Work-nonwork balance: Development and validation of a global and multidimensional measure☆

Julie Holliday Wayne [a,*], Hoda Vaziri [b], Wendy J. Casper [c]

[a] 369 Farrell Hall, School of Business, Wake Forest University, Winston-Salem, NC 27109, USA
[b] Department of Management, G. Brint Ryan College of Business, University of North Texas, USA
[c] University of Texas at Arlington, Department of Management, Box 19467, Arlington, TX 76019-0467, USA

A B S T R A C T

Based on Casper et al.'s (2018) definitions of global balance and involvement, affective, and effectiveness dimensions of work-nonwork balance, we developed and validated a measure, with a total of 20 items – 5 items for global balance and 5 items for each dimension. Using 3 studies, we established the content adequacy, factor structure, reliability, and gender and parental status invariance of the measure. Moreover, we show that our measure converges with prior measures of balance (Carlson et al., 2009; Valcour, 2007) and is distinct from role-specific measures of involvement, satisfaction, and performance, as well as life satisfaction. Finally, we demonstrate that global balance and its dimensions uniquely predict employee engagement, citizenship behaviors, organizational commitment, turnover intentions, emotional exhaustion, and/or health, above and beyond existing measures of balance. Thus, our research provides a comprehensive, validated, multidimensional measure of work-nonwork balance and offers unique explanation of valued attitudes and behaviors. Future theoretical, research and practical implications are discussed.

Popular discussion of work-life balance has generally outpaced scientific inquiry, but research on the meaning, measurement, and correlates of balance is rapidly accumulating. Initially, balance was defined and measured as low conflict and high enrichment (Frone, 2003). Then, scholars began to conceptualize balance as a global appraisal of how work and family fit together (Voydanoff, 2005). Wayne et al. (2017) demonstrated that balance is not a unitary construct and that various measures of balance are not interchangeable. They conceptually differentiated two common measures - "balance satisfaction" (Valcour, 2007) and "balance effectiveness" (Carlson et al., 2009), and empirically differentiated them from one another and from work-family conflict and enrichment. Each measure related differently to employee attitudes and behaviors and did so above and beyond conflict and enrichment. This initial evidence suggested that balance may not be unidimensional, and that different meanings and measures have implications for balance theory and research.

Subsequently, Casper et al. (2018) provided a comprehensive review of all definitions and measures of balance. They found there were approximately 233 conceptual and 204 operational definitions. Frequently, (38% of) measures were developed for a single study and used only once. Of measures drawn from past studies, about 40% used conflict and enrichment rather than measuring balance as a global construct. When global balance measures were used, they often lacked rigorous scale development practices and there were

often misalignments between conceptual and operational definitions. Finally, a substantial portion (19%) of operational definitions used single item measures which is problematic because internal consistency cannot be assessed, and complex constructs, like balance, cannot sufficiently be captured (Loo, 2002). Casper et al. concluded that "relative to work-family research in general, problems with measurement are more significant in the balance literature" (p. 196) and that the observed construct validity problems are an impediment to theory building.

Given organizational researchers must ensure accurate measurement of constructs under examination (Barrett, 1972), care is needed to ensure that items capture the content universe (i.e., content validity), relate to other constructs as expected (convergent and discriminant validity), and are reliable. Without such efforts, it is unclear whether extant balance scales accurately measure the construct of interest. Moreover, most extant scales are unidimensional, despite the fact that work by Wayne et al. (2017) and Casper et al. (2018) suggest balance may be multidimensional. Using a single measure, therefore, is likely construct deficient. Finally, Casper et al. also demonstrated meta-analytically that the relation of balance with family, job, and life satisfaction differed as a function of measure, suggesting it is difficult to aggregate across studies using unique measures.

In sum, most measures of balance suffer from incomplete development or validation, do not capture the complex and multidimensional nature of the construct, and no single scale is used across studies as needed for research to accumulate (Casper et al., 2018). To begin to remedy these problems, Casper and colleagues provided conceptual definitions of global balance and 3 facets: involvement, affective, and effectiveness balance, and called for scholars to develop reliable and valid measures. The present research addresses this critical next step by developing and validating a comprehensive measure of work-nonwork balance.

Content adequacy, or ensuring the content of a measure covers a representative sample of the domain to be assessed, is the first psychometric property of a measure that should be evaluated (Schriesheim et al., 1993). As noted by Colquitt et al. (2019), content validation is surprisingly absent, even from scales published in top journals. From the thoroughly articulated definitions given by Casper et al. (2018), we develop items to tap the meaning of each facet as well as global balance and conduct two studies assessing content adequacy and factorial validity to refine the items in the scale.

Evidence for construct validity is also provided by showing that new measures converge with similar measures and diverge from different ones. Following Hinkin (1995, 1998), in Study 3, we confirm the factor structure of our measure using an independent sample, examine whether the factor structure is invariant across gender and parental status, show convergent validity with other measures of work-family balance, and examine discriminant validity from domain-specific satisfaction, involvement, and performance. A final way of examining construct validity is by demonstrating the extent to which a measure produces results that match theoretical predictions, particularly above and beyond extant measures. Thus, we test the extent to which our balance measure predicts valued outcomes—namely, employee engagement, organizational citizenship behaviors (OCBs), organizational commitment, turnover intentions, emotional exhaustion, and health—above and beyond extant balance measures. In doing so, our research contributes to the literature by remedying identified measurement problems impeding research and theory on balance (Casper et al., 2018).

1. Theories and definitions of work-nonwork balance: a review

In 1996, Marks and MacDermid discussed the organization of multiple roles and challenged the notion that all people organize roles in a salience hierarchy. Instead, they developed a "theory of role balance" suggesting that people take an evenhanded approach to their roles and that people with more balanced role systems experience less strain, greater role ease and well-being and more positive role-specific experiences. They defined "role balance" as approaching every role and role partner with attention and care, conceptualizing role balance as a cognitive-affective structure, or internal working model, that also encompasses a behavioral pattern of acting with a general orientation *across* roles (i.e., not a role-specific construct). They provided an eight-item scale to assess role balance as a unidimensional appraisal of full engagement across all life roles (e.g., "Nowadays, I seem to enjoy every part of my life equally well," p. 422), but no development or validation efforts were described. In two studies, their measure of role balance was associated with greater role ease, role functioning, and well-being.

A few years later, scholars began to draw from existing work-family constructs and validated scales to define and measure balance as the absence of conflict (Duxbury & Higgins, 2001), or a combination of low conflict and high enrichment (Frone, 2003). As these definitions drew on existing constructs, there was limited theory development or measurement advancement on balance. Others, however, used the notion of a balance beam with the burden of family on one side and work on the other (Halpern & Murphy, 2005) and drew on the notion of equality. For example, Greenhaus and colleagues defined work-family balance as being equally engaged in and equally satisfied with work roles and family roles (Greenhaus et al., 2003).

Around this same time and in an effort to expand theory, Voydanoff (2002, 2005) developed integrative models to explain balance. She drew on person-environment (P-E) fit to suggest that work and family demands and resources predict demands-resources fit in which family and work demands are met by resources present in both domains. Fit, in turn, predicts balance which she defined as "a global assessment that work and family resources are sufficient to meet work and family demands such that participation is effective in both domains" (p. 825). Balance, in turn, is theorized to foster better role performance and quality. However, she did not provide a scale to measure balance nor did she conduct empirical tests of her model.

In 2007, Valcour drew from Voydanoff's (2005) work to define and measure "satisfaction with work-family balance." She conceptualized balance as an attitude with affective and cognitive components. Accordingly, she defined balance as "an overall level of contentment resulting from an assessment of one's degree of success at meeting work and family role demands" (p. 1512). She created a 5-item, unidimensional measure where participants rated their satisfaction such as with "the way you divide your time between work and personal or family life" (p. 1517) and found that job complexity and control over work time were associated with greater balance satisfaction.

That same year, Grzywacz and Carlson (2007) defined balance in a fundamentally different way. Rather than viewing balance as a psychological construct such as an attitude (Valcour, 2007) or internal working model (Marks & MacDermid, 1996), they conceptualized balance as a relational construct, defining it as "the accomplishment of role-related expectations that are negotiated and shared between an individual and his/her role-related partners in the work and family domain" (p. 458). Carlson et al. (2009) developed a 6-item, unidimensional measure which asked respondents to self-report the extent to which they meet expectations that are shared with work and family role partners (e.g. "I do a good job of meeting the role expectations of critical people in my work and family life"). Though they demonstrated the discriminant validity of their measure from conflict and enrichment, they did not describe efforts to validate this scale such as content validation with a full list of items, exploratory factor analyses to trim items, or other aspects of construct validation.

In the most recent theoretical explanation, Greenhaus and Allen (2011) developed a definition of balance and proposed a model of how it occurs. Like Voydanoff (2005), they emphasize fit but in terms of fit with values rather than fit of demands and resources (Voydanoff, 2005). They defined balance as when role effectiveness and satisfaction are consistent with life priorities. In their model, they contend that work, family, and personal characteristics relate to conflict and enrichment which then create role satisfaction and effectiveness. In turn, the extent to which a person is satisfied and effective in highly valued roles is a proximal determinant of "feelings of balance." This final notion is contrary to role balance theory (Marks & MacDermid, 1996) and Voydanoff's (2005) model which position role satisfaction and effectiveness as outcomes, rather than determinants of balance. To date, no scales have been developed to measure their definition. Empirically, we are aware of only one study thus far that has tested elements of their model (Pattusamy & Jacob, 2016), but it did not measure their definition of balance nor use a longitudinal design to test the temporal sequencing proposed in their model.

In sum, as this review highlights and as addressed by the work of Casper et al. (2018), there has been a lack of consensus on theoretical models and conceptual definitions of balance. Also, there is not a well-accepted, well developed and validated measure, which is a prerequisite before model testing can be done. As noted, developing and validating a comprehensive measure of balance is the goal of the current research.

2. Conceptual definitions of balance and its facets

A critical first step to developing a new measure is providing a conceptual definition. For this purpose, we draw from Casper et al. (2018). Drawing from their comprehensive literature review, dictionary definitions, and qualitative data about how employees define balance, Casper et al. (2018) offered a thorough definition of balance:

> "*Employees' evaluation of the favorability of their combination of work and nonwork roles, arising from the degree to which their affective experiences and their perceived involvement and effectiveness in work and nonwork roles are commensurate with the value they attach to these roles*."

There are several important aspects to note about their definition. First, unlike many extant definitions or measures of balance which tend to focus on work and family roles, Casper et al. (2018) use the term "work-*nonwork* balance" to reflect that balance incorporates nonwork roles besides family, such as community member, friend, or even the self. Like Valcour (2007) and Marks and MacDermid (1996), they view balance as a psychological construct, an attitude, that refers to an orientation across roles, with the attitude object of "*combination of work and nonwork roles*" (p. 197). A third aspect of their definition is that, like other attitudes, balance has cognitive-affective components (Marks & MacDermid, 1996) such that balance is based on an affective factor and two cognitive factors (i.e., involvement, effectiveness). Finally, they adopt the notion of fit (Greenhaus & Allen, 2011) to propose that balance is a function of the extent to which one is satisfied, effective, and involved across highly valued roles. Because values are central to how experiences translate into balance appraisals and each person values each role differently, distinct levels of affect, effectiveness, and involvement in work and nonwork roles foster balance for each person.

Drawing parallels from the construct of job satisfaction, Casper et al. (2018) suggested that balance be assessed both as a global unidimensional reflective construct (global balance) and as a multidimensional formative construct (affective, involvement, and effectiveness balance). We extend their work by elaborating on their attitude object of the "combination of work and nonwork roles" by using synonyms of the term "combination." In particular, we define balance at a global level as *the overall combination, fit, balance, harmony, or integration of one's work and nonwork roles*. Given their multidimensional construct, Casper et al. (2018) defined each of their three dimensions, with affective balance defined as "*The perception that one experiences sufficiently pleasant emotions in work and nonwork roles commensurate with the value attached to those roles*;" effectiveness balance as "*The perception that one's effectiveness in work and nonwork roles is commensurate with the value attached to the roles*;" and involvement balance as "*The perception that one's involvement in work and nonwork roles is commensurate with the value attached to the roles*" (p. 198).

For the multi-item global assessment of how one combines work and nonwork roles (i.e., global balance), Casper et al. (2018) recommended a semantic-differential approach (Osgood et al., 1957) where respondents rate how work and nonwork roles are combined on bipolar dimensions such as successful-unsuccessful, involved-uninvolved; and satisfying-unsatisfying. To measure facets, they recommended subjective appraisals of fit using a Likert scale and provided example items wherein people would rate their "perceptions of affective balance (e.g., 'Reflecting on combining my work and nonwork roles, I am happy in the roles that I value most'), effectiveness balance (e.g., 'In considering how I combine my work and nonwork roles, I do well in roles that are my biggest priorities'), and involvement balance (e.g., 'The time I spend in my work and nonwork roles reflects what I most value') (p. 200). These conceptual definitions and measurement suggestions guided the first phase of our work: content adequacy, exploratory factor analyses, and item refinement.

3. Study 1 content validation method & results

3.1. Item development

Based on Casper et al.'s (2018) conceptual definitions and exemplar items, three subject matter experts (SMEs), who have written and published extensively on the work-life interface, independently developed items, for a total of 31 items to measure affective balance, 30 items to measure involvement balance, and 29 to measure effectiveness balance. These 3 SMEs, as well as another expert on the work-life interface, provided content adequacy ratings to reduce item content. SMEs were provided the definition of each facet including all proposed items to measure that facet and asked to indicate whether the item was consistent with the definition, rating "yes", "no", or "maybe/not sure". If 75% or more of the SMEs responded yes, items were retained. This cut-off is more stringent than the 70% criteria used in previous content adequacy research (i.e., Carlson et al., 2000; Schriesheim & Hinkin, 1990). Using this criterion, the measure was reduced to 15 items for affective, 22 items for involvement, and 13 items for effectiveness balance. Consistent with recommendations by Casper et al., 24 items were then written as semantic differentials to assess the combination of work and nonwork roles (i.e., global balance), such as "balanced-unbalanced; engaged-disengaged; pleasant-unpleasant; and successful-unsuccessful".

Table 1
Final list of retained items for facets of balance after content validation study (Study 1) and initial exploratory factor analysis results (Study 2).

Items	C[a]	F1	F2	F3	F4	F5
Involvement balance						
My involvement in work and nonwork roles reflects what I care most about.	0.52	**0.45**	−0.06	−0.04	0.21	0.32
I do not engage in my important work and nonwork roles as much as I want.	0.66	0.28	−0.04	**0.77**	−0.28	0.01
I am involved in the roles that are my biggest priorities.	0.54	**0.41**	0.06	0.05	0.18	0.23
I am able to devote enough attention to important work and nonwork activities.	0.55	**0.78**	0.03	0.05	−0.12	−0.05
The time I spend in my work and nonwork roles reflects what I most value.	0.50	**0.51**	0.16	0.08	0.03	0.02
I am able to be adequately involved in the work and nonwork roles that matter most to me.	0.50	**0.69**	−0.10	−0.04	0.01	0.19
The time I spend in work and activities outside of work reflects my life priorities.	0.62	**0.77**	0.00	0.02	0.11	−0.18
When I consider all the ways I invest energy at work and outside of work, my investment reflects what I care about most.	0.50	**0.54**	−0.03	−0.02	0.18	0.16
I spend enough time on important work and nonwork activities.	0.62	**0.65**	0.12	0.04	−0.04	0.12
I do not devote as much time as I want to important work and nonwork activities.	0.70	0.19	0.01	**0.84**	−0.36	−0.05
Based on what matters most to me, I devote the right amount of my time to work and nonwork roles.	0.59	**0.92**	−0.15	−0.02	−0.05	−0.04
I am adequately engaged in my most highly valued work and nonwork roles.	0.49	**0.63**	0.11	−0.01	0.00	−0.03
Overall, I am adequately involved in my most important work and nonwork roles.	0.61	**0.57**	0.29	−0.02	0.06	−0.22
Effectiveness balance						
When I think about the many things I do at work and outside work, I am effective at the things that are most important.	0.62	−0.03	0.32	0.02	0.04	**0.61**
I perform well in the life roles that I really value.	0.52	0.01	**0.67**	0.10	−0.01	0.02
I am effective in combining my most important work and nonwork roles.	0.51	**0.45**	0.20	0.05	−0.01	0.17
I do well in roles that are my biggest priorities.	0.54	−0.25	**0.84**	0.02	−0.02	0.18
I am able to effectively handle important work and nonwork responsibilities	0.60	0.11	**0.72**	−0.07	−0.14	0.18
I am successful in work and nonwork roles that matter to me.	0.61	0.12	**0.63**	0.02	0.13	−0.17
I perform well in my most highly valued work and nonwork roles.	0.57	0.05	**0.62**	−0.04	0.10	0.10
I accomplish what is most important at work and outside of work.	0.50	0.17	0.34	0.13	0.01	0.26
I am effective in work and nonwork roles and activities that matter to me.	0.47	0.14	**0.45**	−0.03	0.11	0.12
I do a good job in the work and nonwork roles that are my biggest priorities.	0.59	0.17	**0.54**	−0.03	0.12	0.02
I am doing well at things that are really important in life.	0.55	0.16	**0.49**	0.05	0.13	0.02
Overall, I am effective in combining my most important work and nonwork roles.	0.60	0.38	**0.53**	−0.15	−0.12	0.11
Affective balance						
I am happy in the roles that I value most.	0.63	0.30	0.29	0.07	0.31	−0.08
I experience a lot of positive emotions in my most highly valued work and nonwork roles.	0.69	−0.04	0.07	−0.03	**0.77**	0.12
I experience a lot of negative emotions in important work and nonwork roles.	0.61	−0.27	0.18	**0.71**	0.25	−0.09
I am happy in the work and nonwork roles that are most important to me.	0.64	0.19	0.34	−0.05	**0.44**	−0.13
I wish I were happier in some of my important roles at work and outside of work.	0.58	−0.05	−0.33	**0.69**	0.35	0.11
I am happy with the work and nonwork aspects of my life that are important to me.	0.60	0.16	0.11	−0.02	**0.59**	0.00
I feel satisfied in the work and nonwork roles that are most important to me.[b]	0.60	**0.41**	0.10	0.09	0.34	−0.05
I am dissatisfied in some high priority roles at work and/or outside work.	0.80	−0.14	0.04	**0.85**	0.15	0.08
I am content with how things are going in the life roles that are my top priorities.[b]	0.52	**0.43**	0.01	−0.06	0.38	0.04
Some important aspects of my work and nonwork life are not pleasant for me.	0.60	0.00	0.13	**0.73**	−0.01	−0.02

Notes. Items in **bold** were retained. Items in *italics* were omitted because they loaded on a pseudo-factor. Plain text items were omitted because they loaded on the wrong factor or had lower loadings.
[a] Extracted communalites.
[b] Items appear to load on effectiveness factor, but after omitting negatively valenced items in *italics*, items loaded on the intended affective factor.

3.2. Content adequacy ratings and item refinement

Following this initial item development and refinement, content adequacy was rated by 11 SMEs for each facet and for global balance items. Per scale development guidelines (Colquitt et al., 2019; Hinkin, 1995; Schriesheim et al., 1993), we recruited 11 faculty and PhD students in Organizational Behavior. At the top of each page of the survey, the Casper et al. (2018) definitions were provided: Global balance: "'Employees' appraisals of how they combine work with nonwork roles' where the attitude object is the 'combination of work and nonwork roles'. Affective balance: 'The perception that one experiences sufficiently pleasant emotions in work and nonwork roles commensurate with the value attached to those roles. People should experience high (low) affective balance when they have more (fewer) positive and fewer (more) negative emotions in highly valued roles.' Involvement balance: 'The perception that one's involvement in work and nonwork roles is commensurate with the value attached to the roles. People experience high (low) involvement balance when they believe that they are adequately (inadequately) engaged in highly valued roles.' Effectiveness balance: 'The perception that one's effectiveness in work and nonwork roles is commensurate with the value attached to the roles. People should experience high (low) effectiveness balance when they believe that they perform well (poorly) in highly valued roles.'"

Following definitions at the top of each page, 49 facet items were randomly presented with approximately 15 items per page. All semantic differential items were presented on the same page in random order. SMEs were instructed to use their expert judgment to indicate which definition each item best fit. They chose either "global balance, affective balance, involvement balance, effectiveness balance, or not sure/none." On each page, we also asked them to provide feedback on any item or on any content that was not included that should be.

Finally, we also examined the extent to which SMEs concurred that the conceptual definitions of balance we are using (as provided by Casper et al., 2018) represented their expert views on the meaning of balance. They were instructed "As an expert, please rate the extent to which YOU believe each of the following definitions represents the meaning of "work-nonwork balance," rated on a 5-point Likert scale ranging from 1 "*Not at all representative of the meaning of work-nonwork balance*" to 5 "*Extremely representative of the meaning of work-nonwork balance.*"

Again, when 75% of SMEs rated the item as consistent with its intended definition, we retained it for 10 affective, 13 involvement, and 12 effectiveness balance items. Of the semantic-differential items intended to assess global balance, only one item met this criterion ("balanced-unbalanced"). Other semantic differential items (e.g., "satisfying-unsatisfying") were divided in ratings among global balance (e.g., 27%), the affective facet (e.g., 36%) or not sure (e.g., 36%). Participants also noted that rating semantic differential items was confusing. Given that generally, the semantic differential items did not meet the agreement criterion, none of these items were retained.

The definition SMEs said most represented the meaning of balance was global balance (M = 4.1, SD = 0.99), followed by affective balance (M = 3.7, SD = 0.82), involvement (M = 3.3, SD = 1.06) and effectiveness balance (M = 3.3, SD = 1.16). All definitions were rated as somewhat to very representative of the meaning of balance. At this point, though we had items with content validity to measure each facet, we did not have items to measure global balance. As such, we developed 8 Likert scale items defining balance per Casper et al. (2018) where the target object is the combination of work and nonwork roles. To differentiate global balance from its facets and to align with the Casper et al. (2018) definition of global balance, we did not reference satisfaction, effectiveness, or involvement in the global balance items. Instead, global balance items referenced the overall combination, fit, balance, harmony, or integration among work and nonwork roles (e.g., "My work and nonwork roles are combined in ways that are harmonious; Overall, my work and nonwork roles fit together"). There was 100% agreement among 3 SMEs that items represented the Casper et al. definition of global balance. At this point, our refined scale contained 43 items (35 facet items and 8 global balance items, see Tables 1 and 2) for the next phase of scale development.

4. Study 2 exploratory factor analyses method & results

We conducted study 2 to perform exploratory factor analysis (EFA) to refine our balance scale based on items that loaded on the intended factor (Hinkin, 1998) and to determine which items performed well enough to be retained.

Table 2
Final list of retained items for global balance after content validation study (Study 1) and initial exploratory factor analysis results (Study 2).

Global balance items	Communality	F1
Overall, my combination of work and nonwork roles is balanced.	0.65	**0.66**
There is harmony in how I blend my work and nonwork roles.	0.59	**0.82**
My work and nonwork roles are combined in ways that are harmonious.	0.58	**0.81**
Overall, my work and nonwork roles are integrated.	0.57	**0.70**
My work and nonwork roles fit together.	0.52	**0.68**
Overall, my work and nonwork roles fit together.	0.65	**0.77**
All things considered, my work and nonwork roles are in balance.	0.59	**0.63**
All in all, my work and nonwork roles are in harmony.	0.70	**0.82**

Notes. Items in **bold** were retained for the final scale. Plain text items were omitted to retain items with higher factor loadings.

4.1. Participants

Participants were recruited via Amazon's Mechanical Turk (MTurk) which evidence suggests yields representative samples that have increased generalizability relative to other samples in organizational research (e.g., Highhouse & Zhang, 2015). To address concerns about data quality, we followed several of the best practice guidelines offered by Aguinis et al. (2020). To avoid misrepresentation, we used screener tools to identify qualified participants (e.g., employed full-time) where participants were presented with a variety of questions (e.g., age, marital status, race, etc.) such that the desired qualifications were not explicit. To address language and data quality concerns, only adult, U.S. participants with a 95% approval rating who had successfully completed at least 1000 tasks were invited to participate. We defined full time employment as 30 or more hours following the Internal Revenue Service (2020) definition (www.irs.gov). Another best practice we followed to ensure effortful responding was the inclusion of two attention checks (e.g., "Please leave this item blank"), and anyone who missed one was omitted from analyses. Participants were paid $1.75 for the survey, a fair price in terms of minimum wage, to motivate their attention to the task. Finally, all scale points were labeled to improve attention.

There were 209 participants, with 39% female, 58% Caucasian, 73% married, and 67% reported at least one child at home and were an average age of 35 years. Participants worked in more than 22 different industries with the most common in business/financial operations (17%), sales (13%), computer/math (14%), and healthcare (8%). They worked, on average, 30.69 h per week and spent 10.81 h on caregiving, 9.26 on household responsibilities, 10.48 in leisure, and 4.67 h in community/volunteer service.

4.2. Measures

The 43 items from Study 1 were provided (see Tables 1 and 2). Respondents were instructed to think about the past month and to rate their agreement with each statement about how they combine their work and nonwork roles. They were told that when responding to items, to consider the work and nonwork roles that mattered most to them. To make sure participants considered nonwork roles beyond family, work and nonwork were each defined. "**Work roles** are those roles and activities that you pursue as part of your paid employment including the responsibilities that are part of your job. Examples include job-related tasks and duties, and anything that is part of your occupational or organizational roles. **Nonwork roles** are roles and activities that you pursue outside of paid employment, including family-related roles and activities and other activities you engage in due to personal interest or for people outside your family. Examples include family relationships and responsibilities, working on a hobby, spending time with friends, volunteering, religious activities, and exercising." They rated agreement with each item using a 5-point Likert scale 1) *Strongly Disagree* to 5) *Strongly Agree*.

4.3. EFA of facets of balance

To examine the dimensionality of facets of balance, we conducted principal axis EFA using promax rotation (Hinkin, 1998). The results revealed 5 factors based on commonly recommended criteria (e.g., eigenvalues and the scree plot; Hayton et al., 2004). The Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) exceeded 0.80, suggesting sampling was adequate (KMO = 0.94). Factor loadings on these 5 factors are provided in Table 1. Items with low loadings (less than 0.40 on the intended factor) or with cross-loading of more than 0.35 on any other factor were omitted (Hinkin, 1998). We chose the cutoff point of 0.40 for factor loadings

Table 3
Final results of exploratory factor analysis of facets of balance – Study 2.

	Item	C[a]	F1	F2	F3
	Involvement balance				
1	I am able to devote enough attention to important work and nonwork activities.	0.56	**0.78**	0.08	−0.13
2	I am able to be adequately involved in the work and nonwork roles that matter most to me.	0.46	**0.64**	−0.07	0.12
3	The time I spend in work and activities outside of work reflects my life priorities	0.56	**0.65**	−0.02	0.16
4	I spend enough time on important work and nonwork activities.	0.60	**0.68**	0.17	−0.05
5	Based on what matters most to me, I devote the right amount of my time to work and nonwork roles.	0.62	**0.90**	−0.14	−0.03
	Effectiveness balance				
6	I perform well in the life roles that I really value.	0.55	0.11	**0.69**	−0.04
7	I do well in roles that are my biggest priorities.	0.57	−0.25	**0.86**	0.06
8	I am able to effectively handle important work and nonwork responsibilities	0.55	0.24	**0.66**	−0.14
9	I am successful in work and nonwork roles that matter to me.	0.53	0.12	**0.44**	0.24
10	I perform well in my most highly valued work and nonwork roles.	0.54	0.03	**0.57**	0.17
	Affective balance				
11	I experience a lot of positive emotions in my most highly valued work and nonwork roles.	0.69	−0.17	−0.02	**0.95**
12	I am happy in the work and nonwork roles that are most important to me.	0.59	0.10	0.24	**0.49**
13	I am happy with the work and nonwork aspects of my life that are important to me.	0.58	0.09	0.02	**0.68**
14	I feel satisfied in the work and nonwork roles that are most important to me.	0.57	0.32	0.11	**0.40**
15	I am content with how things are going in the life roles that are my top priorities.	0.52	0.32	−0.02	**0.47**

Notes.
[a] Extracted communalites.

based on recommendations for minimum factor loadings with a sample of 200 participants (Hair et al., 2006). We also removed negatively valenced items that loaded on a separate factor and a single item that loaded on its own factor, resulting in a 3-factor solution. We further removed the lowest loading items to retain 5 items per factor (see Table 3). These final items were used in subsequent studies of convergent and discriminant validity of these measures.

4.4. EFA of global balance

EFA results with principal axis factoring and promax rotation on global balance items revealed a 1-factor structure reported in Table 2 (KMO = 0.86). Similar to the procedure detailed above, items with the lowest loadings were removed to retain 5 items. The 1-factor structure represented 59.4% of variance. Final items are listed in Table 4. These 5 items were used in subsequent studies to assess convergent and discriminant validity of these measures. Global balance was included in a separate EFA from the balance facets because global balance is not viewed as a fourth dimension of balance but rather, a global factor that influences how each facet of balance is perceived. In other words, global balance shares some variance with all facets. Therefore, we conducted separate exploratory factor analyses. We return to the relations among the dimensions and global balance in our next study.

5. Study 3 construct validation method & results

Following Hinkin (1998), in Study 3, we aimed to confirm the factor structure of our measure on an independent sample. In Study 2, we developed a 3-factor facet measure and a 1-factor global measure. In Study 3, we sought to explore the nature of the relationship between the global measure of balance and the three facets by specifying competing models. We also sought to a) examine whether the factor structure is invariant across gender and parental status, b) demonstrate convergent validity with other balance measures, c) examine discriminant validity from constructs with overlapping meanings that should be sufficiently distinct, and d) examine the predictive value of our measure to employee attachment and health outcomes, above established balance measures.

To demonstrate convergent validity, we examined the relationship between our global and dimension scales with two commonly used balance measures. Valcour (2007) developed a 5-item measure to assess satisfaction with fit and allocation of time and attention across work and family roles (balance satisfaction). Carlson et al. (2009) created six self-report items to assess the extent to which a person perceives they accomplish work and family expectations that are negotiated and shared with role partners (balance effectiveness). We expected our balance measure to relate to both extant measures.

In terms of discriminant validity, given that our balance facets reflect satisfaction, effectiveness, and involvement in work and nonwork roles commensurate with values, we wanted to ensure that our measures are distinct from domain-specific (work, family) satisfaction, effectiveness, and involvement. Moreover, given that global balance is an appraisal of one's combination of work and nonwork roles and how they fit together, it should be related to but distinct from life satisfaction, a broad attitude involving many life roles and activities (Diener et al., 1985). Specifically, we aimed to demonstrate whether (i) affective balance is distinct from job and family satisfaction, (ii) involvement balance is distinct from job and family involvement, (iii) effectiveness balance is distinct from job and family performance and (iv) global balance is distinct from life satisfaction.

A final aspect of construct validation is demonstrating that a new measure's observed relationships with other variables match theoretical predictions about how it should relate to those variables (Cronbach & Meehl, 1955). Thus, in our final step, we seek to demonstrate that our balance measure relates to outcomes as would be expected based on theory, that it does so above and beyond extant balance measures, and that it uniquely contributes to the literature. Role balance theory (Marks & MacDermid, 1996) states that when people experience balance among their role systems, they will experience more positive attitudes and a greater sense of well-being. Similarly, Voydanoff (2005) posited that balance fosters better role performance and quality. Thus, based on prior theories of balance, we expected that our measures of global balance and its facets would relate to greater self-reported organizational citizenship behaviors, engagement, organizational commitment, and health and less emotional exhaustion. We expect our balance measure to relate to these outcomes, above and beyond extant measures.

5.1. Participants and sample

Participants, adults employed full-time (31 h or more per week, IRS, 2020), were recruited using the Qualtrics Web-based survey platform. All participants were recruited by Qualtrics as part of a single, actively managed research panel. Before the launch of the full survey, a small set of participants completed the survey via a "soft launch." After reviewing the data from the soft launch, there appeared to be no problems with the survey, so it was sent to the full number of participants. There were 6 attention check items (e.g.,

Table 4
Final results of exploratory factor analysis of global balance – Study 2.

	Item	Communality	Factor loading
1	There is harmony in how I blend my work and nonwork roles	0.61	0.78
2	My work and nonwork roles are combined in ways that are harmonious	0.60	0.77
3	Overall, my work and nonwork roles are integrated	0.60	0.78
4	Overall, my work and nonwork roles fit together	0.46	0.67
5	All in all, my work and nonwork roles are in harmony	0.70	0.84

Table 5
Means, standard deviations, and correlations among study variables – Study 3.

		Mean (SD)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	Gender	1.57 (0.50)	–																			
2	Parental status	1.39 (0.49)	0.10*	–																		
3	Involvement balance	4.03 (0.83)	−0.09*	−0.09*	(0.91)																	
4	Effectiveness balance	4.30 (0.68)	−0.09*	−0.10*	0.71***	(0.92)																
5	Affective balance	4.00 (0.91)	−0.12**	−0.13**	0.69***	0.76***	(0.95)															
6	Global balance	3.59 (1.05)	−0.06	−0.12**	0.65***	0.58***	0.69***	(0.95)														
7	Balance (Carlson)	4.12 (0.80)	−0.05	−0.11**	0.71***	0.76***	0.74***	0.68***	(0.94)													
8	Balance (Valcour)	3.96 (0.92)	−0.10*	−0.09*	0.72***	0.67***	0.72***	0.76***	0.82***	(0.97)												
9	Job satisfaction	3.94 (1.12)	−0.14***	−0.13**	0.48***	0.52***	0.69***	0.62***	0.59***	0.59***	(0.96)											
10	Family satisfaction	4.37 (0.84)	−0.09*	−0.24***	0.39***	0.48***	0.53***	0.40***	0.51***	0.45***	0.40***	(0.94)										
11	Life satisfaction	3.64 (0.99)	−0.15***	−0.13**	0.50***	0.55***	0.73***	0.54***	0.59***	0.58***	0.59***	0.62***	(0.92)									
12	Job involvement	3.20 (0.92)	−0.15***	−0.10*	0.29***	0.33***	0.37***	0.41***	0.28***	0.30***	0.57***	0.15***	0.32***	(0.87)								
13	Family involvement	4.30 (0.78)	−0.06	−0.36***	0.33***	0.38***	0.39***	0.31***	0.42***	0.35***	0.29***	0.73***	0.47***	0.12**	(0.91)							
14	Job performance	4.59 (0.57)	−0.02	−0.03	0.37***	0.51***	0.39***	0.28***	0.53***	0.40***	0.41***	0.29***	0.32***	0.20***	0.25***	(0.89)						
15	Family performance	4.31 (0.76)	−0.06	−0.20***	0.55***	0.60***	0.57***	0.49***	0.68***	0.59***	0.41***	0.72***	0.54***	0.15***	0.68***	0.45***	(0.93)					
16	Org. commitment	3.06 (1.01)	−0.13**	−0.15***	0.33***	0.33***	0.48***	0.46***	0.37***	0.40***	0.62***	0.25***	0.41***	0.53***	0.22***	0.21***	0.23***	(0.89)				
17	Turnover intentions	2.23 (1.22)	0.11**	0.10*	−0.29***	−0.31***	−0.46***	−0.40***	−0.39***	−0.40***	−0.63***	−0.30***	−0.39***	−0.25***	−0.19***	−0.27***	−0.29***	−0.42***	(0.92)			
18	OCB	3.69 (0.73)	−0.11**	−0.14***	0.29***	0.41***	0.37***	0.36***	0.40***	0.35***	0.46***	0.33***	0.36***	0.38***	0.28***	0.34***	0.33***	0.52***	−0.26***	(0.86)		
19	Vigor	3.10 (1.00)	−0.21***	−0.13**	0.41***	0.44***	0.56***	0.52***	0.46***	0.49***	0.68***	0.30***	0.52***	0.60***	0.24***	0.32***	0.34***	0.65***	−0.36***	0.46***	(0.85)	
20	Emotional exhaustion	2.50 (1.15)	0.16***	0.05	−0.43***	−0.42***	−0.54***	−0.50***	−0.48***	−0.53***	−0.60***	−0.30***	−0.44***	−0.22***	−0.21***	−0.33***	−0.35***	−0.29***	0.67***	−0.24***	−0.47***	(0.93)
21	Health	3.77 (0.78)	−0.14***	−0.01	0.42***	0.50***	0.58***	0.47***	0.52***	0.53***	0.48***	0.44***	0.64***	0.29***	0.34***	0.31***	0.50***	0.32***	−0.33***	0.31***	0.51***	−0.45***

Note. Sample size ranges between 596 and 600 due to missing data. Numbers in diagonal represent scale reliability.

* $p < .05$.
** $p < .01$.
*** $p < .001$.

"Please respond with strongly disagree"), and respondents were omitted if they missed more than one. Participants were also omitted if they responded in less than 7.5 min where the average response time was 23 min, following recommendations from the literature (DeSimone & Harms, 2018; Hohne & Schlosser, 2018) and other researchers (Vaziri et al., 2020). There were 600 respondents who met these criteria with 56.5% female, 61.8% married/living with partner, 85.3% White, and 60.7% with children. The average age was 54 years. Participants worked an average of 42.25 h per week and spent 3.57 h on caregiving, 8.91 on household responsibilities, 10.60 h in leisure, 1.24 h in community or volunteer service, and 7.54 h in self-care. A variety of industries were represented with most being in education (17.3%) or transportation (18.5%) industries.

5.2. Measures

When responding, participants were instructed to think about the past month and respond based on a 5-point Likert scale (1-*Strongly disagree* to 5-*Strongly agree*) unless otherwise noted. All measures exceeded minimal standards of 0.70 for internal consistency (Nunnally, 1978), as can be seen in Table 5 which contains means, standard deviations, and intercorrelations.

5.2.1. Work-family balance
Balance effectiveness was measured with Carlson et al.'s (2009) 6-item scale (e.g., "I am able to accomplish the expectations that my supervisors and family have for me"; α = 0.94). Balance satisfaction was measured using Valcour's (2007) 5-item scale. Respondents reported how satisfied they were from (1) *extremely dissatisfied* to (5) *extremely satisfied* with items such as "the way you divide your time between work and personal or family life" (α = 0.97). Finally, we measured global balance using our 5 items and each of our facets using 5 items, following the same procedure as in Study 2.

5.2.2. Job and family satisfaction
Job satisfaction was measured with 3 items by Cammann et al. (1979) such as "All in all, I am satisfied with my job" (α = 0.96). The 3-item family satisfaction scale adapts these items to the family and has been used in other work–family research (Carlson et al., 2009; α = 0.94).

5.2.3. Job and family involvement
Job involvement was measured with 6 items from Frone et al. (1995) that were derived from Kanungo (1982) such as "The most important things that happen to me involve my work" (α = 0.87). These same items were adapted to measure family involvement (α = 0.91).

5.2.4. Job and family performance
Job performance was measured with 4 items from Williams and Anderson (1991), such as "I adequately complete assigned job duties" (α = 0.89). As done by others (Carlson et al., 2009), family performance was assessed by modifying these items to the family domain (α = 0.93).

5.2.5. Life satisfaction
We measured life satisfaction with 5 items from Diener et al. (1985) including "In general, my life is close to ideal" (α = 0.92).

5.2.6. Work emotional exhaustion
We measured the emotional exhaustion component of burnout using 5 items from Schaufeli et al. (1996) such as "I feel emotionally drained from my work" (α = 0.93).

5.2.7. Employee engagement
We used the shortened version of the Utrecht Engagement Scale (Schaufeli et al. (2006) with 3 items to assess vigor ("At my work, I feel bursting with energy"; α = 0.85).

5.2.8. Organizational commitment
Organizational commitment was measured with 6 items from Meyer and Allen (1991; e.g., "My current organization has a great deal of personal meaning for me"; α = 0.89).

5.2.9. Turnover intent
Turnover intent was measured with 4 items from Bentein et al. (2005) such as "All things considered, I would like to find a comparable job in a different organization," (α = 0.92).

5.2.10. Organizational citizenship behaviors
We used 6 items to measure OCBs. Respondents rated how frequently they engaged in each behavior (e.g., "offer ideas to improve the functioning of the organization; willingly give of your time to help others who have work or nonwork problems"; α = 0.86) from 1-*never* to 5-*very frequently*.

5.2.11. Health

Participants rated their physical, mental/emotional health, and sleep quality from 1-*Terrible* to 5-*Excellent*, and we averaged these three items for overall health ($\alpha = 0.79$).

To ensure discriminant validity of the various correlates of balance, we specified a measurement model for all these variables except balance and its facets (because those CFAs are already reported) and health (because the items are formative rather than reflective). The results suggested that the 12-factor model was a good fit to the data ($\chi^2(1311) = 3503.704$, $p < .001$, CFI = 0.918, RMSEA = 0.053), suggesting discriminant validity.

5.3. Results

5.3.1. Factor structure and measurement invariance

First, we examined the factor structure of facets of balance with CFA using Mplus Version 7.4 (Muthén & Muthén, 1998–2013). Three models were specified: (a) the hypothesized 3-factor structure, (b) 2-factor structure where items for affective and effectiveness balance were set to load on the same factor, and (c) 1-factor structure where all items loaded on the same factor. A comparative fit index (CFI) value of 0.90 or higher, Tucker–Lewis index (TLI) value of 0.90 or higher, a root mean square error of approximation (RMSEA) value of 0.08 or lower, and a standardized root mean squared residual (SRMR) value of 0.08 or lower indicate good model fit (Hu & Bentler, 1999). Fit indices for these models are reported in Table 6. Neither the 1-factor or 2-factor solution provided acceptable fit to the data (CFI and TLI less than 0.9, RMSEA and SRMR greater than 0.08). The 3-factor structure fit the data significantly better than the other models (a vs. b: $\Delta\chi^2(1) = 662.56$, $p < .001$; a vs. c: $\Delta\chi^2(3) = 1434.93$, $p < .001$). Furthermore, all item loadings were significant and greater than 0.6. Thus, this structure was retained. The reliability of balance facets were 0.91 for involvement, 0.92 for effectiveness, and 0.95 for affective balance.

We then sought to explore the nature of the relationship between our global measure of balance and the three facets. Our conceptualization of balance suggests that global balance is a higher order construct that shapes individual perceptions of each facet. Thus, we contrasted our conceptualization of global and facets of balance with other competing models. To do so, we tested three distinct measurement models including global balance and the three facets which are depicted in Fig. 1. These models were chosen as they reflect three viable structural models that capture the possible relation among the facets and global balance.

Model (a) represents a first-order structure with four factors, representing four types of balance (global, affective, effectiveness, and involvement balance). In this model, global balance was considered a first-order construct, just like the three facets of balance – all four constructs were parallel. We chose this model in that it challenges the notion that global balance is an overarching factor and treats it, instead, as a separate facet just like affective, involvement, and effectiveness balance. Model (b), on the other hand, represents a second-order CFA with the same four factors as in Model (a) (global balance, involvement, affective and effectiveness balance), but also specified an unmeasured higher-order balance construct (i.e., general balance factor) which was conceptualized as influencing perceptions on *all four* first-order constructs similarly. As in model (a), model (b) conceptualizes global balance as parallel to the three facets of balance rather than a general factor that influences perception of these facets. Model (b) differs from model (a) in that it specifies an unmeasured general factor of balance that drives responses on all four factors. Selection of this model would indicate that

Table 6
Fit statistics for alternative CFA models and measurement invariance – Study 3.

	χ^2 (df)	AIC	SSA-BIC	CFI	TLI	RMSEA	SRMR
CFA fit indices for facets of balance only							
Model 1A: 3-factor structure	**253.83 (87)*****	**15,723.19****	**15,781.53***	**0.980**	**0.975**	**0.057**	**0.027**
Model 1B: 2-factor structure	916.39 (88)***	16,383.75	16,440.88	0.899	0.879	0.126	0.057
Model 1C: 1-factor structure	1688.76 (90)***	17,152.11	17,206.81	0.804	0.772	0.173	0.074
CFA fit indices for facets of balance and global balance							
Model 2A: first-order CFA	464.56 (164)***	21,684.26	21,764.48	0.974	0.970	0.055	0.030
Model 2B: second-order CFA	502.63 (166)***	21,718.32	21,796.12	0.971	0.967	0.058	0.037
Model 2C: bi-factor (S-1) CFA	**418.64 (152)*****	**21,662.33**	**21,757.15**	**0.977**	**0.971**	**0.054**	**0.027**
Measurement invariance for bi-factor (S-1) model of balance - gender							
Configural invariance	647.67 (304)***	21,641.73	21,831.36	0.971	0.963	0.062	0.030
Metric invariance	693.28 (335)***	21,625.34	21,777.29	0.969	0.965	0.060	0.061
Scalar invariance	718.21 (351)***	21,618.28	21,750.77	0.968	0.966	0.059	0.063
Measurement invariance for bi-factor (S-1) model of balance – parental status							
Configural invariance	673.51 (304)***	21,659.47	21,849.09	0.968	0.960	0.064	0.031
Metric invariance	725.06 (335)***	21,649.02	21,800.96	0.967	0.962	0.063	0.051
Scalar invariance	755.86 (351)***	21,647.81	21,780.31	0.965	0.962	0.062	0.055

Note. AIC = Akaike's Information Criterion; SSA-BIC = Sample Size Adjusted Bayesian Information.
Criteria; CFI = Comparative Fit Index; TLI = Tucker–Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Squared Residual. Best fit models in **bold** text.
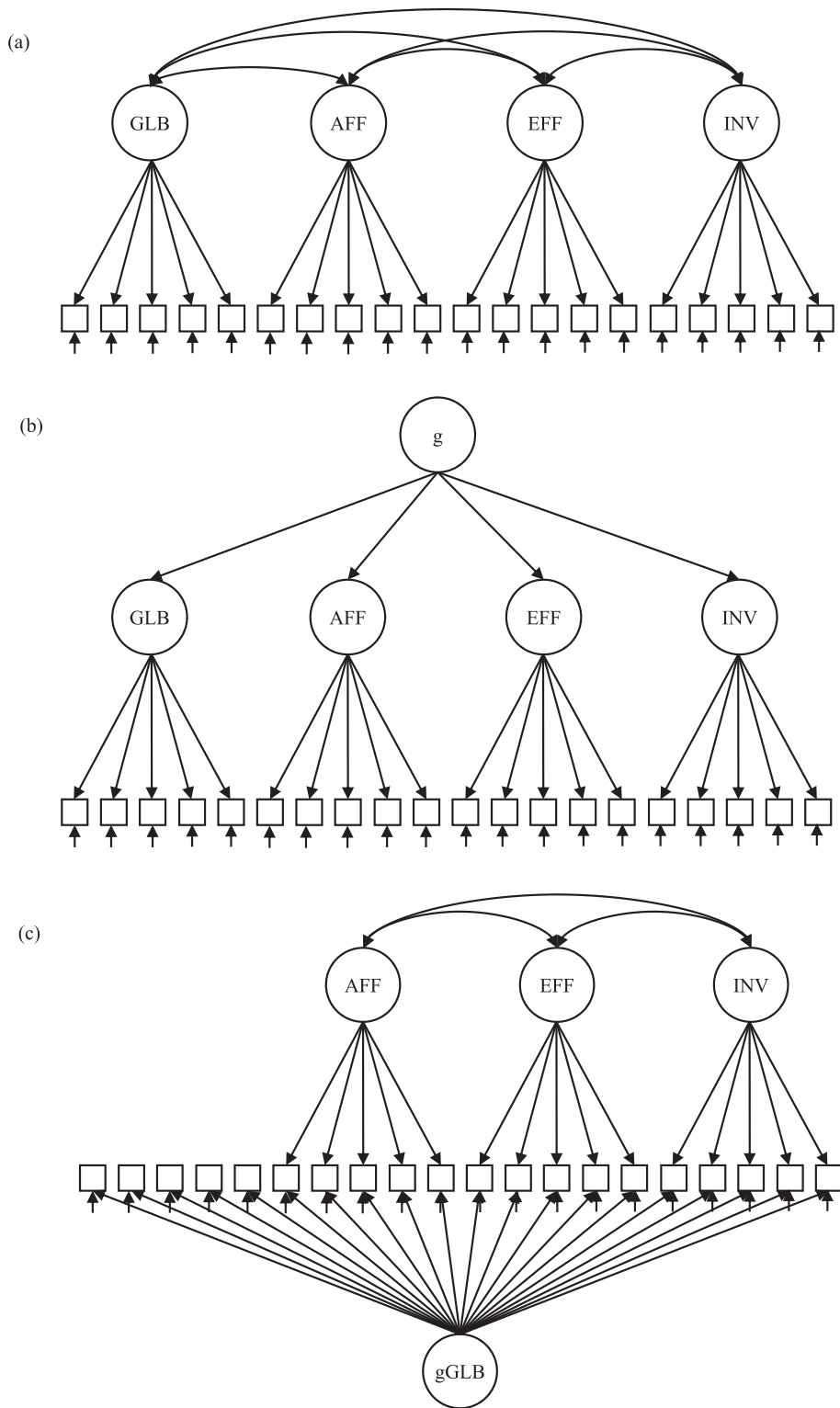* $p < .05$.
** $p < .01$.
*** $p < .001$.

Fig. 1. Visual representation of CFA Models examined. (a) First-order CFA; (b) Second-order CFA (g is unmeasured general balance factor); (c) Bi-factor (S-1) CFA. (gGLB is general balance factor measured by our global balance scale).

our global measure is not measuring the higher-order construct of balance that influences the facets, but is a construct that is distinct from but related to the three facets of balance. Finally, model (c) represents our conceptualization of the relationship between global balance and the three facets, which aligns with Casper et al. (2018). That is, model (c) depicts a general balance factor, as measured by our global balance items, along with specific facets (affective, effectiveness, and involvement balance), similar to models of general cognitive ability and specific abilities (Eid et al., 2017). In short, model (c) is similar to model (b) with the exception that the general balance factor is now measured with items from our global measure of balance. That is, model (b) suggested that our measured global balance is not the overarching factor that influences facets of balance but that there is another unmeasured general factor responsible for variance in the facets. In contrast, model (c) indicates that the items that we developed to assess global balance are, in fact, measuring the general overarching factor.

To test model (c), we conducted a bi-factor (S-1) model. In the bi-factor model, items for global balance only load on the general factor, whereas items for facets of balance load on *both* their respective factor and the general factor, representing the influence of the general factor on these facets (Eid et al., 2017; Zhang et al., 2020). Selection of this model would suggest that our measure of global balance captures the general factor of balance that influences responses people make to items capturing facets of balance, consistent with the conceptualization offered by Casper et al. (2018).

Table 6 reports fit statistics for these alternative CFA models. All three models exhibited acceptable fit to the data (CFI and TLI greater than 0.95, RMSEA and SRMR less than 0.08). As models are not nested, χ^2 difference test cannot be used to compare models, but absolute fit indices (e.g., AIC, sample size adjusted BIC) can identify best fit among non-nested models. Thus, model (b) exhibited a decrement in fit compared to the first-order CFA model (model a). On the other hand, the bi-factor model (model c) was a better fit to the data than model a or b, given both AIC and sample size adjusted BIC decreased slightly. Table 7 details CFA results for the bi-factor model. As indicated, all factor loadings on the general balance factor were relatively large (from 0.47 to 0.92 across all factors), as were factor loadings on the facets of balance (ranging from 0.45 to 0.70). In addition, consistency values (the proportion of variance in facets of balance accounted for by the general factor; Eid et al., 2017) were low to moderate (ranging between 0.08 and 0.60), whereas specificity values (the proportion of variance in facets of balance that is *not* accounted for by the general factor; Eid et al., 2017), were moderate to high ranging between 0.40 and 0.92. These results suggest that while a large proportion of the variance in items are accounted for by each facet, there was still a moderate level of variance accounted by the general factor. Altogether, results support model c in that balance is a multidimensional construct with our global balance items operating as the general factor at the apex, influencing facets of balance (i.e., involvement, effectiveness, and affective). Furthermore, each balance facet has its own variation not captured by the general factor as measured by our global balance items.

Table 7
Standardized factor loadings for bi-factor (S-1) confirmatory factor analysis of balance – Study 3.

	Item	Facet	G-factor loadings	S-factor loadings	Consistency	Specificity
1	There is harmony in how I blend my work and nonwork roles		0.92			
2	My work and nonwork roles are combined in ways that are harmonious		0.90			
3	Overall, my work and nonwork roles are integrated		0.79			
4	Overall, my work and nonwork roles fit together		0.89			
5	All in all, my work and nonwork roles are in harmony		0.91			
6	I am able to devote enough attention to important work and nonwork activities.	Inv	0.58	0.62	0.26	0.74
7	I am able to be adequately involved in the work and nonwork roles that matter most to me.	Inv	0.60	0.60	0.29	0.71
8	The time I spend in work and activities outside of work reflects my life priorities	Inv	0.47	0.45	0.18	0.82
9	I spend enough time on important work and nonwork activities.	Inv	0.59	0.64	0.25	0.75
10	Based on what matters most to me, I devote the right amount of my time to work and nonwork roles.	Inv	0.65	0.57	0.47	0.53
11	I perform well in the life roles that I really value.	Eff	0.50	0.66	0.09	0.91
12	I do well in roles that are my biggest priorities.	Eff	0.50	0.61	0.08	0.92
13	I am able to effectively handle important work and nonwork responsibilities	Eff	0.51	0.58	0.08	0.92
14	I am successful in work and nonwork roles that matter to me.	Eff	0.58	0.65	0.14	0.86
15	I perform well in my most highly valued work and nonwork roles.	Eff	0.53	0.70	0.10	0.90
16	I experience a lot of positive emotions in my most highly valued work and nonwork roles.	Aff	0.63	0.51	0.33	0.67
17	I am happy in the work and nonwork roles that are most important to me.	Aff	0.67	0.64	0.36	0.64
18	I am happy with the work and nonwork aspects of my life that are important to me.	Aff	0.65	0.68	0.36	0.64
19	I feel satisfied in the work and nonwork roles that are most important to me.	Aff	0.65	0.64	0.37	0.63
20	I am content with how things are going in the life roles that are my top priorities.	Aff	0.69	0.56	0.60	0.40

Note. $N = 595$. Standardized factor loadings are reported here. All factor loadings are significant at $p < .001$. Inv = Involvement Balance; Eff = Effectiveness Balance; Aff = Affective Balance; Consistency is the proportion of true-score variance in the non-reference domain (i.e., involvement, effectiveness, and affective balance) that is shared with the general factor (i.e., global balance). Specificity is the proportion of true-score variance in the non-reference domain that is *not* shared with the general factor.

Table 8
Results of structural equation modeling, predicting outcomes of balance – Study 3.

	Organizational commitment		Turnover intentions		Organizational citizenship behavior		Vigor		Emotional exhaustion		Health	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Gender[a]	−0.08*	−0.06	0.06	0.04	−0.08*	−0.07	−0.18***	−0.15***	0.12**	0.11**	−0.14***	−0.11**
Parental status[b]	−0.10**	−0.09*	0.08	0.06	−0.10*	−0.09*	−0.07	−0.05	−0.02	−0.04	0.06	0.08*
Balance (Carlson)	0.17*	0.03	−0.21*	−0.19	0.44***	0.24*	0.24**	0.03	−0.12	−0.01	0.37***	0.15
Balance (Valcour)	0.32***	0.03	−0.20*	−0.08	0.00	−0.02	0.31***	0.05	−0.45***	−0.25**	0.29***	0.20*
Involvement balance	–	−0.11*	–	0.15**	–	−0.17**	–	−0.08	–	0.03	–	−0.19***
Effectiveness balance	–	−0.06	–	0.13	–	0.28***	–	0.03	–	0.09	–	0.13
Affective balance	–	0.31***	–	−0.30***	–	−0.01	–	0.27***	–	−0.24***	–	0.31***
Global balance	–	0.49***	–	−0.20*	–	0.21*	–	0.50***	–	−0.33***	–	0.28***
R^2	0.24***	0.37***	0.17***	0.25***	0.21***	0.26***	0.32***	0.42***	0.32***	0.38***	0.42***	0.53***
ΔR^2		0.13***		0.08***		0.05***		0.10***		0.06***		0.11***
χ^2	359.54	1434.21	283.65	1253.24	379.97	1458.16	191.84	1151.53	241.75	1246.83	207.79	1162.31
df	148	668	115	595	148	668	100	560	131	631	100	560
CFI	0.979	0.966	0.982	0.970	0.975	0.964	0.989	0.972	0.989	0.973	0.987	0.972
TLI	0.976	0.963	0.979	0.967	0.971	0.960	0.987	0.969	0.987	0.970	0.985	0.968
RMSEA	0.049	0.044	0.050	0.043	0.051	0.045	0.039	0.042	0.038	0.041	0.042	0.042
SRMR	0.053	0.045	0.048	0.044	0.049	0.046	0.038	0.037	0.038	0.037	0.040	0.038

Note. Standardized effect sizes are reported. CFA = Comparative Fit Index; TLI = Tucker–Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Squared Residual.

[a] 1 = Male, 2 = Female.
[b] 1 = Yes, 2 = No.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

We also examined whether the bi-factor structure of balance is invariant across different groups following Vandenberg and Lance (2000) and Muthén and Asparouhov (2013). Tests of measurement invariance ensure that respondents from different groups have similar conceptions of a construct and perceive items and rating scales similarly. This is important to confirm that observed group differences indicate true differences rather than different interpretations of items or scales. As work-family relationships may vary by gender (Eby et al., 2005) and parental status (Wayne et al., 2013), theoretically, these factors may relate to how balance is perceived or rated, so we examined measurement invariance due to gender and parental status. Results are reported in Table 6. As indicated, the configural models of balance for both factors (i.e., gender, parental status) exhibited excellent fit to the data (CFI and TLI greater than 0.95, RMSEA and SRMR less than 0.08), suggesting that our balance measure exhibits configural invariance across groups. To establish metric invariance, the metric model should be compared to the configural model. To do this, we used ΔCFI to compare models instead of chi-squared difference test, because chi-square is too sensitive to sample size (Cheung & Rensvold, 2002). Cheung and Rensvold (2002) suggested that the model would be acceptable if the decrease in CFI is less than 0.01. Results suggested that the metric invariance model had acceptable fit and the change in CFI compared to the configural model was 0.002 and 0.001 for gender and parental status, respectively, suggesting metric invariance. Similarly, the scalar model had acceptable fit to the data and the change in CFI compared to the metric model was 0.001 and 0.003 for gender and parental status, respectively, suggesting scalar invariance.

5.3.2. Convergent validity

Correlational analyses indicated strong relationships among our measure of balance and extant balance satisfaction (Valcour, 2007) and effectiveness measures (Carlson et al., 2009). As shown in Table 5, the weakest relationship was between our measure of effectiveness balance and Valcour's balance satisfaction ($r = 0.67$) and the strongest relationship was between our measure of effectiveness balance and the balance effectiveness scale by Carlson et al. ($r = 0.76$; 2009) and our measure of global balance and Valcour's balance satisfaction ($r = 0.76$). Thus, our measure converges with other measures of the same construct as needed for evidence of construct validity.

5.3.3. Discriminant validity

Correlations provided initial evidence that our measures of affective, effectiveness, and involvement *balance across work and nonwork roles* are distinct from satisfaction, performance, and involvement *in work and family roles*. The correlations between affective balance and job ($r = 0.69$) and family ($r = 0.53$) satisfaction are below 0.70, suggesting they are distinct. To confirm, we conducted CFAs to demonstrate that a 3-factor model fits the data better than 1- or 2-factor models. The three factor model (affective balance, job and family satisfaction) was an excellent fit to the data ($\chi^2(41) = 100.6, p < .001$, CFI = 0.992, RMSEA = 0.049), and both the 2-factor and the 1-factor models exhibited a significant decrement in fit relative to the 3-factor (2-factor 1: $\Delta\chi^2(1) = 1222.72, p < .001$; 2-factor 2: $\Delta\chi^2(1) = 1201.99, p < .001$; 1-factor: $\Delta\chi^2(3) = 2424.50, p < .001$), suggesting discriminant validity.

Similarly, involvement balance was moderately related to job involvement ($r = 0.29$) and family involvement ($r = 0.33$). CFAs confirmed their distinctiveness. Specifically, while the 3-factor structure exhibited a good fit to the data ($\chi^2(101) = 642.13, p < .001$, CFI = 0.912, RMSEA = 0.095), neither of the 2-factor structures or the 1-factor structure exhibited adequate fit to the data and both were significantly worse than the 3-factor structure (2-factor 1: $\Delta\chi^2(1) = 1717.57, p < .001$; 2-factor 2: $\Delta\chi^2(1) = 1335.44, p < .001$; 1-factor: $\Delta\chi^2(3) = 3024.96, p < .001$).

Effectiveness balance was moderately related to self-reported job performance ($r = 0.51$) and family performance ($r = 0.60$). Again, CFAs confirmed their distinctiveness. While the 3-factor structure exhibited an excellent fit to the data ($\chi^2(62) = 115.97, p < .001$, CFI = 0.991, RMSEA = 0.038), neither of the 2-factor nor the 1-factor structures exhibited adequate fit to the data and fit significantly worse than the 3-factor structure (2-factor 1: $\Delta\chi^2(1) = 963.43, p < .001$; 2-factor 2: $\Delta\chi^2(1) = 1022.65, p < .001$; 1-factor: $\Delta\chi^2(3) = 1941.52, p < .001$).

Finally, the correlation between global balance and life satisfaction was $r = 0.54$ and CFAs confirmed that the 2-factor structure ($\chi^2(34) = 106.998, p < .001$, CFI = 0.987, RMSEA = 0.06) better fit the data compared to the 1-factor structure ($\Delta\chi^2(1) = 1558.88, p < .001$).

In sum, our measure of balance converges with extant measures of balance, but is sufficiently distinct from role-specific measures of satisfaction, effectiveness, and involvement as well as overall life satisfaction.

5.3.4. Theorized nomological network and incremental validity

Next, we conducted structural equation modeling to predict outcomes of balance. For each outcome (i.e., organizational commitment, turnover intentions, OCB, vigor, work emotional exhaustion, and health), we examined two models. The first model included balance satisfaction (Valcour, 2007) and balance effectiveness (Carlson et al., 2009) as predictors of outcomes while controlling for gender and parental status. In the second model, we added the bi-factor structure of balance developed in this paper. Comparing the two models allows us to examine 1) how well the newly developed measure predicts outcomes, and 2) whether our measure of balance predicts outcomes above and beyond extant balance measures (i.e., balance effectiveness and satisfaction). Results, (see Table 8) reveal that our measure predicts a 5–13% *increment* in variance in outcomes. Often, if one or both extant balance measures were significant in model 1, they were not significant in model 2 when our measure was added. The affective dimension of balance and global balance were consistently unique predictors of vigor, turnover intent, organizational commitment, emotional exhaustion, and health. Effectiveness balance also predicted OCBs. Involvement balance was significantly related to organizational commitment, turnover intentions, OCB, and overall health, but there was indication of suppression in that the sign was negative - opposite expectations and the bivariate correlations. Thus, results for the involvement balance dimension should be interpreted cautiously.

We then ran analyses with our measure in model 1 and added extant measures of balance (Carlson et al., 2009; Valcour, 2007) in model 2 (see Table 8) to determine whether established measures explained additional variance in outcomes above and beyond our measure. The change in R^2 was not significant for vigor or organizational commitment, suggesting that extant balance measures do not explain variance in these outcomes above and beyond our measure. Though the change was statistically significant, the magnitude of the change in R^2 was very small for turnover intent (0.01), burnout (0.01), OCB (0.01), and overall health (0.02), suggesting that the predictive ability of the extant balance measures above and beyond our new measure is very small.

In short, our measure contributed to the prediction of every outcome we measured, above and beyond established measures of balance. In contrast, established measures of balance predicted little, if any, additional variance in outcomes above and beyond our measure. These empirical findings are consistent with Casper et al.'s (2018) contention that extant measures do not fully capture the construct domain and suggests that our measure is more comprehensive.

6. Discussion

An important aspect of organizational research is accurate measurement of key constructs (Barrett, 1972); yet thorough validation efforts, particularly content validation, are noticeably absent, even in top journals (Colquitt et al., 2019). Construct validity concerns are especially problematic in the study of work-nonwork balance (Casper et al., 2018), a burgeoning topic that has significant practical importance. Because prior balance measures have not been rigorously developed to capture its complex and multi-faceted nature, a well-developed, validated, and widely accepted scale is sorely needed for theory and research to advance (Casper et al., 2018). Based on conceptual definitions and measurement strategies provided by Casper et al., we conducted three studies to develop, refine, and validate a 20-item scale with five items for global balance and for each facet- affective, effectiveness, and involvement balance. We demonstrate that balance, as measured by our scale, explains significant variance in employee attitudes and health that is not captured by extant measures. Ours is the first rigorously developed scale to fully capture the construct domain of balance as both a global and dimensional construct. As such, it has great potential to advance the scientific study of work-nonwork balance.

6.1. Contributions to theoretical understanding of balance

The scale developed in this research offers numerous advantages over existing balance scales and offers contributions to theoretical understanding of balance. First, as discussed in our literature review, prior theories and measures of balance have treated it as uni-dimensional. Yet, as noted by Wayne et al. (2017), work-family balance has multiple, distinct meanings and measures. Casper et al. (2018) posited balance as a construct with multiple dimensions (affective, effectiveness, and involvement balance) that can also be measured as a global construct, noting that misalignment between conceptual and operational definitions of balance is a common problem. Based on our comprehensive item development and content validation efforts, our measure is the first to fully map the construct domain by including global balance as well as its multiple dimensions. Subject matter experts agreed that the global and facet definitions represented the meaning of balance and that our retained items mapped onto these definitions. Our exploratory and confirmatory factor analyses demonstrated that each facet is a unique, discernable dimension of balance that is distinct from but related to global balance. As such, our measure has strong alignment with the conceptual definition of balance which enables stronger tests of theoretical models.

Our findings also indicate that current theories of balance are insufficient. That is, prior theories do not elaborate or explain balance as a multidimensional construct, so theories are needed that account for this fact. To our knowledge, thus far, only Wayne et al. (2017) has adapted theoretical explanations to account for the fact that balance satisfaction and balance effectiveness have unique relations to satisfaction and performance-based consequences. It is also likely that each dimension of balance has its own set of primary ante-cedents, and theories are needed to explain how and why this occurs. For example, perhaps factors that contribute to being adequately involved in valued work and nonwork roles (such as fewer time demands or greater mindfulness, schedule flexibility) are distinct from those that contribute to performing adequately across valued roles (such as conscientiousness, role ambiguity or situational con-straints, Gilboa et al., 2008) or being satisfied across them (such as positive affectivity, task identity or significance). The time is ripe for scholars to develop theory about the unique antecedents of each dimension of balance to guide future research.

Theoretically, each dimension should be differentiated given the compatibility principle (Ajzen & Fishbein, 1977) suggests the outcomes of each may differ, not only their antecedents. Researchers may choose the global balance measure when the theoretical, empirical, or practical emphasis is on overall fit between work and nonwork roles. Global evaluations of balance may relate to global evaluations of job, family, life, and health. Researchers can measure facets of balance to assess the extent to which positive emotions, effectiveness, or involvement across roles is commensurate with values. We would expect each facet to relate differentially to types of correlates. For example, affective balance may relate more strongly to affective correlates whereas effectiveness balance may relate more strongly to performance-based correlates (Wayne et al., 2017). In studies presented here, we show that our measure of balance is distinct from job and family satisfaction, involvement, and performance. Beyond this, though, consistent with the compatibility principle, our findings show that role-specific constructs, such as job or family satisfaction and performance, have higher correlations with the parallel content of balance, such as affective and effectiveness balance, respectively, than with another facet of balance with different content. For example, job, family, and life satisfaction have significantly higher correlations with affective balance than they do with involvement or effectiveness balance, whereas job and family performance have significantly higher correlations with effectiveness balance than with affective or involvement balance (Lee & Preacher, 2013). This suggests that correlates of each facet may differ in magnitude based on compatibility of content.

An exception to this, however, was the involvement dimension which did not have higher correlations with job and family

involvement than did affective or effectiveness balance (Lee & Preacher, 2013). This may be due to the measure of job and family involvement used in the current study (Kanungo, 1982). Examination of these items (e.g., "The most important things that happen to me involve my work") suggest they reflect the priority one places on and/or extent to which one values work or family roles. Our involvement balance items reference the extent to which people perceive they are adequately involved in terms of devoting the right amount *of time, energy, and attention* to their most important *work and nonwork roles*. Because the job and family involvement variables may capture how important a role is more than time and energy involvement in a role, these measures may not be as parallel in substance as is the case for role satisfaction and performance, resulting in different patterns with facets of balance.

Our research also provides insight into the nature of balance which can inform theory. Specifically, the bi-factor model of balance (Fig. 1c) was the best fitting model to the data. This would suggest that balance is a hierarchical construct with global balance operating as the general factor that directly influences facets of balance (i.e., involvement, effectiveness, and affective). Thus, the hierarchical construct of global balance is a general factor that accounts for substantial variance in facets of balance and operates at the apex of the balance hierarchy. Further, each facet would represent a specific type of balance that is not captured in the general or global form of balance. The results also suggest that effectiveness balance had the largest unique variance after controlling for global balance. This is evident from the higher factor loadings on the specific compared to general factor and the higher average specificity value for its items (average specificity for effectiveness balance was 0.90 compared to 0.71 and 0.60 for involvement and affective balance, respectively). However, results suggest that involvement and affective balance also tap unique variation beyond global balance and should be considered.

6.2. Contributions to research and practice

Our research also contributes to empirical research on balance. Prior balance measures have not been subjected to rigorous validation efforts. In fact, many measures use single items and/or items developed solely for use in that research without best practices in scale development (Casper et al., 2018). Ours is the first balance scale, to our knowledge, to demonstrate content, convergent, discriminant, and predictive validity, as considered essential for construct validation (Cronbach & Meehl, 1955). It has the potential, therefore, to remedy the construct validity concerns raised by Casper et al. (2018). Also, by using independent samples for each study, the potential for sample-specific bias has been reduced, and we have demonstrated scale invariance across gender and parental status. Given these advantages, it is our hope that researchers would adopt this scale so that research can accumulate in a meaningful way and advance understanding of the work-nonwork interface.

Finally, our measure has practical advantages which can facilitate insights with practical implications. Our measure predicts valued outcomes over two commonly used balance measures (Carlson et al., 2009; Valcour, 2007). Notably, our measure explains additional variance above and beyond extant measures, meaning that our measure contributes uniquely to predicting key outcomes. In contrast, extant measures generally did not predict incremental variance in outcomes above and beyond our measure. Our measure is comprehensive and sufficient to capture the construct space and predict valued outcomes. Importantly, global balance and affective balance, the two types of balance that SMEs indicated were most representative, were also the most consistent predictors of outcomes. Practically, our findings indicate that having greater overall fit among work and nonwork roles and feeling more positive emotions in valued roles is associated with greater work engagement, organizational commitment, health, and less burnout and turnover intentions, all desirable outcomes for organizations. Moreover, effectiveness balance was associated with better mental and physical health and sleep quality. Our measure may be useful to practitioners in that they could use the brief, 5-item measure of global balance to meaningfully predict employee attitudes, or could delve more deeply into facets of balance that might predict other outcomes such as well-being.

6.3. Limitations and future research directions

As with all research, ours should be considered in the context of its limitations. Though we used two heterogeneous, working adult samples across various industries, both samples are from the US. Cross-cultural investigations of our scale to assess invariance across cultures are needed. Also, though we examined differential relationships of our measure with theoretically relevant consequences, other correlates, especially antecedents, need to be considered. Personality and values are characteristics underexamined in the study of balance. For example, conscientiousness may provide the necessary planning and organizational skills to combine work and nonwork roles in harmonious ways. Extraversion, characterized by energy, positive affectivity, and sociability (Lucas et al., 2008), likely provides pleasant emotions, energy for involvement and performance for combining valued roles. Family and work centrality are also likely relevant to balance, particularly the facets which incorporate one's most valued work and nonwork roles.

Opportunity also exists for examining how unique situational antecedents might relate to different types of balance. For example, given the distinct facets of balance, theories pertaining to involvement, satisfaction, and effectiveness within roles may be applied to make predictions about unique factors that might promote each. For example, given their relevance to satisfaction and performance, motivation theories suggest that job or family design factors (e.g., task variety, significance; Hackman & Oldham, 1976) may relate to affective and effectiveness balance more strongly than to involvement balance. On the other hand, factors that limit role involvement such as structural demands (e.g., work hours, number and age of children, marital status, spouse employment status) might more strongly affect the extent to which involvement in work and nonwork roles fit values. We encourage researchers to develop theory to meaningfully test potential differences.

A limitation of our research is the use of a cross-sectional design. As such, future research should collect data over time to help demonstrate temporal ordering of relationships. We show empirical associations between our measure of balance and a host of

potential "outcomes" including job and family satisfaction, performance, life satisfaction, as well as employee attachment and health. Though we tentatively suggest that balance temporally precedes these outcomes, these relationships may be more complicated. That is, role satisfaction and performance may be cues that people use in determining their work-nonwork balance (Greenhaus & Allen, 2011), suggesting relationships may not be unidirectional but instead, form feedback loops over time. Now that there is a thoroughly developed measure, we urge scholars to ensure research designs allow them to test presumed temporal conclusions.

A final limitation in our data has to do with our involvement balance dimension. Our content validation process showed that SMEs had most agreement on the items fitting into this conceptual definition and felt this definition represented the meaning of balance. We also found that, at the bivariate level, involvement balance was associated with outcomes in the expected direction such that greater involvement balance was associated with more desirable outcomes. Despite this, in the analyses for each outcome, when all 4 of our balance scales and the two extant measures were included together, the unique relationship of involvement balance to outcomes was negative. It is unclear why this pattern emerged, but this is likely a suppressor effect which occurred due to multicollinearity between involvement balance and other measures. We encourage researchers to continue to incorporate this facet, particularly in longitudinal designs, to examine its predictive ability over time to determine whether this pattern persists or whether this was idiosyncratic to our data collection.

In conclusion, given the practical and empirical interest in work-nonwork balance, the identified construct validity problems (Casper et al., 2018), and the lack of a well-validated and widely accepted measure on which research can build, we provide such a measure of work-nonwork balance. Building on conceptual definitions provided by Casper et al. (2018), we developed items and demonstrated construct validity using best practices in scale development. We also demonstrated that it comprehensively and sufficiently predicts employee attachment and health and well-being. By focusing on the unique antecedents and outcomes of each aspect of balance (global, affective, involvement, and effectiveness), research can better inform the science and practice of work-nonwork balance.

CRediT authorship contribution statement

Julie Holliday Wayne- Conceptualization; Design of methodology; Writing of introduction, method, and discussion.
Hoda Vaziri- Conceptualization; Statistical analyses, reporting of results in text, tables, and figures.
Wendy Casper- Conceptualization; Writing - reviewing and editing; Data collection.

Declaration of competing interest

There are no conflicts of interests.

Acknowledgements

References

Aguinis, H., Villamor, I., & Ramani, R. S. (2020). MTurk research: Review and recommendations. *Journal of Management.* https://doi.org/10.1177/0149206320969787.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84,* 888–918.

Barrett, G. V. (1972). New research models of the future for industrial and organizational psychology. *Personnel Psychology, 25,* 1–17.

Bentein, K., Vandenberg, R., Vandenberghe, C., & Stinglhamber, F. (2005). The role of change in the relationship between commitment and turnover: A latent growth modeling approach. Journal of Applied Psychology, 90, 468–482. Adapted from Hom and Griffith (1991).

Cammann, C., Fichman, M., Jenkins, D., & Klesh, J. (1979). *The Michigan Organizational Assessment Questionnaire.* Ann Arbor: University of Michigan.

Carlson, D. S., Grzywacz, J. G., & Zivnuska, S. (2009). Is work-family balance more than conflict and enrichment? *Human Relations, 62,* 1459–1486. https://doi.org/10.1177/0018726709336500.

Carlson, D. S., Kacmar, K. M., & Williams, L. J. (2000). Construction and initial validation of a multi-dimensional measure of work–family conflict. *Journal of Vocational Behavior, 56,* 249–276. https://doi.org/10.1006/jvbe.1999.1713.

Casper, W. J., Vaziri, H., Wayne, J. H., DeHauw, S., & Greenhaus, J. (2018). The jingle-jangle of work–nonwork balance: A comprehensive and meta-analytic review of its meaning and measurement. *Journal of Applied Psychology, 103,* 182–214.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255.

Colquitt, J. A., Sabey, T. B., Rodell, J. B., & Hill, E. T. (2019). Content validation guidelines: Evaluation criteria for definitional correspondence and definitional distinctiveness. *Journal of Applied Psychology, 104*(10), 1243–1265. https://doi-org.go.libproxy.wakehealth.edu/10.1037/apl0000406.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

DeSimone, J. A., & Harms, P. D. (2018). Dirty data: The effects of screening respondents who provide low-quality data in survey research. *Journal of Business and Psychology, 33,* 559–577.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49,* 71–75.

Duxbury, L., & Higgins, C. (2001). *Work-life balance in the new millennium: Where are we? Where do we need to go?* CPRN: Ottawa.

Eby, L. T., Casper, W. J., Lockwood, A., Bordeaux, C., & Brinley, A. (2005). A twenty-year retrospective on work and family research in IO/OB: A review of the literature. *Journal of Vocational Behavior, 66,* 124–197.

Eid, M., Geiser, C., Koch, T., & Heene, M. (2017). Anomalous results in G-factor models: Explanations and alternatives. *Psychological Methods, 22*(3), 541–562. https://doi.org/10.1037/met0000083.

Frone, M. R. (2003). Work-family balance. In *In Handbook of occupational health psychology: 143–162.* Washington, DC: American Psychological Association.

Frone, M. R., Russell, M., & Cooper, M. L. (1995). Job stressors, job involvement and employee health: A test of identity theory. *Journal of Occupational and Organizational Psychology, 68,* 1–11.

Gilboa, S., Shirom, A., Fried, Y., & Cooper, C. (2008). A meta-analysis of work demand stressors and job performance: Examining main and moderating effects. *Personnel Psychology, 61*(2), 227–271. https://doi.org/10.1111/j.1744-6570.2008.00113.x.

Greenhaus, J. H., & Allen, T. D. (2011). Work-family balance: A review and extension of the literature. In *In Handbook of occupational health psychology (2nd ed.): 165–183*. Washington, DC: American Psychological Association.

Greenhaus, J. H., Collins, K. M., & Shaw, J. D. (2003). The relation between work-family balance and quality of life. *Journal of Vocational Behavior, 63*, 510–531.

Grzywacz, J. G., & Carlson, D. S. (2007). Conceptualizing work-family balance: Implications for practice and research. *Advances in Developing Human Resources, 9*, 455–471.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance, 16*, 250–279.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Pearson/Prentise Hall.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods, 7*, 191–205. https://doi.org/10.1177/1094428104263675.

Highhouse, S., & Zhang, D. (2015). The new fruit fly for applied psychological research. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 8*, 179–183. https://doi-org.go.libproxy.wakehealth.edu/10.1017/iop.2015.22.

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*, 967–988.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1*, 104–121.

Hohne, J. K., & Schlosser, S. (2018). Investigating the adequacy of response time outlier definitions in computer-based web surveys using paradata SurveyFocus. *Social Science Computer Review, 36*, 369–378.

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. https://doi.org/10.1080/10705519909540118.

Kanungo, R. N. (1982). Measurement of job and work involvement. *Journal of Applied Psychology, 67*, 341–349.

Lee, I. A., & Preacher, K. J. (2013, September). Calculation for the test of the difference between two dependent correlations with one variable in common. [Computer software]. Available from http://quantpsy.org.

Loo, R. (2002). A caveat on using single-item versus multiple-item scales. *Journal of Managerial Psychology, 17*, 68–75.

Lucas, R. E., Le, K., & Dyrenforth, P. S. (2008). Explaining the extraversion/positive affect relation: Sociability cannot account for extraverts' greater happiness. *Journal of Personality, 76*, 385–414. https://doi-org.go.libproxy.wakehealth.edu/10.1111/j.1467-6494.2008.00490.x.

Marks, S. R., & MacDermid, S. M. (1996). Multiple roles and the self: A theory of role balance. *Journal of Marriage and the Family, 58*, 417–432.

Meyer, J. P., & Allen, N. J. (1991). A three-component conceptualization of organizational commitment. *Human Resource Management Review, 1*, 61–89.

Muthén, B. O., & Asparouhov, T. (2013). New methods for the study of measurement invariance with many groups. Retrieved from http://www.statmodel2.com/download/PolAn.pdf.

Muthén, L. K., & Muthén, B. O. (1998–2013). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

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

Osgood, C. E., Suci, G., & Tannenbaum, P. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Pattusamy, M., & Jacob, J. (2016). Testing the mediation of work-family balance in the relationship between work-family conflict and job and family satisfaction. *South Africa Journal of Psychology, 46*, 218–231.

Schaufeli, W., Bakker, A. B., & Salanova, M. (2006). The measurement of work engagement with a short questionnaire: A cross-national study. *Educational and Psychological Measurement, 66*, 701–716.

Schaufeli, W. B., Leiter, M. P., Maslach, C., & Jackson, S. E. (1996). The Maslach Burnout Inventory general survey. In C. Maslach, S. E. Jackson, & M. P. Leiter (Eds.), *MBI manual* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press. ADAPTED.

Schriesheim, C. A., & Hinkin, T. R. (1990). Influence tactics used by subordinates: A theoretical and empirical analysis and refinement of the Kipnis, Schmidt, and Wilkinson subscales. *Journal of Applied Psychology, 75*, 246–257.

Schriesheim, C. A., Powers, K. J., Scandura, T. A., Gardiner, C. C., & Lankau, M. J. (1993). Improving construct measurement in management research: Comments and a quantitative approach for assessing the theoretical content adequacy of paper-and-pencil survey-type instruments. *Journal of Management, 19*, 385–417. https://doi.org/10.1016/0149-2063(93)90058-U.

Valcour, M. (2007). Work-based resources as moderators of the relationship between work hours and satisfaction with work-family balance. *Journal of Applied Psychology, 92*, 1512–1523.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4–70.

Vaziri, H., Casper, W. J., Wayne, J. H., & Matthews, R. A. (2020). Changes to the work–family interface during the COVID-19 pandemic: Examining predictors and implications using latent transition analysis. *Journal of Applied Psychology, 105*(10), 1073–1087. https://doi.org/10.1037/apl0000819.

Voydanoff, P. (2002). Linkages between the work-family interface and work, family, and individual outcomes: An integrative model. *Journal of Family Issues, 23*, 138–164.

Voydanoff, P. (2005). Toward a conceptualization of perceived work-family fit and balance: A demands and resources approach. *Journal of Marriage and Family, 67*, 822–836.

Wayne, J. H., Butts, M. B., Casper, W. J., & Allen, T. D. (2017). In search of balance: A conceptual and empirical integration of multiple meanings of work-family balance. *Personnel Psychology, 70*, 167–210.

Wayne, J. H., Casper, W. J., Mathews, R., & Allen, T. D. (2013). Family-supportive organization perceptions and organizational commitment: The mediating role of work-family conflict and enrichment and partner attitudes. *Journal of Applied Psychology, 98*, 606–622.

Williams, L. J., & Anderson, S. E. (1991). Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. *Journal of Management, 17*, 601–617.

Zhang, B., Sun, T., Cao, M., & Drasgow, F. (2020). *Using bifactor models to examine the predictive validity of hierarchical*. Constructs: Pros, Cons, and Solutions. Organizational Research Methods. https://doi.org/10.1177/1094428120915522.