POSTER

TITLE

Engagement vs. attitude: Measurement invariance across item orderings

SHORTENED TITLE

Engagement vs. attitude

 ABSTRACT

The most common current definition of attitudinal engagement implicates sub-dimensions of vigor, dedication, and absorption. The tripartite model of attitudes specifies cognitive, behavioral, and affective components. The current study investigates measurement invariance across a new measure of engagement that intentionally crosses the substantive and attitudinal components. Through counterbalancing the order of item presentation, we provide response cues regarding structural priority. Analyses of measurement invariance reveal generally positive support for invariance across item orderings.

2,657 words

The roots of employee (aka work; e.g., Schaufeli & Bakker, 2010) engagement research likely started with theoretical expansions of forms of employee participation (see, for example, Ferris & Hellier, 1984) and job involvement (e.g., Elloy, Everett, & Flynn, 1991). This exploration extended into broader considerations of attitudes and emotions (Staw, Sutton, & Pelled, 1994) and were informed by further exploration of the dimensionality of constructs such as organizational commitment (Meyer & Allen, 1991). The 1990’s saw focused development and refinement (for example, a dissertation; Leone [1995] or actual semantic reference; Kahn [1990]). Staw et al. (1994) investigated the relationships between *positive emotions* and favorable work outcomes, and although they do not use the word, “engagement”, their distinction between felt and expressed emotion likely held influence upon the burgeoning interest in the engagement construct.

**Employee engagement**

Although occasionally referred to as residing on the opposing pole to *burnout* (Maslach & Leiter, 2008), these two constructs are currently most commonly conceptualized as being distinct (Goering, Shimazu, Zhou, Wada, & Sakai, 2017; Kim, Shin, & Swanger, 2009; Schaufeli, Taris, & Van Rhenen, 2008; Timms, Brough, & Graham, 2012), although certainly not universally (Cole, Walter, Bedeian, & O’Boyle, 2012; Taris, Ybema, & Beek, 2017). Comparing the two, Goering et al. (2017) concluded that they have a moderate (negative) association, but also distinct nomological networks. Schaufeli et al. (2008) investigated both internal and external association indicators, concluding that engagement and burnout (as well as *workaholism*) should be considered three distinct constructs.

Burnout can be defined as a psychological syndrome characterized by exhaustion (low energy), cynicism (low involvement), and inefficacy (low self-efficacy), which is experienced in response to chronic job stressors (e.g., Leiter & Maslach, 2004; Maslach & Leiter, 1997). Alternatively, engagement refers to an individual worker’s involvement and satisfaction as well as enthusiasm for work (Harter, Schmidt, & Hayes, 2002). Schaufeli and Bakker (2003) further specify a “positive, fulfilling, work-related state of mind that is characterized by vigor, dedication, and absorption” (p. 74). Via their conceptualization, vigor is described as high levels of energy and mental resilience while working. Dedication refers to being strongly involved in one’s work and experiencing a sense of significance, enthusiasm, inspiration, pride, and challenge. Absorption is characterized by being fully concentrated and happily engrossed in one’s work, whereby time passes quickly and one has difficulties with detaching oneself from work (Schaufeli, Salanova, González-Romá, & Bakker, 2002). The dimension of absorption has been noted as being influenced in conceptual specification by (Csikszentmihalyi, 1990)’s concept of “flow”.

**Measures of engagement.** Like many other constructs within the psychology literature, our knowledge of engagement has been at least partially informed by its measurement and documented association with other work attitudes and behaviors. There are currently many engagement scales used for either academic or applied purposes (and sometimes both). In addition to the aforementioned Gallup Q12, popular measures include the Utrecht Work Engagement Scale (UWES; Schaufeli et al., 2002), the ISA engagement scale (Phuangthuean et al., 2018), the Saks (2006; 2019) scale, and May et al. (2004) scale. These scales differ in the number and content of dimensions. For example, the ISA divides engagement into intellectual, social, and affective components. Saks specifies two scales. The Utrecht Work Engagement Scale (UWES) has been subject to criticism focused on the three subscales (vigor, dedication, and absorption) being so highly correlated that it may be argued that a unidimensional structure may be a better structural representation than the three-factor model (Kulikowski, 2017; Willmer et al., 2019).

We retain the vigor-dedication-absorption structure put forth in Schaufeli et al. (2002) and also specified in the UWES as the foundation for our construct definition. This model is not without criticism, however. Some researchers have questioned its structural validity by pointing out that vigor, dedication and absorption all correlate highly with each other (Kulikowski, 2017). We believe that some of the inter-construct association may be able to be accounted for when attitudinal components are additionally and simultaneously recognized. That is, it is possible that the constructs have distinction but the most commonly leveraged measure does not. The current study focuses on a *new* measure of engagement that crosses the substantive construct dimensions of vigor, dedication, and absorption with the attitudinal components of affect, cognition, and behavior.

**Engagement as an attitude.** The first (to our knowledge) use of the word “engagement” as a construct of organizational relevance came with Kahn (1990), who defined engagement as: “the harnessing of organization members’ selves to their work roles; in engagement, people employ and express themselves physically, cognitively, and emotionally during role performances.” Although this definition was quickly outpaced by subsequent respecifications (see, for example, Baumruk, 2004 and Shaw, 2005), Kahn (1990)’s definition is notable in that it conforms to the then-ascendant tripartite model of attitudes proposed by Rosenberg (1960). This model frames attitudes as latent variables that manifest cognitively, affectively and behaviorally. Even though it is not specifically a work engagement model, the tripartite model has helped researchers define and deconstruct attitudes to gain a better understanding of individuals’ reactions towards specific attitude objects (Kaiser and Wilson, 2019).

**Response cues in Psychological assessment**

Assessment response cues in general and order effects in particular have their origins in the early days of Cognitive Psychology. Primacy and recency (whether an item is presented at the beginning or end of a list) are known to elicit differences in response (e.g., Krosnick & Alwin, 1987). This effect has also been noted in methodological contexts in the form of differential carryover effects. For example, the order of presentation of samples in a product taste test (see, for example, Dean, 1980). Ackerman, Spray, Reckase, and Carlson (1989) found only small differences in response patterns when presenting *test* items in a fixed versus random ordering. Mashburn, Meyer, Allen, and Pianta (2014) experimentally controlled the presentation of rated material, finding higher indices of reliability and validity of ratings when content was administered randomly (e.g., order effects were controlled for).

Knowles (1988) and Hamilton and Shuminsky (1990) both administered exhaustively crossed orderings of items, noting better item discrimination (e.g., corrected item-total correlations) *later* in the assessment, regardless of actual item content. Steinberg (1994) provides a description of this effect: “…literature converges on the view that responding repeatedly to items representing a single, unidimensional psychological construct increases the accessibility of relevant beliefs or feelings, which in turn, increases the relation between the item response and the underlying construct… …This attentional focus influences item responses through such processes as item interpretation and ease of retrieval of relevant feelings that are applied to the item” (p. 341). Across investigations, item order effects seem to have a small but consistent effect on psychometric indicators of response quality (see also Steinberg, 1994 and Weinberg et al., 2018).

**Current Study**

Our presentation investigates the impact of item ordering on factor structure, and we do so from the unique perspective of an instrument that is intentionally saturated with two different structural components: 1) an attitudinal structure, and 2) a substantive construct structure. This structure permits the strategic ordering of items to respondents such that one structure may be emphasized over the other.

Regarding the substantive structure, this measure operationalizes work engagement as a mental state wherein employees: a) feel energized (*Vigor*), b) are enthusiastic about the content of their work and the things they do (*Dedication*), and c) are so immersed in their work activities that time seems compressed (*Absorption*). The attitudinal structure is comprised of three components: d) feeling (e.g., affect), e) thought (e.g., cognition), and f) action (e.g., behavior). Development and construct validation of this 18-item measure of engagement is described in Russell et al. (2022) whereas the current study focuses on administrative response cues in the form of order of item presentation. The expectation is that either model (attitudinal or substantive) will exhibit stronger factorial validity when item administration parallels latent structure.

# 1 Methods

## 1.1 Participants

Data was obtained from two sources. In the first sampling, 282 individuals responded to a snowball sampling initiated by Industrial and Organizational Psychology faculty and graduate students. There were four counterbalanced orderings of item presentations within this administration, as well as an additional 18 contextual items - this sample constituted the original scale development sample, and at the time of administration the additional contextual items were candidates for item retention. In the second data collection initiative, Qualtrics panels were solicited. These US workforce representative adults responded to two counterbalanced orderings of the focal 18 items along with 2 additional contextual items. These Qualtrics respondents included 343 who responded to attitudinally clustered items and 404 who responded to substantively clustered items.

## 1.2 Materials

Our 18-item engagement measure was crafted to be intentionally complex (each item is intended to load on two constructs). This complexity, however, derives from a crossing of the attitudinal components of affect, cognition, and behavior with the substantive engagement components of vigor, dedication, and absorption. The 6-point response scale is: *Strongly Disagree, Disagree, Somewhat Disagree, Somewhat Agree, Agree, Strongly Agree*. The item stems as well as their intended scale associations are presented in Table 1.

# Results

We used R (Version 4.0.3; R Core Team, 2021) and the R-packages *apaTables* (Version 2.0.8; Stanley, 2021), *dplyr* (Version 1.0.2; Wickham et al., 2021), *DT* (Version 0.16; Xie, Cheng, & Tan, 2021), *forcats* (Version 0.5.0; Wickham, 2021a), *ggplot2* (Version 3.3.2; Wickham, 2016), *labourR* (Version 1.0.0; Kouretsis, Bampouris, Morfiris, & Papageorgiou, 2020), *lavaan* (Version 0.6.8; Rosseel, 2012), *magrittr* (Version 2.0.1; Bache & Wickham, 2020), *papaja* (Version 0.1.0.9997; Aust & Barth, 2020), *purrr* (Version 0.3.4; Henry & Wickham, 2020), *readr* (Version 1.4.0; Wickham & Hester, 2020), *sem* (Version 3.1.11; Fox, Nie, & Byrnes, 2020; Epskamp, 2019), *semPlot* (Version 1.1.2; Epskamp, 2019), *stringr* (Version 1.4.0; Wickham, 2019), *tibble* (Version 3.1.0; Müller & Wickham, 2021), *tidyr* (Version 1.1.2; Wickham, 2021b), and *tidyverse* (Version 1.3.0; Wickham, Averick, et al., 2019) for all our analyses.

Two omnibus confirmatory factor analyses (CFAs) were imposed on the data conforming to the two scale definitions reported in Table 1. Latent covariances were also freed (e.g., associations across latent constructs were estimated). Regardless of item ordering, across 1,025 respondents, both models showed fair fit (=995.34, *df*=132, *RMSEA*=0.09; =1,103.47, *df*=132, *RMSEA*=0.09). Additional fit indices for the two models (as well as models run separately within each condition) are presented in Table 2. Figures 1 and 2 present the omnibus models visually (standardized coefficients displayed). Note that the primary source of misfit for both models within the omnibus analyses is item 14, which is the lone reverse-coded item within the inventory, “Thinking about work saps my energy”. Table 3 presents unit-weighted scale scores across the three substantive and three attitudinal scales. Omnibus data ’s (across all conditions) were 0.81 (Absorption), 0.91 (Dedication), 0.78 (Vigor), 0.78 (Affect), 0.89 (Cognition), and 0.83 (Behavior).

### 2.0.1 Condition effects.

The order of item presentation was: 1) random within attitudinal dimension, 2) random within substantive dimension, 3) parcels of attitudinal within substantive (36-item attitudinal context), 4) parcels of substantive within attitudinal (36-item substantive context), 5) parcels of substantive within attitudinal (20-item attitudinal context), and 6) parcels of attitudinal within substantive (20-item substantive context). For example, in condition 1, the first items presented were all associated with one attitudinal dimension (for example, “Affect”). Once the Affect item list was fully exhausted, the respondent was then administered the full set of Behavioral items, and once these were completed the respondent was then administered the Cognitive item set[[1]](#footnote-1). We view these orderings as cues regarding factor structure, and anticipated empirical factor structures to reflect these cues. The effects did emerge in the anticipated direction in all conditions, but were quite moderate (for example, = 9.55, = 10.53). Given the variety of item orderings administered, this should be considered somewhat comforting regarding the effect of contextual embeddedness within multidimensional inventories. Yes there is consistently better fit with item orderings conforming to factor structure, however, the magnitude of those differences is quite small. To further explore degree of similarity, we applied explicit tests of measurement invariance.

#### 2.0.1.1 Measurement invariance.

We investigated measurement invariance across a series of progressively greater restrictive conditions (Meredith, 1993; Steinmetz et al., 2009). First, we constrained latent variable means (this is a test of “configural” invariance), then further restricted factor loadings (this is our “weak” invariance test). For our test of “strong” invariance, we constrained item intercepts to be equal across conditions, and, as a most extreme restraint we finally restricted item residual variances to be equal across administration conditions (this is referred to as a test of “strict” invariance). To evaluate the extent to which invariance is plausible, chi-square difference tests are applied across restraints (for example, a chi-square difference from configural to weak), with non-significant chi-square changes indicating plausibility for the more restrictive model. That is, a significant change in fit implicates non-invariance (in our case, an item-ordering effect).

Because our six conditions were realized across two qualitatively different sampling procedures, we apply our analyses of measurement invariance twice - once within our initial snowball sampling and then secondly extending to the follow-up Qualtrics panel respondents. Because of the different sample sizes, our chi-square comparisons would not be directly interpretable if we were to evaluate measurement invariance across all six item ordering conditions. Thus, we applied invariance constraints separately across four snowball sampling conditions and two Qualtrics panel conditions. Stated differently, we applied tests of invariance to: our snowball sample only (4 item ordering conditions) and Qualtrics sample only (2 item ordering conditions).

The initial snowball sample across four counterbalanced conditions can be reasonably characterized as exhibiting strong/scalar invariance for the attitudinal structure and marginally attaining scalar invariance for the substantive CFA (see Tables 4 and 5). Tables 6 and 7 provide evidence of weak/metric invariance for the substantive CFA across the two Qualtrics item orderings (ΔΧ2 = 13.84, *df* = 15, *p* = .54), but a lack of invariance for the attitudinal model.

# Discussion

When attending to individual model fit indices within item presentation order conditions, we did realize consistently better fit in the direction of the item consistency cue. When items are presented “together” in terms of conceptual cohesion (either grouped by attitudinal dimension or substantive dimension), the CFAs yielded slightly better fit indicators than when items were not presented in conceptually cohesive groupings. This was true across different levels of cohesive emphasis (e.g., number of cohesive items, randomization or static sequencing of the alternative model indicators). Exploring the possibility of measurement invariance gave somewhat mixed results. Certainly there is evidence of at least partial invariance within both administrations*.* However, the idea of invariance has greater support within our initial four administration conditions than it does within our Qualtrics panel conditions. It is worth noting that these two sampling procedures may be qualitatively distinct enough to question whether respondent effects may account for some of these findings

Practically, these findings should provide some comfort to assessment specialists. The practical implication is that “item ordering perhaps does not matter so much”. We were in a unique circumstance regarding the opportunity to manipulate different meaningful groupings because our focal measure is intentionally complex, affording the luxury of differentially stressing different structures. The item ordering literature does fairly consistently claim better reliability of assessment scores later in a sequence of similar items, and this was not evaluated in the current study, but could be the focus of future investigations. Additionally, it is possible that ordering effects would be more prominent within longer scales. We attempted to account for this possibility by administering different numbers of “contextual” but unscored content-cohesive companion items, although our longest instrument was yet only 36 items long.

It is certainly feasible that the content cue becomes more powerful as respondents become more fatigued (e.g., with truly longer assessments). In a similar consideration, it is possible that our scales were SO short that the respondents did not pick up on the cue (“hey – these are all asking about how I *feel* [or *think*, or *act*]”). Longer scales would represent a stronger manipulation in the sense that there is greater opportunity for the cue to be perceived. In addition to questions of item order effects, the current investigation additionally provides some evidence of internal factorial validity to this newly developed measure of engagement, and it is hoped that others further investigate the psychometric properties of this unique engagement measure.

**References**

Ackerman, T. A., Spray, J. A., Reckase, M. D., & Carlson, J. E. (1989). *A comparison of the effects of random versus fixed order of item presentation via the computer*. AMERICAN COLL TESTING PROGRAM IOWA CITY IA.

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown*. Retrieved from <https://github.com/crsh/papaja>

Bache, S. M., & Wickham, H. (2020). *Magrittr: A forward-pipe operator for r*. Retrieved from <https://CRAN.R-project.org/package=magrittr>

Baumruk, R. (2004). The missing link: The role of employee engagement in business success, *47*, 48–52.

Coffman, C., & Harter, J. (1999). A hard look at soft numbers. *Position Paper, Gallup Organization*.

Cole, M. S., Walter, F., Bedeian, A. G., & O’Boyle, E. H. (2012). Job burnout and employee engagement: A meta-analytic examination of construct proliferation. *Journal of Management*, *38*(5), 1550–1581.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience* (Vol. 1990). Harper & Row New York.

Dean, M. L. (1980). Presentation order effects in product taste tests. *The Journal of Psychology*, *105*(1), 107–110.

Elloy, D. F., Everett, J. E., & Flynn, W. R. (1991). An examination of the correlates of job involvement. *Group & Organization Studies*, *16*(2), 160–177. <https://doi.org/10.1177/105960119101600204>

Epskamp, S. (2019). *SemPlot: Path diagrams and visual analysis of various sem packages’ output*. Retrieved from <https://CRAN.R-project.org/package=semPlot>

Ferris, R., & Hellier, P. (1984). Added value productivity schemes and employee participation. *Asia Pacific Journal of Human Resources*, *22*(4), 35–44. <https://doi.org/10.1177/103841118402200406>

Fox, J., Nie, Z., & Byrnes, J. (2020). *Sem: Structural equation models*. Retrieved from <https://CRAN.R-project.org/package=sem>

Goering, D. D., Shimazu, A., Zhou, F., Wada, T., & Sakai, R. (2017). Not if, but how they differ: A meta-analytic test of the nomological networks of burnout and engagement. *Burnout Research*, *5*, 21–34.

Hamilton, J. C., & Shuminsky, T. R. (1990). Self-awareness mediates the relationship between serial position and item reliability. *Journal of Personality and Social Psychology*, *59*(6), 1301.

Harter, J. K., Schmidt, F., & Hayes, T. (2002). Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis. *The Journal of Applied Psychology*. <https://doi.org/10.1037/0021-9010.87.2.268>

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. Retrieved from <https://CRAN.R-project.org/package=purrr>

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, *33*(4), 692–724.

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, *33*(4), 692–724.

Kaiser, F. G., & Wilson, M. (2019). The Campbell Paradigm as a Behavior-Predictive Reinterpretation of the Classical Tripartite Model of Attitudes. *European Psychologist*, *24*(4), 359–374. <https://doi.org/10.1027/1016-9040/a000364>

Kim, H. J., Shin, K. H., & Swanger, N. (2009). Burnout and engagement: A comparative analysis using the Big Five personality dimensions. *International Journal of Hospitality Management*, *28*(1), 96–104. <https://doi.org/10.1016/j.ijhm.2008.06.001>

Knowles, E. S. (1988). Item context effects on personality scales: Measuring changes the measure. *Journal of Personality and Social Psychology*, *55*(2), 312.

Kouretsis, A., Bampouris, A., Morfiris, P., & Papageorgiou, K. (2020). *LabourR: Classify multilingual labour market free-text to standardized hierarchical occupations*. Retrieved from <https://CRAN.R-project.org/package=labourR>

Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly*, *51*(2), 201–219.

Kulikowski, K. (2017). Do we all agree on how to measure work engagement? Factorial validity of utrecht work engagement scale as a standard measurement tool–a literature review. *International Journal of Occupational Medicine and Environmental Health*, *30*(2), 161–175.

Leiter, M., & Maslach, C. (2004). Areas of worklife: A structured approach to organizational predictors of job burnout. In *Research in occupational stress and well-being* (Vol. 3, pp. 91–134). <https://doi.org/10.1016/S1479-3555(03)03003-8>

Leone, D. R. (1995). *The relation of work climate, higher order need satisfaction, need salience, and causality orientations to work engagement, psychological adjustment, and job satisfaction* (PhD thesis). ProQuest Information & Learning.

Mashburn, A. J., Meyer, J. P., Allen, J. P., & Pianta, R. C. (2014). The effect of observation length and presentation order on the reliability and validity of an observational measure of teaching quality. *Educational and Psychological Measurement*, *74*(3), 400–422.

Maslach, C., & Leiter, M. (1997). What causes burnout. *Maslach C, Leiter MP. The Truth About Burnout: How Organizations Cause Personal Stress and What to Do About It. San Francisco, CA: Josey-Bass Publishers*, 38–60.

Maslach, C., & Leiter, M. P. (2008). Early predictors of job burnout and engagement. *Journal of Applied Psychology*, *93*(3), 498–512.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543.

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

Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames*. Retrieved from <https://CRAN.R-project.org/package=tibble>

Phuangthuean, P., Kulachai, W., Benchakhan, K., Borriraksuntikul, T., & Homyamyen, P. (2018). Employee engagement: validating the ISA engagement scale. In *Conference of the International Journal of Arts and Sciences* (Vol. 11, No. 1, pp. 99-108).

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

Rosenberg, M. J. (1960). *Cognitive, affective, and behavioral components of attitudes*. *Attitude organization and change*.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved from <https://www.jstatsoft.org/v48/i02/>

Russell, M., Ossorio Duffoo, C., Garcia Prieto Palacios Roji, R., & Kulas, J. (2022). Development of an intentional bifactor measure of engagement. In *The seattle edition of siop* (pp. 1–14). SIOP.

Saks, A. M. (2006). Antecedents and consequences of employee engagement. *Journal of Managerial Psychology*, *21*(7), 600–619.

Saks, A. M. (2019). Antecedents and consequences of employee engagement revisited. *Journal of Organizational Effectiveness: People and Performance*, *6*(1), 19–38.

Schaufeli, W., & Bakker, A. (2010). The conceptualization and measurement of work engagement. In W. Schaufeli, A. Bakker, & M. Leiter (Eds.), *Work engagement: A handbook of essential theory and research* (pp. 10–24). New York: Psychology Press.

Schaufeli, W. B., & Bakker, A. B. (2003). UWES–utrecht work engagement scale: Test manual. *Unpublished Manuscript: Department of Psychology, Utrecht University*, *8*.

Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies*, *3*(1), 71–92.

Schaufeli, W. B., Taris, T. W., & Van Rhenen, W. (2008). Workaholism, burnout, and work engagement: Three of a kind or three different kinds of employee well-being? *Applied Psychology*, *57*(2), 173–203.

Shaw, K. (2005). An engagement strategy process for communicators. *Strategic Communication Management*, *9*(3), 26.

Soane, E., Truss, C., Alfes, K., Shantz, A., Rees, C., & Gatenby, M. (2012). Development and application of a new measure of employee engagement: The isa engagement scale. *Human Resource Development International*, *15*(5), 529–547.

Stanley, D. (2021). *ApaTables: Create american psychological association (apa) style tables*. Retrieved from <https://CRAN.R-project.org/package=apaTables>

Staw, B. M., Sutton, R. I., & Pelled, L. H. (1994). Employee positive emotion and favorable outcomes at the workplace. *Organization Science*, *5*(1), 51–71.

Steinberg, L. (1994). Context and serial-order effects in personality measurement: Limits on the generality of measuring changes the measure. *Journal of Personality and Social Psychology*, *66*(2), 341–349.

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S. H. (2009). Testing measurement invariance using multigroup cfa: Differences between educational groups in human values measurement. *Quality & Quantity*, *43*(4), 599–616.

Taris, T. W., Ybema, J. F., & Beek, I. van. (2017). Burnout and engagement: Identical twins or just close relatives? *Burnout Research*, *5*, 3–11.

Timms, C., Brough, P., & Graham, D. (2012). Burnt-out but engaged: The co-existence of psychological burnout and engagement. *Journal of Educational Administration*, *50*(3), 327–345.

Weinberg, M. K., Seton, C., & Cameron, N. (2018). The Measurement of Subjective Wellbeing: Item-Order Effects in the Personal Wellbeing Index—Adult. Journal of Happiness Studies, 19(1), 315–332. <https://doi-org.ezproxy.montclair.edu/10.1007/s10902-016-9822-1>

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from <https://ggplot2.tidyverse.org>

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations*. Retrieved from <https://CRAN.R-project.org/package=stringr>

Wickham, H. (2021a). *Forcats: Tools for working with categorical variables (factors)*. Retrieved from <https://CRAN.R-project.org/package=forcats>

Wickham, H. (2021b). *Tidyr: Tidy messy data*. Retrieved from <https://CRAN.R-project.org/package=tidyr>

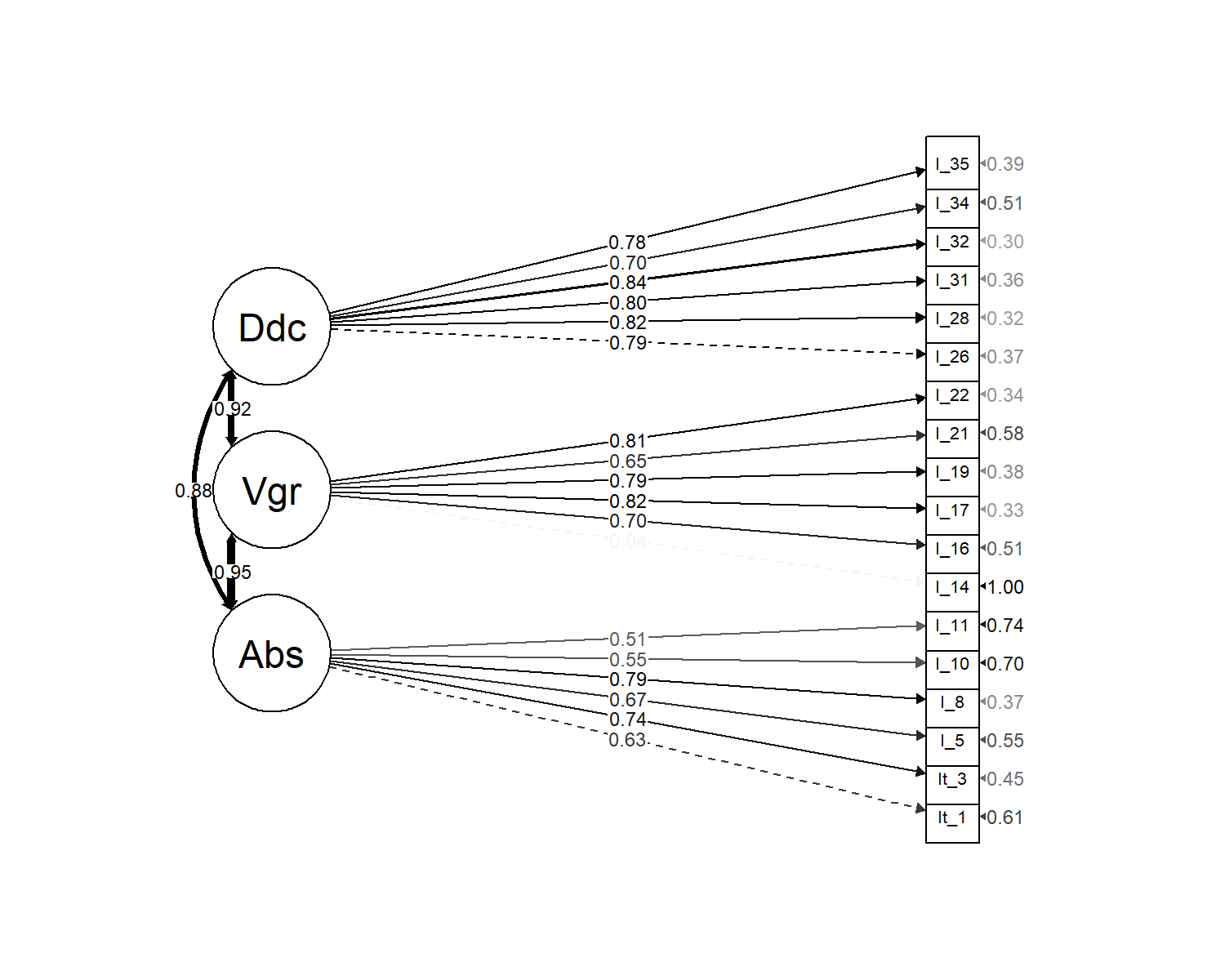
Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. <https://doi.org/10.21105/joss.01686>

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation*. Retrieved from <https://CRAN.R-project.org/package=dplyr>

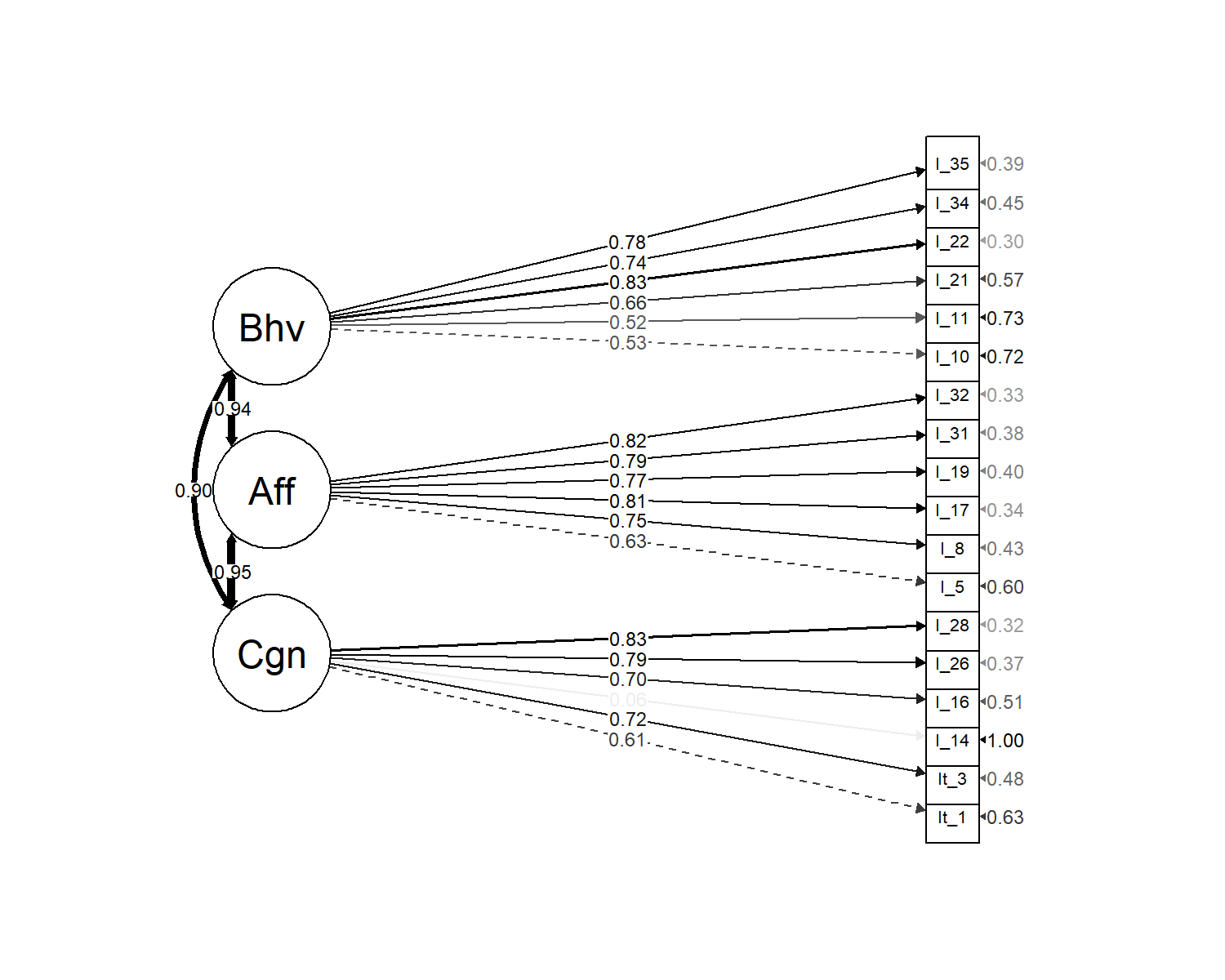
Wickham, H., & Hester, J. (2020). *Readr: Read rectangular text data*. Retrieved from <https://CRAN.R-project.org/package=readr>

Willmer, M., Westerberg Jacobson, J., & Lindberg, M. (2019). Exploratory and confirmatory factor analysis of the 9-item utrecht work engagement scale in a multi-occupational female sample: A cross-sectional study. *Frontiers in Psychology*, *10*. <https://doi.org/10.3389/fpsyg.2019.02771>

Xie, Y., Cheng, J., & Tan, X. (2021). *DT: A wrapper of the javascript library ’datatables’*. Retrieved from <https://CRAN.R-project.org/package=DT>



*Figure* *1.*  Omnibus Confirmatory Factor Analysis substantive structure.



*Figure* *2.*  Omnibus Confirmatory Factor Analysis attitudinal structure.

Table 1

Focal items and their scale associations

|  |  |  |
| --- | --- | --- |
| Substantive | Attitudinal | Item |
| Absorption | Cognitive | Item 1. I’m able to concentrate on my work without distractions |
| Absorption | Cognitive | Item 3. Time passes quickly while I’m working |
| Absorption | Affective | Item 5. I enjoy thinking about work even when I’m not at work |
| Absorption | Affective | Item 8. I love starting my workday |
| Absorption | Behavioral | Item 10. I have to be reminded to take breaks while I’m at work |
| Absorption | Behavioral | Item 11. I never miss a work deadline |
| Vigor | Cognitive | Item 14. Thinking about work saps my energy |
| Vigor | Cognitive | Item 16. I’m able to maintain good levels of energy throughout the workday |
| Vigor | Affective | Item 17. I enjoy spending time completing my job tasks |
| Vigor | Affective | Item 19. I feel motivated to go beyond what is asked of me at work |
| Vigor | Behavioral | Item 21. When work is slow I find ways to be productive |
| Vigor | Behavioral | Item 22. I express enthusiasm for my job while at work |
| Dedication | Cognitive | Item 26. I believe this company cares about my career goals |
| Dedication | Cognitive | Item 28. This organization challenges me to work at my full potential |
| Dedication | Affective | Item 31. I feel proud of my accomplishments within this organization |
| Dedication | Affective | Item 32. My job makes me feel like I’m part of something meaningful |
| Dedication | Behavioral | Item 34. I embrace challenging situations at work |
| Dedication | Behavioral | Item 35. I speak positively about this organization to others |

Table 2

*Summary fit indices across item ordering conditions*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Condition | Model |  | *df* | RMSEA | SRMR | CFI | TLI | AIC |
| Condition 1 | 3-factor substantive | 300.86 | 132 | 0.14 | 0.11 | 0.68 | 0.63 | 3,282.88 |
|  | 3-factor attitudinal | 290.33 | 132 | 0.14 | 0.11 | 0.70 | 0.65 | 3,272.35 |
| Condition 2 | 3-factor substantive | 310.01 | 132 | 0.15 | 0.10 | 0.71 | 0.66 | 3,257.45 |
|  | 3-factor attitudinal | 322.52 | 132 | 0.15 | 0.10 | 0.69 | 0.64 | 3,269.96 |
| Condition 3 | 3-factor substantive | 252.07 | 132 | 0.12 | 0.10 | 0.78 | 0.74 | 3,510.32 |
|  | 3-factor attitudinal | 275.74 | 132 | 0.13 | 0.10 | 0.73 | 0.69 | 3,534.00 |
| Condition 4 | 3-factor substantive | 224.96 | 132 | 0.10 | 0.09 | 0.82 | 0.79 | 3,421.64 |
|  | 3-factor attitudinal | 228.99 | 132 | 0.10 | 0.09 | 0.81 | 0.78 | 3,425.66 |
| Condition 5 | 3-factor substantive | 549.80 | 132 | 0.10 | 0.05 | 0.90 | 0.89 | 14,932.57 |
|  | 3-factor attitudinal | 497.90 | 132 | 0.10 | 0.04 | 0.92 | 0.90 | 14,880.67 |
| Condition 6 | 3-factor substantive | 468.02 | 132 | 0.09 | 0.05 | 0.91 | 0.90 | 17,953.02 |
|  | 3-factor attitudinal | 610.60 | 132 | 0.10 | 0.06 | 0.88 | 0.86 | 18,095.61 |
| Overall | 3-factor substantive | 995.34 | 132 | 0.09 | 0.05 | 0.91 | 0.90 | 46,915.05 |
|  | 3-factor attitudinal | 1,103.47 | 132 | 0.09 | 0.05 | 0.90 | 0.88 | 47,023.18 |

Table 3

Unit-weighted scale intercorrelations (all conditions)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 |  |  |
| 1. Absorption | - |  |  |  |  | 4.00 | 1.02 |
| 2. Vigor | .76\*\*\* | - |  |  |  | 4.23 | 0.90 |
| 3. Dedication | .76\*\*\* | .80\*\*\* | - |  |  | 4.41 | 1.12 |
| 4. Affect | .81\*\*\* | .84\*\*\* | .85\*\*\* | - |  | 4.09 | 0.95 |
| 5. Cognition | .87\*\*\* | .86\*\*\* | .89\*\*\* | .79\*\*\* | - | 4.16 | 1.12 |
| 6. Behavior | .84\*\*\* | .83\*\*\* | .84\*\*\* | .71\*\*\* | .81\*\*\* | 4.39 | 0.96 |

*Note.* \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001

Table 4

*Measurement invariance summary statistics (attitudinal structure snowball sample only)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Df | AIC | BIC | Chisq | Chisq diff | Df diff | Pr(>Chisq) |
| Configural | 528 | 13,645.98 | 14,453.38 | 1,117.59 | NA | NA | NA |
| Weak | 573 | 13,614.45 | 14,262.50 | 1,176.06 | 58.48 | 45 | 0.09 |
| Strong | 618 | 13,569.48 | 14,058.18 | 1,221.09 | 45.03 | 45 | 0.47 |
| Strict | 672 | 13,538.35 | 13,835.82 | 1,297.96 | 76.87 | 54 | 0.02 |

*Note.* \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001

Table 5

Measurement invariance summary statistics (substantive structure snowball sample only).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Df | AIC | BIC | Chisq | Chisq diff | Df diff | Pr(>Chisq) |
| Configural | 528 | 13,616.29 | 14,423.69 | 1,087.90 | NA | NA | NA |
| Weak | 573 | 13,588.39 | 14,236.44 | 1,150.00 | 62.10 | 45 | 0.05 |
| Strong | 618 | 13,546.51 | 14,035.20 | 1,198.12 | 48.12 | 45 | 0.35 |
| Strict | 672 | 13,521.28 | 13,818.74 | 1,280.89 | 82.77 | 54 | 0.01 |

*Note.* \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.00

Table 5

Measurement invariance summary statistics (attitudinal structure only conditions 5 and 6).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Df | AIC | BIC | Chisq | Chisq diff | Df diff | Pr(>Chisq) |
| configural.a2 | 264 | 33,048.28 | 33,557.42 | 1,108.50 | NA | NA | NA |
| weak.a2 | 279 | 33,046.22 | 33,488.37 | 1,136.44 | 27.94 | 15 | 0.02 |
| strong.a2 | 294 | 33,085.13 | 33,460.29 | 1,205.35 | 68.91 | 15 | 0.00 |
| strict.a2 | 312 | 33,100.74 | 33,395.51 | 1,256.96 | 51.61 | 18 | 0.00 |

*Note.* \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001

Table 6

Measurement invariance summary statistics (substantive structure only conditions 5 and 6).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Df | AIC | BIC | Chisq | Chisq diff | Df diff | Pr(>Chisq) |
| configural.s2 | 264 | 32,957.60 | 33,466.74 | 1,017.81 | NA | NA | NA |
| weak.s2 | 279 | 32,941.43 | 33,383.58 | 1,031.65 | 13.84 | 15 | 0.54 |
| strong.s2 | 294 | 32,969.93 | 33,345.09 | 1,090.15 | 58.50 | 15 | 0.00 |
| strict.s2 | 312 | 32,956.32 | 33,251.08 | 1,112.54 | 22.39 | 18 | 0.22 |

*Note.* \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001

1. Across conditions, the order of presentation of item “blocks” was also randomized. For example, not all respondents in Condition 1 was administered the Affect item block first - roughly 1/3 was presented the Behavioral block first and roughly 1/3 was presented the Cognitive block first. [↑](#footnote-ref-1)