatigue Exhaustion Recovery (OFER) Scale. *Journal of Occupational and Environmental Medicine*, 48, 381–389. <http://dx.doi.org/10.1097/01.jom.0000194164.14081.06>
- Winwood, P. C., Winefield, A. H., Dawson, D., & Lushington, K. (2005). Development and validation of a scale to measure work-related fatigue and recovery: The Occupational Fatigue Exhaustion/Recovery Scale (OFER). *Journal of Occupational and Environmental Medicine*, 47, 594–606. <http://dx.doi.org/10.1097/01.jom.0000161740.71049.c4>
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34, 806–838. <http://dx.doi.org/10.1177/0011000006288127>

Received August 22, 2016

Revision received December 7, 2016

Accepted December 7, 2016 ■