Research Article

Measuring the hybrid workplace: development and cross-sector validation of the Hybrid Work Resources and Outcomes Inventory (HWROI) among Indian knowledge workers

Subhadhra Srinivasan*

Ph.D Research scholar, Department of Management, Bharath Institute of Higher Education and Research, 173, Agharam Main Road, selaiyur chennai 600073, Tamilnadu, India, Email: subhadra_jai@yahoo.co.in

Abstract

Purpose – Research on hybrid work routinely combines arrangement features, the resources employees draw from them and the outcomes they report, yet few instruments measure all three with documented psychometric evidence, and almost none have been validated across sectors that differ in how portable their work is. This study reports the development and validation of a 49-item, nine-dimension inventory for this purpose. M**ethodology:** The items were developed from literature, subjected to content validity ratio by five-member expert panel and pre-tested with 25 employees. The inventory was subsequently given to the staff of 21 information technology (IT), finance and shipping companies in Chennai, India (N = 417 usable responses, usable response rate 67.2%). The evidence for item loadings, composite reliability, convergent validity, discriminant validity, common method variance, measurement invariance of composites (MICOM) across the three sectors and known-groups and nomological was gathered. **Findings**: Forty-eight of the 49 standardised loadings exceeded .70 (range .67–.95). The range of composite reliability was between .88 to .95 and average variance extracted from .59 to .79.All constructs met the Fornell–Larcker criterion; the largest disattenuated correlation was .72 (upper 95% bound ≈ .79). Common method diagnostics did not indicate dominant method variance (single-factor share 34.2%; full-collinearity VIF 1.64–2.26). Compositional invariance held across sectors while composite means differed, establishing partial measurement invariance. Sector differences were large for arrangement features (ω^2 = .26–.30) and smaller for resources and outcomes (ω^2 = .05–.13), as expected. **Originality/value** – The inventory contains a validated, sector-portable assessment of the hybrid work feature–resource–outcome chain for an emerging-economy workforce, and the paper highlights certain design challenges — the mixed format of responses and the orientation of the outcome statements to the arrangement — that future users will need to address.

Keywords: hybrid work; scale development; content validity; PLS-SEM; measurement invariance; MICOM; work–life balance; India

1. Introduction

Hybrid work, in which employees divide paid work between home and an employer's premises, has settled into a lasting arrangement for a large share of professional employees since the COVID-19 pandemic (Allen et al., 2015; Bloom et al., 2024). A field experiment in a large technology firm found that two home-working days a week reduced attrition without measurable harm to performance (Bloom et al., 2024), and meta-analytic evidence links remote work to modestly better autonomy, job satisfaction and work–family outcomes (Gajendran & Harrison, 2007). The research agenda has accordingly moved from asking whether hybrid work is beneficial to asking how, for whom and under what conditions it is beneficial. Answering those questions depends on measurement that keeps three things apart: the features of the arrangement an organisation offers, the resources an employee is actually able to draw from it, and the outcomes that follow.

These levels are not usually distinguished in existing instruments. Many studies have focused on one question regarding hybrid work and that has been on "days worked from home" or in "home and office" environments, combining the hybrid option with the mode of use (Allen et al., 2015). Others take general scales of work-family balance from Greenhaus et al., 2003, or well-being, and do not measure the resources that are assumed to be involved in the relationship between the two. Resource based theories turn the omission into a significant issue. Conservation of resources theory suggests that people's well-being is linked to their acquisition, safeguarding, and investment in valuable resources (Hobfoll, 1989), and the job demands–resources model posits that job resources are the immediate antecedents of engagement and health (Bakker & Demerouti, 2017). If a hybrid arrangement is beneficial to well-being, it should do so because it provides resources (such as usable workspace, control over time, freedom from commuting) and an instrument that is not a measure of resources can not test the provision/receipt account.

There's a second gap, which is contextual. Validation efforts have been focused on western sample and where as much as possible can be validated by working remotely. While some sectors are more mobile than others,

software development is easily mobile, financial services are subject to supervisory and data-protection rules, and shipping has a mix of mobile administrative operations and physical presence. For comparative studies, any instrument used must demonstrate that its composites have the same meaning in both contexts to explain any level differences and/or relationships (Henseler et al., 2016; Putnick & Bornstein, 2016).

2. Construct domain

The inventory is organised in three tiers (Table 1). The first tier measures *hybrid work features* — what the arrangement makes available: the type and regularity of the hybrid arrangement, the implementation of work-from-home policy, and the availability of flexible scheduling. The second tier measures *resource mechanisms* — what the employee obtains: an adequate home workspace, a sense of control over working time, and relief from commuting. The third tier measures *outcomes*: satisfaction with work–life balance, well-being, and successful adaptation to hybrid working. The tiers follow the distinction drawn in boundary-management research between the formal availability of flexibility and the control employees perceive over it (Kossek et al., 2006), and the resource logic of conservation of resources theory (Hobfoll, 1989).

Table 1. HWROI dimensions, item counts and sample items

Tier	Dimension (abbreviation)	Items	Sample item
Features	Hybrid arrangement type (HAT)	5	"My work arrangement allows me to choose my work location on a daily basis"
	Work-from-home policy implementation (WPI)	5	"My organisation provides adequate technology support for home-based work"
	Flexible schedule availability (FSA)	5	"I have significant control over when I start and end my work day"
Resources	Home workspace adequacy (HWA)	6	"My home workspace provides adequate desk and seating arrangements"
	Schedule flexibility perception (SFP)	6	"I feel I have adequate control over my work schedule timing"
	Reduced commuting time (RCT)	6	"I save significant time by not commuting as frequently"
Outcomes	Work–life balance satisfaction (WLB)	5	"I am satisfied with the balance I achieve between my work and personal life"
	Employee well-being (EWB)	6	"I have adequate energy for both work and personal activities"
	Adaptation success (EAS)	5	"I have successfully adapted to working in hybrid arrangements"

Note. Each item from the agreement was given a rating of 1 to 7 (from strongly disagree to strongly agree). One HAT item asks for the percentage of working time spent away from the workplace, and two RCT items ask for hours and days of commuting avoided per week; these are addressed later in the "Limitations" data.

3. Method

3.1 Item generation and content validity

Items were written to cover each dimension's definition, drawing on established work–family balance (Greenhaus et al., 2003) and job resources (Bakker & Demerouti, 2017) measures and adapted to hybrid work. Following recommended practice for scale development (Boateng et al., 2018; Hinkin, 1998), a five-member panel judged each item: two academic researchers in work–life and organisational behaviour, an industrial–organisational psychologist, a human resources practitioner with hybrid implementation experience in Chennai, and a survey methodologist. Panellists rated each item as essential, useful but not essential, or not necessary, rated clarity on a four-point scale, and commented on coverage.

Item necessity was summarised with Lawshe's (1975) content validity ratio, CVR = (n_e − N/2)/(N/2), where n_e is the number of panellists rating an item essential. With N = 5, CVR can take only the values −1.00, −0.60, −0.20, 0.20, 0.60 and 1.00, and the critical value for one-tailed α = .05 is 1.00 (Ayre & Scally, 2014), so only unanimous endorsement is statistically distinguishable from chance agreement. In the first round 89% of items reached CVR ≥ 0.60. The panel made 23 suggestions: 12 wording changes, six response-option changes, three

additional items and two deletions for redundancy. After revision, a second review round was conducted, and every retained item was rated essential by all five panellists (CVR = 1.00).

3.2 Pilot test

A field test sample of 25 employees (26-58 years; 52% female) were drawn using the same recruitment strategy as the main study (IT 9, Finance 8, Shipping 8). Testing included rendering on various devices and browsers, completion time, frequency of correct answers and correct/incorrect response distributions. The time taken to complete the means was 17.3 minutes (range 13–24). The following changes were made to make the items more comprehensible, to revise response options, and to improve section transitions: High internal consistency coefficients for all the pilots are given in Table 2 and the complete sequence of development is summarized in Figure 1.

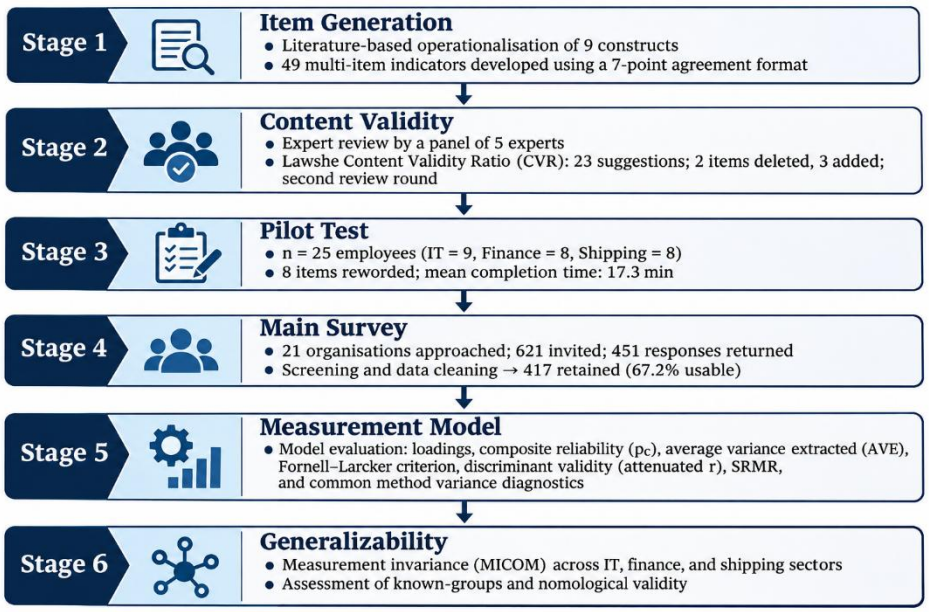


Figure 1. *Development and validation sequence of the Hybrid Work Resources and Outcomes Inventory.*

3.3 Main survey: sample and procedure

The validation sample was drawn from 21 organisations in Chennai, India, stratified by sector (IT; banking, insurance and other financial services; shipping lines, freight forwarding, port and maritime services) and size (50–250, 251–1,000 and more than 1,000 employees). Within each organisation, a liaison distributed the online questionnaire (Google Forms) by e-mail to employees in roles eligible for hybrid working; in shipping, this restricted the frame to administrative, documentation, customer-service and management positions.

Of 621 employees invited, 451 returned questionnaires. Cases were screened for excessive missing data (more than 10% of items), failed attention checks, invariant responding across scales, logically inconsistent answers and implausible completion times. After screening, 417 responses were retained (usable response rate 67.2%; IT 147/206, finance 142/208, shipping 128/207). Remaining item-level missingness was low (1.05% of all item responses; highest for workspace items at 1.8%), and Little's test did not reject the hypothesis that data were missing completely at random ($\chi^2 = 23.45$, $p = .891$). Univariate ($|z| > 3.29$) and multivariate (Mahalanobis, $p < .001$) screening flagged 12 and 8 cases respectively; none exerted undue influence and all were retained.

The respondents ranged in age from 22 to 64 years (M = 38.4, SD = 9.2); 52.4% were female and 57.3% had a master's or professional degree; and the mean length of time in the organisation was 7.2 years. Of the sample, 75.5% worked under a formal hybrid arrangement at the time of the survey, 25.4% had other family care responsibilities, and 31.4% had no dependent children; 59.8% felt the hybrid work model was the ideal arrangement, 15.5% were satisfied with it and 24.7% were not.

3.4 Analysis

The measurement model was estimated using partial least squares structural equation modelling (SmartPLS 3.3.3), the path weighting scheme and 5,000 bootstrap subsamples, and assessed based on the existing reporting guidelines (Hair et al., 2019; Hair et al., 2022). The loading of .708 or greater was used to assess the reliability of the indicators, indicating that the construct measures more than 50% of the variance of the indicator. Cronbach's α coefficient and composite reliability (ρc) were used to assess the internal consistency,

and values > .95 were deemed as a warning of item redundancy (Hair et al., 2019). The average variance extracted (AVE) had to be .50 or higher as per the convergent validity criterion. Standardised outer loadings were directly calculated to get composite reliability and AVE.

Discriminant validity was assessed with the Fornell–Larcker criterion (Fornell & Larcker, 1981). Because item-level correlations required for the heterotrait–monotrait ratio (Henseler et al., 2015) were not available for this report, inter-construct correlations were also corrected for unreliability ($r_{ij}/\sqrt{(\alpha_i\alpha_j)}$). Under tau-equivalence this disattenuated correlation is the quantity HTMT estimates, and the correction allows the more stringent recommendation that the upper confidence bound remain below .80–.85 to be applied (Rönkkö & Cho, 2022). Upper bounds were obtained by Fisher-*z* transformation of the observed correlation before correction.

Common method variance was controlled for by the commonly recommended procedural remedies (Podsakoff et al., 2003 2012), a full collinearity test (Kock, 2015) and an unmeasured latent method factor plus a marker variable supposed to be theoretically unrelated to the measurement (workplace aesthetics, Lindell & Whitney, 2001) with the caution that the latter test is insensitive (Fuller et al., 2016). The three-step MICOM (configural invariance, compositional invariance, equality of composite means and variances; 5000 permutations; Henseler et al., 2016) was performed to assess the measurement invariance across sectors. One-way analysis of variance (ANOVA) was performed on the composite scores (calculated from means and standard deviations) for one-way validation, known-groups validity, was evaluated with the effect size (ω^2).

4. Results

4.1 Indicator reliability, internal consistency and convergent validity

All 49 loadings were statistically significant (all $t > 12$, $p < .001$). Forty-eight exceeded .708; the single exception was the first hybrid arrangement type item (HAT1, .67), the percentage-of-time item, which was retained because its removal did not raise ρ_c or AVE above thresholds that were already met (Figure 2a). Loadings were highest for the outcome dimensions (WLB .81–.95; EWB .76–.92) and lowest for the feature dimensions (HAT .67–.84; FSA .69–.86).

A summary of reliability and convergent validity is given in Table 2. Cronbach's α values on the main sample ranged between .82 and .92, while the values in the pilot ranged from .78 to .89. All dimensions of composite reliability and AVE exceeded the convergent validity threshold (from .88 (HAT) to .95 (WLB) and from .59 (HAT) to .79 (WLB) respectively).

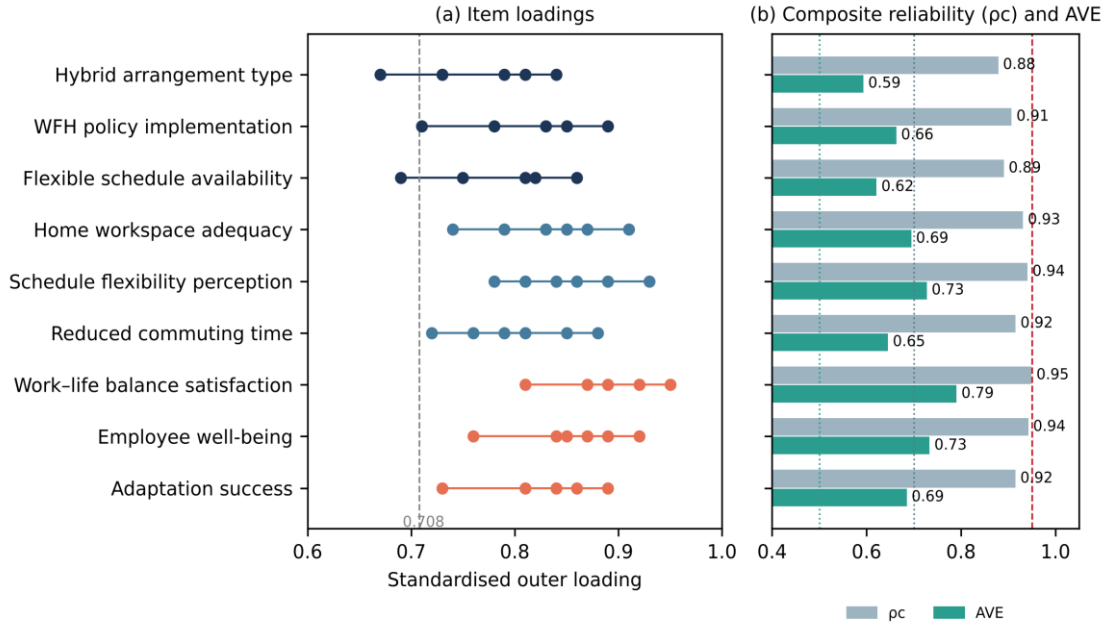


Figure 2. *(a) Standardised outer loadings of the 49 items by dimension (dashed line = .708); (b) composite reliability (ρ_c) and average variance extracted (AVE); dotted lines mark .50 and .70, the dashed red line marks .95.*

Table 2. Reliability and convergent validity of the HWROI dimensions (N = 417)

Dimension	Items	Loading range	α (pilot, n = 25)	α (main)	ρ_c	AVE	Test–retest *r*
Hybrid arrangement type	5	.67–.84	.78	.82	.88	.59	.78
WFH policy implementation	5	.71–.89	.85	.87	.91	.66	.83
Flexible schedule availability	5	.69–.86	.79	.83	.89	.62	.79
Home workspace adequacy	6	.74–.91	.86	.89	.93	.69	.85
Schedule flexibility perception	6	.78–.93	.88	.91	.94	.73	.88
Reduced commuting time	6	.72–.88	.84	.87	.92	.65	.83
Work–life balance satisfaction	5	.81–.95	.89	.92	.95	.79	.89
Employee well-being	6	.76–.92	.87	.90	.94	.73	.86
Adaptation success	5	.73–.89	.83	.86	.92	.69	.81

Note. ρ_c = composite reliability; AVE = average variance extracted; both computed from the standardised outer loadings. All loadings $p < .001$.

4.2 Discriminant validity

To meet the Fornell–Larcker criterion, the square root of each dimension's AVE (.77–.89) was greater than their correlation with all other dimensions (.32–.63) in all 36 comparisons (Figure 3). Correcting for unreliability raised the inter-construct correlations to between .37 and .72. The three highest disattenuated correlations involved conceptually adjacent pairs: flexible schedule availability with schedule flexibility perception (.72; upper 95% bound .79), Implementation of work from home policy, availability of flexible work schedule (.72; .79) and hybrid arrangement type with work-from-home policy implementation (.69; .76). All upper bounds were below .80, the most conservative threshold proposed for discriminant validity (Rönkkö & Cho, 2022). The availability of flexible scheduling and the experience of schedule control are therefore empirically distinct, which matters because the inventory's purpose is to keep provision and receipt apart. Among the outcomes, the largest disattenuated correlation was between work–life balance satisfaction and well-being (.65), indicating related but separable constructs.

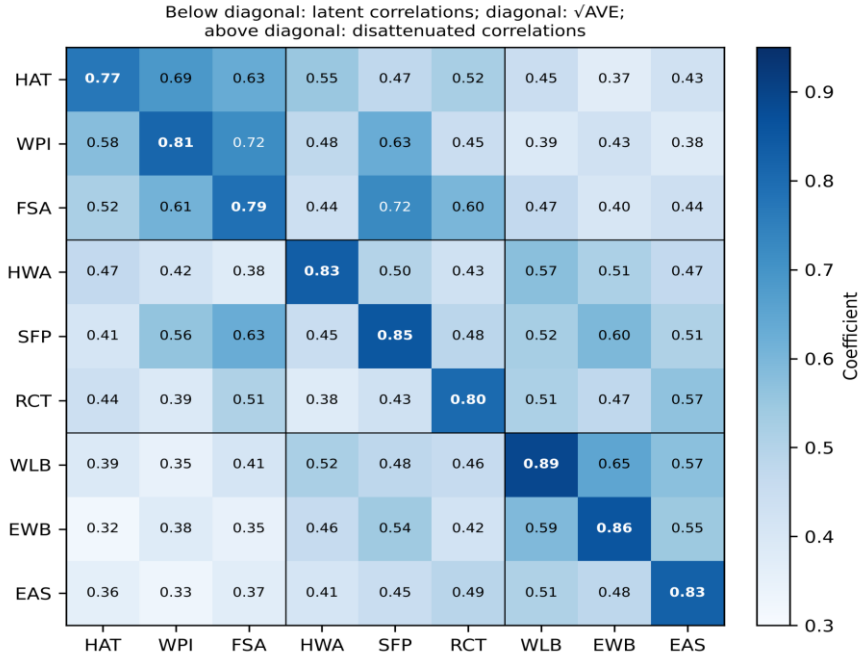


Figure 3. *Discriminant validity matrix. The correlations are latent inter-construct, the square root of the AVEs are on the diagonal (bold), and the correlations corrected for unreliability are above the diagonal. HA = hybrid arrangement type; WPI = work-from-home policy implementation; FSA = flexibility in the work schedule; HWA = home workspace adequacy; SFP = perception of flexibility in the work schedule; RCT = reduced commuting time;*

WLB = satisfaction with work–life balance; EWB = employee well-being; EAS = success in adapting to the home environment.

4.3 Model fit and common method variance

The nine-dimension measurement model had an SRMR of .065, below the .08 benchmark, and a normed fit index of .91 (Table 3). Alternative structures that merged conceptually adjacent dimensions or collapsed the three outcomes into a single factor fitted worse (single-outcome model SRMR = .156; NFI = .794). Because SRMR in composite models is a descriptive rather than an inferential index, these comparisons are offered as supporting rather than decisive evidence.

Table 3. Model fit and common method variance diagnostics

Assessment	Result	Benchmark	Interpretation
SRMR (nine-dimension model)	.065	< .08	Acceptable
Normed fit index	.91	> .90	Acceptable
RMS_{theta}	.089	< .12	Acceptable
SRMR, single-outcome alternative	.156	—	Inferior to nine-dimension model
Harman single-factor share	34.2%	< 50%	No single dominant factor
Full-collinearity VIF (range)	1.64–2.26	< 3.3	No pathological collinearity
Unmeasured latent method factor	12.8% of indicator variance (8.3–18.7%)	—	Modest method variance
Marker variable correlation	*r* = .08, *p* = .10	ns	No marker contamination

Note. Full-collinearity VIFs are the diagonal elements of the inverse of the latent correlation matrix.

The variance of the items accounted for by a single factor was low (34.2%). The full-collinearity VIFs ranged from 1.64 to 2.26, which is not considered to be common method contamination by Kock (2015) at 3.3. The average variance accounted for by a latent method factor was 12.8% and the marker variable was not significantly associated with the substantive constructs (r = .08, p = .10). Overall, these diagnostics indicate that there is some shared method variance, as would be anticipated with instruments of a single source, but not sufficient to explain the correlations found.

4.4 Measurement invariance across sectors

Configural invariance was achieved by design: all indicators were identical, and both data treatment and algorithm settings were the same for each sector. All nine dimensions were compositionally invariant, as the correlations between the group-specific and pooled scores were between .996 and .999 and were not significantly less than 1 (permutation p = .134–.334) (Table 4). Variances were not significantly different across sectors for composite (p = .098–.267) but were significantly different across sectors for composite means (all p < .001) in step 3. This pattern is not complete measurement invariance (Henseler et al., 2016) but partial measurement invariance. Partial invariance is enough for comparison of structural relationships between sectors, however since there is variation in the composite means, the sectors are not pooled without modelling the sector as a grouping or control variable.

Table 4. Measurement invariance of composites (MICOM) and known-groups comparison by sector

Dimension	Step 2 *c*	Step 2 *p*	Step 3 variance *p*	Step 3 mean *p*	IT M (SD)	Finance M (SD)	Shipping M (SD)	*F*(2, 414)	ω²
Hybrid arrangement type	.998	.267	.178	< .001	6.1 (1.0)	5.2 (1.3)	4.1 (1.5)	84.58	.29
WFH policy implementation	.999	.189	.267	< .001	6.3 (0.9)	5.4 (1.2)	4.6 (1.4)	72.18	.26
Flexible schedule availability	.997	.334	.123	< .001	5.8 (1.2)	4.9 (1.4)	3.7 (1.3)	89.24	.30
Home workspace adequacy	.996	.245	.234	< .001	5.9 (1.0)	5.4 (1.2)	4.9 (1.3)	25.17	.10
Schedule flexibility	.999	.156	.189	< .001	5.6	5.0 (1.3)	4.4 (1.4)	30.77	.12

Dimension	Step 2 *c*	Step 2 *p*	Step 3 variance *p*	Step 3 mean *p*	IT M (SD)	Finance M (SD)	Shipping M (SD)	*F*(2, 414)	ω²
perception					(1.1)				
Reduced commuting time	.998	.223	.098	< .001	5.4 (1.4)	4.9 (1.6)	4.3 (1.7)	16.90	.07
Work–life balance satisfaction	.999	.134	.134	< .001	5.8 (1.1)	5.3 (1.3)	4.8 (1.4)	21.37	.09
Employee well-being	.997	.198	.156	< .001	5.5 (1.0)	5.1 (1.2)	4.7 (1.3)	16.11	.07
Adaptation success	.998	.267	.167	< .001	5.4 (1.2)	5.0 (1.3)	4.6 (1.4)	13.02	.05

Note. c = correlation between composite scores using group-specific and pooled weights; *p* values from 5,000 permutations. *F* and ω² recomputed from group means and standard deviations (IT n = 147; finance n = 142; shipping n = 128); all *F* tests *p* < .001.

4.5 Known-groups and nomological validity

If the feature dimensions capture what organisations make available, they should differ most between sectors whose work is more and less portable; if the resource and outcome dimensions capture what employees experience, they should differ in the same direction but less strongly, because individual circumstances intervene. Both expectations were met (Table 4). Sector accounted for 26–30% of the variance in the three feature dimensions, with IT–shipping differences of 1.47–1.68 standard deviations, but for only 5–12% of the variance in the resource and outcome dimensions (IT–shipping *d* = 0.62–0.96).

The correlations among dimensions were also ordered as the construct domain implies (Figure 3). Each feature dimension correlated most strongly with the resource it is meant to deliver or with a close neighbour (for example, flexible schedule availability with schedule flexibility perception, *r* = .63), and the resource dimensions correlated more strongly with the outcomes (*r* = .38–.54) than the feature dimensions did (*r* = .32–.41). This gradient — features relate to outcomes mainly through resources — is the pattern a resource account predicts (Hobfoll, 1989) and supports the inventory's nomological validity.

5. Discussion

5.1 Principal findings

The HWROI met conventional Indicators' reliability, internal consistency, convergent validity and discriminant validity in a sample of 417 employees across three sectors. Its greatest comparative research strength was that despite the sizeable disparities between the amount of hybrid work in IT, finance and shipping, composites were created in an equivalent manner. The researcher can thus with some assurance make comparisons of the relations between the dimensions as they exist within the sectors, and interpret mean differences within each sector as substantive rather than measurement artifacts.

It is important to note that the discriminant validity results, which indicate that the inventory differentiates between provision and receipt, are of special significance since the value of the inventory lies in its ability to distinguish provision from receipt. The difference between what the organization provides and what the employee gets would vanish if the employees' responses to the flexible schedule availability items were identical to their responses to the schedule flexibility perception items. The disattenuated correlation is .72 which means that the two are significantly correlated, which is appropriate, but that about half of the variance in each is reliable. This aligns with findings that formal flexibility policies and employees' perceptions of control over their work are related but differ (Kossek et al., 2006), as it is through this that implementation gaps can be identified in the inventory. The second line of support are the known-groups pattern. There was greater variance explained in feature scores (roughly three times) than resource and outcome scores. An instrument that only recorded general satisfaction with the employer would not generate this attenuation.Such an instrument needs to capture a sequence from provision to experience.

5.2 Weaknesses and recommended revisions

Validation is incomplete without an account of where an instrument falls short, and four weaknesses should guide future use.

First, three indicators do not share the seven-point agreement format of the rest: the percentage of working time spent at home (HAT1) and the weekly hours and days of commuting avoided (two RCT items). Combining open-ended quantities with agreement ratings in a reflective composite assumes that all indicators reflect the

same latent attribute on comparable metrics, an assumption that is hard to defend (Jarvis et al., 2003; MacKenzie et al., 2011). HAT1 was also the only item below the .708 loading benchmark. We recommend that these quantities be retained as separate single-indicator variables describing arrangement intensity, and that the composites be formed from the agreement items alone.

Second, several outcome items attribute the outcome to the arrangement — for example, "my work arrangement positively affects my overall stress levels". Attribution-laden wording builds the hypothesised causal link into the measure and may inflate associations between arrangement and outcome measures. Users wishing to test the effects of hybrid work on well-being should pair the inventory with an attribution-free well-being measure or rephrase these items as state descriptions.

Third, the work–life balance satisfaction dimension approached the redundancy threshold (ρ_c = .95). A four-item version omitting the highest-loading item would probably retain adequate reliability, but this should be verified with item-level data.

Fourth, the content validity panel was small. With five panellists, only unanimous endorsement distinguishes item necessity from chance (Ayre & Scally, 2014). The second-round result that all retained items were unanimously endorsed meets this bar, but a larger panel, or a content validity index with clarity ratings (Polit & Beck, 2006), would provide more informative evidence.

5.3 Implications for research and practice

For researchers, the inventory provides a ready means of testing resource-based accounts of hybrid work in which features act through resources, and of comparing those processes across sectors. For organisations, the feature subscales can serve as an audit of what is formally available, and the resource subscales as a check on whether provision reaches employees. A large difference between the two is a practical signal that policy is not translating into experience, the situation the inventory was designed to detect.

5.4 Limitations and further validation

Evidence is based upon a cross sectional survey of the professional workers in one metropolitan labour market and all dimensions were self-reported. While method variance diagnostics were confirmed, it is important to note that these cannot completely rule out the possibility of shared-method inflation (Podsakoff, 2012) Variance at the organisational level was not modelled, as employees were nested in only 21 organisations. Invariance was tested across sectors only; tests across gender, age, family composition and job level are needed before group comparisons on those characteristics are interpreted. Criterion validity against objective indicators — badge or calendar data on location, commuting records, or health measures — and predictive validity for retention remain to be established. Finally, the inventory was administered in English; translation and validation in Tamil and other Indian languages would extend its reach to employees who are less comfortable completing English questionnaires.

6. Conclusion

The Hybrid Work Resources and Outcomes Inventory measures three features of hybrid arrangements, three resources employees draw from them and three outcomes, with reliability, convergent validity and discriminant validity that meet current standards, and with composites that are formed equivalently in IT, finance and shipping. Its main contribution is to make the gap between what organisations offer and what employees receive measurable. With the revisions recommended here — separating quantitative intensity items, removing outcome attributions and testing a shorter work–life balance scale — it can serve comparative research on hybrid work in emerging economies.

References

1. Allen, T. D., Golden, T. D., & Shockley, K. M. (2015). How effective is telecommuting? Assessing the status of our scientific findings. *Psychological Science in the Public Interest, 16*(2), 40–68. https://doi.org/10.1177/1529100615593273
2. Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development, 47*(1), 79–86. https://doi.org/10.1177/0748175613513808
3. Bakker, A. B., & Demerouti, E. (2017). Job demands–resources theory: Taking stock and looking forward. *Journal of Occupational Health Psychology, 22*(3), 273–285. https://doi.org/10.1037/ocp0000056
4. Bloom, N., Han, R., & Liang, J. (2024). Hybrid working from home improves retention without damaging performance. *Nature, 630*(8018), 920–925. https://doi.org/10.1038/s41586-024-07500-2

5. Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149
6. Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104
7. Fuller, C. M., Simmering, M. J., Atinc, G., Atinc, Y., & Babin, B. J. (2016). Common methods variance detection in business research. *Journal of Business Research, 69*(8), 3192–3198. https://doi.org/10.1016/j.jbusres.2015.12.008
8. Gajendran, R. S., & Harrison, D. A. (2007). The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences. *Journal of Applied Psychology, 92*(6), 1524–1541. https://doi.org/10.1037/0021-9010.92.6.1524
9. 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*(3), 510–531. https://doi.org/10.1016/S0001-8791(02)00042-8
10. Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). SAGE. ISBN 9781544396408
11. Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203
12. Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8
13. Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using partial least squares. *International Marketing Review, 33*(3), 405–431. https://doi.org/10.1108/IMR-09-2014-0304
14. Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1*(1), 104–121. https://doi.org/10.1177/109442819800100106
15. Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist, 44*(3), 513–524. https://doi.org/10.1037/0003-066X.44.3.513
16. Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research, 30*(2), 199–218. https://doi.org/10.1086/376806
17. Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration, 11*(4), 1–10. https://doi.org/10.4018/ijec.2015100101
18. Kossek, E. E., Lautsch, B. A., & Eaton, S. C. (2006). Telecommuting, control, and boundary management: Correlates of policy use and practice, job control, and work–family effectiveness. *Journal of Vocational Behavior, 68*(2), 347–367. https://doi.org/10.1016/j.jvb.2005.07.002
19. Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology, 86*(1), 114–121. https://doi.org/10.1037/0021-9010.86.1.114
20. MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly, 35*(2), 293–A5. https://doi.org/10.2307/23044045
21. Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*(1), 539–569. https://doi.org/10.1146/annurev-psych-120710-100452
22. Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879
23. Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*(5), 489–497. https://doi.org/10.1002/nur.20147
24. Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004
25. Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational Research Methods, 25*(1), 6–14. https://doi.org/10.1177/1094428120968614