

Measurement Invariance of a BiFactor Assessment of Engagement

Casey Osorio-Duffoo<sup>2</sup>, Renata Garcia Prieto Palacios Roji<sup>1</sup>, Morgan Russell<sup>1</sup>, & John  
Kulas<sup>1</sup>

<sup>1</sup> Montclair State University

<sup>2</sup> Harver

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

Correspondence concerning this article should be addressed to Casey Osorio-Duffoo,

1 Normal Ave, Montclair, NJ 07043. E-mail: caseyosorio@gmail.com

## Abstract

Employee engagement has, in recent years, enjoyed a surge in popularity as a positive employee outcome. Despite this burgeoning interest, disagreement still remains regarding its factor structure and nomological relationship with similar concepts, such as burnout. The classical definition of engagement divides it between the dimensions of vigor, dedication, and absorption, while the tripartite model of attitudes divides engagement between its cognitive, behavioral, and affective manifestations. Using bifactor analysis, study 1 proposes a scale that reconciles these two models. Study 2 convergently and discriminantly validates this scale.

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline.

Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines.

One sentence clearly stating the **general problem** being addressed by this particular study.

One sentence summarizing the main result (with the words “**here we show**” or their equivalent).

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

*Keywords:* Engagement, engagement

Word count: X

## Measurement Invariance of a BiFactor Assessment of Engagement

Renata's SEM paper will come in handy

Recent decades have seen a proliferation of interest and research in the construct of employee engagement.

more on why we're looking at tripartite model

The roots of employee (aka work; e.g., W. 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; William A. Kahn (1990a)). Staw, Sutton, and Pelled (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.

burnout

Although occasionally referred to as residing on the opposing pole to *burnout* (Christina Maslach & Leiter, 2008), these two constructs are currently most commonly conceptualized as being distinct (Goering, Shimazu, Zhou, Wada, & Sakai, 2017; Kim, Shin, & Swanger, 2009; Wilmar B. 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, Shimazu, Zhou, Wada, and Sakai (2017) concluded that they have a moderate (negative) association, but also distinct nomological networks. Wilmar B. Schaufeli, Taris, and Van Rhenen (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; C. Maslach & Leiter, 1997). Alternatively, engagement refers to an individual worker's involvement and satisfaction as well as enthusiasm for work (Harter, Schmidt, & Hayes, 2002). W. B. 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 (Wilmar B. 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."

Regarding measurement, Gallup is widely acknowledged as an early pioneer in the measurement of the construct (see, for example, Coffman & Harter, 1999). The Utrecht Work Engagement Scale (UWES) is another self-report questionnaire developed by W. B. Schaufeli and Bakker (2003) that directly assesses the vigor, dedication, and absorption elements.

The first, to our knowledge, use of the word “engagement” as a construct came in William A. Kahn (1990b), defining it 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 bypassed by subsequent papers (see, for example, (Baumruk, 2004) and (Shaw, 2005), who framed it in terms of one’s cognitive and affective *commitment* to one’s organization), William A. Kahn (1990b)’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.

Although falling out of favor in the decades following its construction, interest in the tripartite model was revived by Kaiser and Wilson (2019),

we need to do some market research on the Q12: 1. what’s the feedback report look like? (google images show one overall “satisfaction” score and/or one overall “engagement” score), 2. how much does it cost, 3. what are the 200 pulse items Gallup refers to? (6/7/21)

Our conceptualization of work engagement is a mental state wherein employees...

- ...feel energized (**Vigor**)
- ...are enthusiastic about the content of their work and the things they do (**Dedication**)
- ...are so immersed in their work activities that time seems compressed (**Absorption**)

This model is not without criticism, however. Some critics question its structural validity by pointing out that vigor, dedication and absorption all correlate highly with each other (Kulikowski, 2017).

need more on criticisms of model

The present article explores two methods for constructing a scale that incorporates both the substantive and attitudinal models into one, a more classical one based on corrected item-total correlations and one based on modification indices.

## Methods

Choice of focus on BIC versus AIC discussed in Dziak, Coffman, Lanza, Li, and Jermini (2020).

#Condition 1

#Condition 2

#Condition 3

#Condition 4

## Participants

330 individuals provided ratings across 36 candidate items. These participants were gathered via snowball sampling, with an initial population of undergraduate and graduate students, as well as professional acquaintances of faculty members.

## Material

Our survey was administered on Qualtrics.

**Item generation.** We generated a set of 36 items for our engagement measure, with the ultimate goal of reducing them to a final set of 18. These items were generated according to a review of extant tripartite engagement measures, as well as *WHAT RESEARCH DID WE USE FOR ATTITUDINAL WORDING? WAS IT LITERALLY*

*JUST “I THINK,” “I FEEL,” “I DO?”* Each item was worded to reflect both a substantive dimension as well as an attitudinal dimension. For example, the item “My job makes me feel like I’m part of something meaningful” reflects the affective dimension with “feel” and the dedication dimension with “I’m part of something meaningful.”

Our 3x3 bifactor model produced nine pairs of dimensions (e.g., Vigor-Cognitive, Vigor-Affective, Vigor-Behavioral, etc.). With 36 initial items, this left four items per pair of substantive and attitudinal dimensions.

The substantive scale definitions provided for ratings were:

- *Absorption*: Being fully immersed in one’s work, where time passes quickly and one has difficulty detaching from work tasks
- *Vigor*: Experiencing persistent levels of energy, effort, and enthusiasm while working
- *Dedication*: Experiencing pride and challenge in ones work, as well as strong feelings of support from and loyalty toward the organization

The attitudinal scale definitions were:

- *Cognitive*: Pertaining to thoughts or general mental processes (for example what someone thinks)
- *Affective*: Pertaining to feelings or emotions (for example, how someone feels)
- *Behavioral*: Pertaining to acts or actions (for example, what someone does)

See table X for a full list of items and their respective dimensions.

## Procedure

Looking into the specification of polychoric covariances (Jöreskog, 1994). This seems to be not very commonly leveraged (only package that seems to estimate these is `semPlot`).

The effective result of this was two divergent quasi-experimental approaches: 1) focus on corrected item-total correlations, and 2) focus on CFA modification indices.

### **Corrected item-total correlations.**

To Casey: document your process here

We conduct a correct item-total correlation on our original 36-items set. Base off, the r. drops that the corrected item-total correlations provide us we narrowed it down by selecting that items that had the best r. drops off removing one item at a time. For example, each cell division contain 4 items, therefore, we remove one of the four items creating 6 potential 3 item corrected item correlations, and from there we choose the items with the best r. drops. We continue the same process when narrowing our three items down to two items. An example is shown below:

**CFA Modification Indices.** We followed two parallel stepwise item-reduction processes centered around eliminating items in decreasing order of modification indices. Looking at the 36-item substantive and attitudinal models independently, we requested modification indices from each, with the intent of retaining indicators whose fixed shared residual covariances were associated with high modification indices (indicating better model fit if the paths were freed). The item pair with the highest modification index was scrutinized, with a subjective group judgment made on wording/semantics content domain coverage. The less preferred item was removed from the model. In cases where the highest modification index was between the only two remaining items in a substantive-attitudinal



pair, these items were passed over for scrutiny in favor of the items with the next-highest index. This process was repeated until 18 items remained (i.e., 2 items for each of the 9 substantive-attitudinal pairs)

For example, the path with the highest modification index across both CFAs was between item 2 and item 4, which are both indicators of “Absorption” and “Cognition.” One of these items was therefore a candidate for deletion, and semantic preference was given to item 4, “I find it difficult to mentally disconnect from work” over item 2. After item 2 was excluded from both scale definitions (substantive and attitudinal), the CFAs were re-run and modification indices re-checked for bi-factor structure optimizing modifications.<sup>1</sup>

The end result was two separate final scale definitions (one optimized for the substantive model and one for the attitudinal).

We prioritized item deletions such that an item was implicated for deletion if: 1) modification index was high (relative to others) and 2) error residual was within same “cell.” The choice of item to delete was based on author preference for wording/semantics as well as construct element coverage (considering the possible consequences for construct deficiency). Item variance was also consulted (retention more likely with greater item variance).

**Single factor versus bifactor approaches.** We conducted correct-item total correlations using

## Data analysis

We used R [Version 4.1.1; R Core Team (2021)] and the R-packages *apaTables* (Stanley, 2021), *dplyr* [Version 1.0.7; Wickham, François, Henry, and Müller (2021)], *DT*

---

<sup>1</sup> Probably put a table in here highlighting certain modification indices (with a key to intended factor-item association). Look at “modincides1”

[Version 0.19; Xie, Cheng, and Tan (2021)], *forcats* [Version 0.5.1; Wickham (2021a)], *ggplot2* [Version 3.3.5; Wickham (2016)], *kableExtra* [Version 1.3.4; Zhu (2021)], *labourR* (Kouretsis, Bampouris, Morfiris, & Papageorgiou, 2020), *lavaan* [Version 0.6.9; Rosseel (2012)], *magrittr* [Version 2.0.1; Bache and Wickham (2020)], *papaja* [Version 0.1.0.9997; Aust and Barth (2020)], *purrr* [Version 0.3.4; Henry and Wickham (2020)], *readr* [Version 2.0.1; Wickham and Hester (2020)], *sem* [Version 3.1.11; Fox, Nie, and Byrnes (2020); Epskamp (2019)], *semPlot* [Version 1.1.2; Epskamp (2019)], *stringr* [Version 1.4.0; Wickham (2019)], *tibble* [Version 3.1.2; Müller and Wickham (2021)], *tidyr* [Version 1.1.3; Wickham (2021b)], and *tidyverse* [Version 1.3.1; Wickham et al. (2019)] for all our analyses.

## Results

CFA drafts below

### Study 2

Construct validation was accomplished via administration of the 17-item UWES as well as the Saks (2006) 12-item scale. Saks (2006) aggregates to two scales: job and organizational engagement.

## Discussion

## References

- 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.
- Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in Bioinformatics*, *21*(2), 553–565.
- 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.
- Harter, J. K., Schmidt, F. L., & Hayes, T. L. (2002). Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis. *Journal of Applied Psychology*, 87(2), 268.
- Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. Retrieved from <https://CRAN.R-project.org/package=purrr>
- Jöreskog, K. G. (1994). On the estimation of polychoric correlations and their asymptotic covariance matrix. *Psychometrika*, 59(3), 381–389.
- Kahn, William A. (1990b). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, 33(4), 692–724.
- Kahn, William A. (1990a). 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>

- 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>
- 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](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.
- 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, Christina, & Leiter, M. P. (2008). Early predictors of job burnout and engagement. *Journal of Applied Psychology*, 93(3), 498–512.
- 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>
- 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/>
- Saks, A. M. (2006). Antecedents and consequences of employee engagement. *Journal of Managerial Psychology*, 21(7), 600–619.
- Schaufeli, W. B., & Bakker, A. B. (2003). UWES–utrecht work engagement scale: Test manual. *Unpublished Manuscript: Department of Psychology, Utrecht University*, 8.
- Schaufeli, Wilmar 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, Wilmar 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.
- 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.
- Shaw, K. (2005). An engagement strategy process for communicators. *Strategic Communication Management*, 9(3), 26.
- 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.
- 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.
- 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>
- Xie, Y., Cheng, J., & Tan, X. (2021). *DT: A wrapper of the JavaScript library 'DataTables'*. Retrieved from <https://CRAN.R-project.org/package=DT>

Zhu, H. (2021). *kableExtra: Construct complex table with 'kable' and pipe syntax*.

Retrieved from <https://CRAN.R-project.org/package=kableExtra>



Table 1

*Scale intercorrelations (students low-stakes).*

	chisq	df	pvalue	rmsea	srmr	cfi	tli	aic
Substantive	300.86	132.00	0.00	0.14	0.11	0.68	0.63	3,282.88
Attitudinal	290.33	132.00	0.00	0.14	0.11	0.70	0.65	3,272.35

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

Model	$\chi^2$	<i>df</i>	RMSEA	SRMR	CFI	TLI	AIC
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

Model	$\chi^2$	<i>df</i>	RMSEA	SRMR	CFI	TLI	AIC
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

Model	$\chi^2$	<i>df</i>	RMSEA	SRMR	CFI	TLI	AIC
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	3534

Model	$\chi^2$	<i>df</i>	RMSEA	SRMR	CFI	TLI	AIC
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