Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Validation of the Arabic Basic Psychological Need Satisfaction and Frustration Scale Among Pre-Service and In-Service Teachers in Kuwait

Fahad Ameen

fahad.ameen@ku.edu.kw

Kuwait University

Yousef Alardhi

Kuwait University

Research Article

Keywords: Basic Psychological Need Satisfaction and Frustration Scale, Arabic validation, teacher motivation, Self-Determination Theory, measurement invariance

Posted Date: June 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9994486/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

This study examined the psychometric properties of an Arabic version of the Basic Psychological Need Satisfaction and Frustration Scale among pre-service and in-service teachers in Kuwait. Data were collected from 1,755 participants, including 780 pre-service and 975 in-service teachers. Confirmatory factor analysis compared competing models, while reliability, convergent validity, discriminant validity, criterion validity, measurement invariance, and differential item functioning were examined across career stage. The six-factor model showed the strongest fit and was retained. Internal consistency was acceptable across the six subscales, and criterion associations were broadly consistent with Self-Determination Theory. Configural and metric invariance were supported, and partial scalar invariance was established after freeing selected intercepts. Most differential item functioning effects were negligible, although five items showed medium effects. The findings provide initial support for the 23-item Arabic BPNSFS in teacher motivation research, with caution needed for career-stage comparisons.

Introduction

Teacher motivation is increasingly examined through Self-Determination Theory because the theory offers a clear account of how social and institutional contexts may support or undermine professional functioning. Within Basic Psychological Need Theory, autonomy, competence, and relatedness are understood as central psychological needs whose satisfaction is associated with more adaptive functioning, whereas their frustration is associated with poorer adjustment and ill-being (Ryan & Deci, 2020; Vansteenkiste et al., 2020). The Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS) was developed to measure these six dimensions separately: autonomy satisfaction, autonomy frustration, relatedness satisfaction, relatedness frustration, competence satisfaction, and competence frustration (Chen et al., 2015). This distinction is important for teacher research because low satisfaction and active frustration are not simply opposite ends of the same continuum. A teacher may experience limited autonomy support, for example, without necessarily feeling directly controlled or pressured.

Although the BPNSFS has been validated across several languages and contexts, each new cultural and professional setting requires careful psychometric evaluation. Translation alone cannot establish that the same construct is being measured in the same way across groups, particularly when the target population differs from the samples used in earlier validation studies. Zayed et al. (2021) provided important evidence for the Arabic BPNSFS across five Arab countries, but their validation focused on undergraduate students rather than teachers. Evidence from the work domain also indicates that need satisfaction and frustration can be meaningfully assessed in occupational settings, yet such measures still require context-sensitive testing when applied to specific professional groups (Olafsen et al., 2021). This issue is especially relevant in teacher education and school contexts, where pre-service teachers and in-service teachers may interpret autonomy, competence, and relatedness through different institutional experiences.

The present study therefore validates an Arabic version of the BPNSFS among pre-service and in-service teachers in Kuwait. It examines the factorial validity, reliability, convergent validity, discriminant validity, and criterion validity of the scale. It also tests measurement invariance and differential item functioning across career stage. By doing so, the study evaluates whether the Arabic BPNSFS can be used to assess basic psychological need satisfaction and frustration among teachers and whether comparisons between pre-service and in-service teachers are psychometrically defensible. The study was guided by the following research questions:

RQ1: Does the Arabic BPNSFS demonstrate adequate factorial validity, reliability, convergent validity, discriminant validity, and criterion validity among pre-service and in-service teachers in Kuwait?

RQ2: Does the Arabic BPNSFS demonstrate measurement invariance across career stage (pre-service vs. in-service teachers)?

RQ3: Do individual items of the Arabic BPNSFS exhibit differential item functioning across career stage?

Method

Participants and Procedure

Data were collected during the Fall 2025 semester in Kuwait as part of a larger study examining teacher motivation (Authors, under review). The sample comprised 1,755 teachers: 780 pre-service teachers enrolled in undergraduate teacher education programmes at Kuwait University and 975 in-service teachers recruited from public schools across Kuwait. Pre-service teachers were recruited through course instructors who distributed the survey link via Microsoft Teams; in-service teachers were recruited through direct school visits by the research team. Participation was voluntary and anonymous, and all participants provided informed consent prior to completing the survey. Ethical approval was granted by the Head of the Educational Psychology Department at Kuwait University. No formal institutional review board was available; all procedures followed APA ethical guidelines and the principles of the Declaration of Helsinki. Participant demographic characteristics are presented in Table 1.

Table 1
Demographic Characteristics of Pre-Service and In-Service Teachers

Characteristic	**Pre-service Teachers (n = 780)**	**In-service Teachers (n = 975)**
Age (years)	*M* = 22.9, *SD* = 5.07 (18–46)	*M* = 35.1, *SD* = 8.60 (18–65)
Gender		
Male	144 (18.5%)	89 (9.1%)
Female	635 (81.4%)	880 (90.3%)
Prefer not to say	1 (0.1%)	6 (0.6%)
Nationality		
Kuwaiti	691 (88.6%)	768 (78.8%)
Non-Kuwaiti	80 (10.3%)	180 (18.5%)
Prefer not to say	9 (1.1%)	27 (2.8%)
Marital status		
Single	578 (74.1%)	218 (22.4%)
Married	175 (22.4%)	665 (68.2%)
Divorced	18 (2.3%)	68 (7.0%)
Widowed	1 (0.1%)	4 (0.4%)
Prefer not to say	8 (1.0%)	20 (2.1%)
Teaching experience (years)	-	*M* = 10.7, *SD* = 7.75 (0–34)
Teaching level	-	Kindergarten: 22 (2.3%); Elementary: 342 (35.1%); Intermediate: 223 (22.9%); High school: 388 (39.8%)

Note. Teaching experience is reported in years; a value of 0 indicates teachers in their first year of teaching. Percentages may not sum to 100 due to rounding.

Instrument and Adaptation Process

The Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS; Chen et al., 2015) was used to assess the six dimensions of basic psychological need satisfaction and frustration, autonomy, relatedness, and competence, each measured across satisfaction and frustration. The original scale comprises 24 items rated on a five-point Likert scale ranging from 1 (*not true at all*) to 5 (*completely true*).

The scale was adapted into Arabic following established cross-cultural adaptation guidelines (Hambleton et al., 2005). Items were first translated into Arabic by a bilingual researcher, then back-translated into English by an independent bilingual expert. A

panel of experts reviewed both versions to resolve discrepancies and ensure semantic and psychological equivalence with the original. Minor wording adjustments were made to reflect the educational contexts of each group: items were worded to reference the university setting for pre-service teachers and the school setting for in-service teachers, while preserving the psychological meaning of each item. The adaptation process was approved by the Head of the Educational Psychology Department at Kuwait University.

Prior to finalising the scale, an initial confirmatory factor analysis identified one item from the Relatedness Frustration subscale (item 22) as exhibiting a weak standardised loading ($\lambda = .26$), substantially below the acceptable threshold. This item was removed, yielding a final 23-item version used in all subsequent analyses.

Criterion Measures

Three criterion variables were included to assess the criterion validity of the BPNSFS subscales. Occupational commitment was measured using four items adapted from Hackett et al. (2001), rated on a nine-point Likert scale (1 = *strongly disagree*, 9 = *strongly agree*; $\alpha = .82–.88$). Intentions to quit the teaching profession were assessed using three items from the same source, also on a nine-point scale ($\alpha = .87–.91$). Teacher self-efficacy was measured using 12 items from the Teachers' Sense of Efficacy Scale (TSES; Tschannen-Moran & Woolfolk Hoy, 2001), rated on a nine-point Likert scale ($\alpha = .98–.99$).

Data Analysis

All analyses were conducted in R. Confirmatory factor analysis (CFA) was conducted using the *lavaan* package with the robust maximum likelihood estimator (MLR). Model fit was evaluated using the comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA) with 90% confidence intervals, and standardised root mean square residual (SRMR). Thresholds of CFI/TLI ≥ .95, RMSEA ≤ .06, and SRMR ≤ .08 were used to indicate good fit, with CFI/TLI ≥ .90 and RMSEA ≤ .08 indicating acceptable fit (Hu & Bentler, 1999).

Three sets of analyses were conducted to address the three research questions. For RQ1, four competing models were compared: a one-factor model, a two-factor satisfaction/frustration model, a three-factor need-type model, and the theoretically hypothesised six-factor model. A bifactor model was additionally estimated as a supplementary check using the lavaan package. Convergent validity was assessed using average variance extracted (AVE), with values ≥ .50 indicating adequate convergence (Fornell & Larcker, 1981). Discriminant validity was assessed using the heterotrait-monotrait ratio of correlations (HTMT), with values below .85 indicating adequate discriminant validity and values below .90 representing the more lenient threshold (Henseler et al., 2015). Internal consistency was assessed using Cronbach's alpha and McDonald's omega. Criterion validity was examined by inspecting the pattern of correlations between BPNSFS subscales and occupational commitment, self-efficacy, and intentions to quit, separately for each career stage group.

For RQ2, measurement invariance across career stage was tested sequentially: configural, metric, full scalar, and partial scalar invariance. Model comparisons followed the $\Delta CFI \leq .010$ and $\Delta RMSEA \leq .015$ criteria recommended by Chen (2007). Non-invariant intercepts were identified using modification indices and freed sequentially to establish partial scalar invariance.

For RQ3, differential item functioning (DIF) across career stage was examined using the *mirt* package. A constrained multigroup graded response model was estimated as the baseline, and item-level DIF was tested using likelihood ratio tests. Effect sizes were computed as $\chi^2/(2N)$ and classified using the ETS A/B/C framework, where A indicates negligible DIF, B indicates moderate DIF, and C indicates large DIF (Zwick & Ercikan, 1989).

Results

RQ1: Factorial Validity, Reliability, and Validity of the Arabic BPNSFS

Model Comparison

Four competing models were estimated and compared; fit indices are presented in Table 2. The one-factor and three-factor models fit the data poorly, and the two-factor model showed improvement but remained inadequate. The six-factor model showed the strongest overall fit across the reported indices and was clearly superior to the competing models. A bifactor model was additionally estimated as a supplementary check. Although the bifactor model showed a marginally better AIC, the specific factor loadings were largely small and non-significant. The six-factor model was therefore retained for all subsequent analyses.

Table 2
Fit Indices for Competing Confirmatory Factor Analysis Models

Model	χ^2	df	p	CFI	TLI	RMSEA	90% CI	SRMR	AIC
1-factor	5238.71	230	< .001	.632	.595	.131	[.128, .134]	.143	120372.2
2-factor (Sat/Frus)	2309.58	229	< .001	.848	.832	.084	[.081, .087]	.081	116312.9
3-factor (Need type)	5025.90	227	< .001	.651	.611	.128	[.125, .131]	.141	120011.2
6-factor	609.95	215	< .001	.972	.967	.037	[.034, .041]	.029	114007.4
Bifactor	—	—	—	.977	.971	.035	—	.026	113925.4

Note. CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; SRMR = standardised root mean square residual; AIC = Akaike information criterion. χ^2 values are Satorra–Bentler scaled test statistics. All models estimated using MLR with FIML.

Factor Loadings

All 23 items loaded significantly on their intended factors ($p < .001$), with standardised loadings ranging from .611 to .834. Full estimates are presented in Table 3.

Reliability and Convergent Validity

Internal consistency was acceptable to strong across all six subscales. Cronbach's alpha ranged from .769 (Autonomy Satisfaction) to .874 (Competence Satisfaction), and McDonald's omega estimates were closely aligned (range: .770–.874), indicating that alpha did not overestimate reliability. Because items were contextually adapted across career stage groups and partial scalar non-invariance was observed, reliability coefficients are reported for the full sample only. AVE was adequate for three of the six subscales (range: .533–.634). Three subscales, Autonomy Satisfaction (AVE = .457), Autonomy Frustration (AVE = .485), and Competence Frustration (AVE = .463), fell marginally below the .50 threshold; however, omega values for all three subscales exceeded .75, suggesting acceptable construct reliability despite the marginal AVE.

Table 3

Standardised Factor Loadings, Reliability, and Average Variance Extracted for the Six-Factor BPNSFS

Subscale / Item	λ	SE	α	ω	AVE
Autonomy Satisfaction			.769	.770	.457
Item 1	.617	.020			
Item 7	.611	.020			
Item 13	.732	.016			
Item 19	.745	.016			
Autonomy Frustration			.788	.790	.485
Item 2	.623	.022			
Item 8	.694	.019			
Item 14	.756	.017			
Item 20	.707	.019			
Relatedness Satisfaction			.840	.839	.567
Item 3	.651	.019			
Item 9	.822	.016			
Item 15	.792	.015			
Item 21	.752	.017			
Relatedness Frustration			.771	.773	.533
Item 4	.662	.024			
Item 10	.746	.024			
Item 16	.783	.020			
Competence Satisfaction			.874	.874	.634
Item 5	.773	.017			
Item 11	.833	.014			
Item 17	.834	.014			
Item 23	.747	.016			
Competence Frustration			.773	.775	.463
Item 6	.618	.022			
Item 12	.667	.021			
Item 18	.737	.019			
Item 24	.714	.021			

Note. λ = standardised factor loading; SE = standard error; α = Cronbach's alpha; ω = McDonald's omega; AVE = average variance extracted. All loadings significant at $p < .001$. Reliability coefficients are reported for the full sample only; items were contextually adapted across career stage groups and partial scalar non-invariance was observed.

Discriminant Validity

HTMT values indicated adequate discriminant validity for most subscale pairs. Values between satisfaction and frustration subscales within the same need dimension were all well below .85 (range: .045–.382). However, HTMT values among the three

satisfaction subscales were elevated: Autonomy Satisfaction/Relatedness Satisfaction = .907, Autonomy Satisfaction/Competence Satisfaction = .906, and Relatedness Satisfaction/Competence Satisfaction = .885. These values exceed the strict .85 threshold, though two of three fall below the more lenient .90 criterion. This pattern reflects the well-documented positive intercorrelation among need satisfaction dimensions in SDT research and is consistent with findings from other BPNSFS validation studies rather than indicating construct overlap.

Criterion Validity

The criterion validity pattern was theoretically coherent across both career stage groups. Satisfaction subscales correlated positively with occupational commitment and self-efficacy, and negatively with intentions to quit. Frustration subscales showed the opposite pattern, correlating negatively with commitment and positively with quit intentions. Full correlations are presented in Tables 4a and 4b.

Table 4

a. *Correlations Among Study Variables for Pre-Service Teachers*

Variable	1	2	3	4	5	6	7	8	9
1. Autonomy Satisfaction	1.000	.691***	.714***	− .040	− .157***	− .175***	.247***	.499***	− .206***
2. Relatedness Satisfaction	.691***	1.000	.728***	.105**	− .249***	− .186***	.146***	.492***	− .095**
3. Competence Satisfaction	.714***	.728***	1.000	.075*	− .237***	− .305***	.160***	.568***	− .117**
4. Autonomy Frustration	− .040	.105**	.075*	1.000	.212***	.450***	− .273***	.020	.219***
5. Relatedness Frustration	− .157***	− .249***	− .237***	.212***	1.000	.588***	− .206***	− .322***	.123***
6. Competence Frustration	− .175***	− .186***	− .305***	.450***	.588***	1.000	− .316***	− .273***	.270***
7. Commitment	.247***	.146***	.160***	− .273***	− .206***	− .316***	1.000	.081*	− .712***
8. Self-Efficacy	.499***	.492***	.568***	.020	− .322***	− .273***	.081*	1.000	− .063
9. Intentions to quit	− .206***	− .095**	− .117**	.219***	.123***	.270***	− .712***	− .063	1.000

Note. Values represent Pearson correlations. *p < .05, **p < .01, ***p < .001

Table 4

b. *Correlations Among Study Variables for In-Service Teachers*

Variable	1	2	3	4	5	6	7	8	9
1. Autonomy Satisfaction	1.000	.762***	.768***	− .001	− .160***	− .191***	.125***	.478***	− .167***
2. Relatedness Satisfaction	.762***	1.000	.789***	.068*	− .318***	− .236***	.045	.505***	− .109***
3. Competence Satisfaction	.768***	.789***	1.000	.123***	− .234***	− .317***	− .011	.557***	− .050
4. Autonomy Frustration	− .001	.068*	.123***	1.000	.316***	.321***	− .459***	.025	.387***
5. Relatedness Frustration	− .160***	− .318***	− .234***	.316***	1.000	.702***	− .101**	− .265***	.141***
6. Competence Frustration	− .191***	− .236***	− .317***	.321***	.702***	1.000	− .094**	− .355***	.107***
7. Commitment	.125***	.045	− .011	− .459***	− .101**	− .094**	1.000	− .165***	− .729***
8. Self-Efficacy	.478***	.505***	.557***	.025	− .265***	− .355***	− .165***	1.000	.102**
9. Intentions to quit	− .167***	− .109***	− .050	.387***	.141***	.107***	− .729***	.102**	1.000

Note. Values represent Pearson correlations. *p < .05, **p < .01, ***p < .001

RQ2: Measurement Invariance Across Career Stage

Measurement invariance was tested sequentially across pre-service and in-service teachers; full results are presented in Table 5. The configural model confirmed that the six-factor structure holds in both groups. Metric invariance was supported (ΔCFI = .003, ΔRMSEA = .001), indicating equivalent factor loadings across career stage. The full scalar model showed a meaningful decline in fit (ΔCFI = .016, ΔRMSEA = .007), indicating that not all item intercepts are equivalent. Modification indices identified four non-invariant intercepts: item 7 and item 19 (Autonomy Satisfaction), item 8 (Autonomy Frustration), and item 3 (Relatedness Satisfaction), which were freed. The resulting partial scalar model showed a negligible decline relative to the metric model (ΔCFI = .004, ΔRMSEA = .001), supporting partial scalar invariance.

Table 5

Measurement Invariance Sequence Across Career Stage

Model	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA	Conclusion
Configural	.963	.957	.044	.037	—	—	Supported
Metric	.961	.955	.044	.042	.003	.001	Supported
Scalar (full)	.944	.939	.052	.046	.016	.007	Not supported
Scalar (partial)▯	.957	.953	.046	.043	.004	.001	Supported

Note. ΔCFI and ΔRMSEA are relative to the preceding model. Threshold: ΔCFI ≤ .010, ΔRMSEA ≤ .015 (Chen, 2007). ▯ Items 7, 19 (Autonomy Satisfaction), item 8 (Autonomy Frustration), and item 3 (Relatedness Satisfaction) intercepts freed.

RQ3: Differential Item Functioning Across Career Stage

DIF was examined for all 23 items across pre-service and in-service teachers. Likelihood ratio tests indicated statistically significant DIF for 19 of 23 items. However, effect sizes were negligible (ETS Class A) for 18 of the 23 items (effect size range: .0032–.0143), indicating that the statistically significant results reflect the sensitivity of the test to the large sample size rather than meaningful measurement bias. Five items showed medium DIF (ETS Class B): items 7 (effect size = .046) and 19 (.026) from Autonomy Satisfaction, item 3 (.025) from Relatedness Satisfaction, and items 12 (.024) and 6 (.018) from Competence Frustration. No item reached the large DIF threshold (ETS Class C). Four items showed no statistically significant DIF: items 2 and 20 (Autonomy Frustration), item 21 (Relatedness Satisfaction), and item 23 (Competence Satisfaction). Full DIF results are presented in Table 6.

Table 6

Differential Item Functioning Results Across Career Stage (Pre-Service vs. In-Service)

Item	Subscale	χ^2	df	p	Effect size	ETS class
Item 1	Autonomy Satisfaction	16.815	5	.005	.0048	Negligible (A)
Item 7	Autonomy Satisfaction	161.331	5	< .001	.0460	Medium (B)
Item 13	Autonomy Satisfaction	22.757	5	< .001	.0065	Negligible (A)
Item 19	Autonomy Satisfaction	90.856	5	< .001	.0259	Medium (B)
Item 2	Autonomy Frustration	5.562	5	.351	.0016	Negligible (A)
Item 8	Autonomy Frustration	32.038	5	< .001	.0091	Negligible (A)
Item 14	Autonomy Frustration	11.287	5	.046	.0032	Negligible (A)
Item 20	Autonomy Frustration	8.256	5	.143	.0024	Negligible (A)
Item 3	Relatedness Satisfaction	89.168	5	< .001	.0254	Medium (B)
Item 9	Relatedness Satisfaction	15.366	5	.009	.0044	Negligible (A)
Item 15	Relatedness Satisfaction	37.893	5	< .001	.0108	Negligible (A)
Item 21	Relatedness Satisfaction	9.955	5	.077	.0028	Negligible (A)
Item 4	Relatedness Frustration	27.242	5	< .001	.0078	Negligible (A)
Item 10	Relatedness Frustration	14.802	5	.011	.0042	Negligible (A)
Item 16	Relatedness Frustration	12.877	5	.025	.0037	Negligible (A)
Item 5	Competence Satisfaction	23.820	5	< .001	.0068	Negligible (A)
Item 11	Competence Satisfaction	27.288	5	< .001	.0078	Negligible (A)
Item 17	Competence Satisfaction	12.919	5	.024	.0037	Negligible (A)
Item 23	Competence Satisfaction	2.327	5	.802	.0007	Negligible (A)
Item 6	Competence Frustration	63.647	5	< .001	.0181	Medium (B)
Item 12	Competence Frustration	82.986	5	< .001	.0236	Medium (B)
Item 18	Competence Frustration	50.220	5	< .001	.0143	Negligible (A)
Item 24	Competence Frustration	28.438	5	< .001	.0081	Negligible (A)

Note. Effect size = $\chi^2/(2N)$. ETS classification: A = negligible (< .017), B = moderate (.017–.064), C = large (> .064; Zwick & Ercikan, 1989).

Discussion

This study examined the psychometric properties of an Arabic version of the Basic Psychological Need Satisfaction and Frustration Scale among pre-service and in-service teachers in Kuwait. Overall, the findings provide encouraging evidence for the use of the scale in this context. The six-factor model was supported, reliability was acceptable across the retained subscales, and the pattern of criterion associations was broadly consistent with Self-Determination Theory. At the same time, the findings also point to important qualifications. Discriminant validity among the satisfaction subscales was not fully clear, full scalar invariance was not achieved, and several items showed evidence of medium differential item functioning. These results suggest that the Arabic BPNSFS can be used in teacher motivation research in Kuwait, but comparisons across career stage should be made with appropriate psychometric caution.

The support for the six-factor model is an important finding because it is consistent with the theoretical structure of the BPNSFS. The scale was designed to distinguish autonomy, relatedness, and competence satisfaction from autonomy, relatedness, and competence frustration (Chen et al., 2015). The present findings therefore support the view that need satisfaction and need frustration should not be treated simply as opposite ends of the same dimension. This distinction is central to Basic Psychological Need Theory, where the absence of need satisfaction is not equivalent to the active experience of need frustration (Ryan & Deci, 2020; Vansteenkiste et al., 2020). In teacher research, this distinction is particularly useful. A teacher may report limited opportunities for professional choice without necessarily experiencing direct pressure or control. Similarly, a teacher may feel only moderately connected to colleagues without necessarily feeling excluded or rejected. The present findings therefore support the value of retaining separate satisfaction and frustration dimensions when examining teacher motivation in Arabic-speaking educational settings.

The results also extend existing Arabic validation evidence. Zayed et al. (2021) provided important evidence for the Arabic BPNSFS across several Arab countries, but their work focused on undergraduate students. The present study adds evidence from a teacher population and from two career stage groups. This matters because pre-service and in-service teachers are positioned differently within the educational system. Pre-service teachers encounter teaching mainly through university preparation and early professional socialisation. In-service teachers respond to the daily conditions of school work, institutional expectations, workload, and professional accountability. These differences make it necessary to test whether the scale functions adequately across career stage rather than assuming equivalence from prior student-based validation evidence. In this respect, the study contributes to the contextual validation of the BPNSFS for Arabic-speaking teacher education and school contexts.

The reliability evidence was generally acceptable. Alpha and omega estimates showed a consistent pattern across the six subscales, which supports the internal consistency of the retained 23-item version. The convergent validity evidence was more qualified. Some subscales met the expected AVE criterion, whereas Autonomy Satisfaction, Autonomy Frustration, and Competence Frustration fell slightly below the conventional threshold. This does not necessarily undermine the scale, particularly because the reliability estimates remained acceptable. However, it does indicate that these subscales may require further examination in future Arabic teacher samples. The weaker AVE values may reflect the limited number of items per subscale, subtle translation effects, or differences in how autonomy and competence frustration are experienced across university and school settings. These possibilities should be treated as hypotheses for future research rather than conclusions from the present study.

The discriminant validity findings also require a balanced interpretation. The distinction between satisfaction and frustration within the same need dimension was supported. This is theoretically important because it indicates that satisfaction and frustration were not simply mirror images of one another. However, the satisfaction subscales were strongly interrelated. This pattern is understandable within Self-Determination Theory because autonomy, relatedness, and competence satisfaction often co-occur in supportive environments (Ryan & Deci, 2020; Vansteenkiste et al., 2020). For example, teachers who feel trusted to make professional decisions may also feel more competent and more connected to colleagues. Nevertheless, from a measurement perspective, the elevated HTMT values mean that the three satisfaction subscales should be interpreted with caution. The findings support their theoretical distinction, but they also suggest that researchers should avoid overstating their empirical separability in this sample.

The criterion validity evidence was broadly consistent with theoretical expectations. Satisfaction subscales were generally associated with more adaptive professional outcomes, including stronger occupational commitment and teacher self-efficacy,

and lower intentions to quit. Frustration subscales generally showed the opposite pattern. This supports the relevance of the Arabic BPNSFS for studying teacher motivation because the scale related meaningfully to professional variables that matter in teacher education and school systems. The pattern is also consistent with the broader SDT assumption that need-supportive experiences are associated with more adaptive functioning, whereas need-frustrating experiences are associated with less adaptive outcomes (Ryan & Deci, 2020; Vansteenkiste et al., 2020). At the same time, the present findings are correlational. They should not be interpreted as evidence that need satisfaction causes commitment or that need frustration causes intentions to quit.

The measurement invariance findings provide partial support for career stage comparisons. Configural and metric invariance indicated that the same broad factor structure and factor loadings were sufficiently similar across pre-service and in-service teachers. This suggests that both groups understood the general structure of the scale in comparable ways. Full scalar invariance, however, was not supported. After freeing selected intercepts, partial scalar invariance was established. This means that comparisons across career stage may be possible, particularly at the latent level, but they should not be treated as straightforward without qualification. Researchers using the Arabic BPNSFS should report the level of invariance achieved and should avoid making strong claims based on raw mean differences unless the relevant measurement assumptions have been checked. This is especially important because career stage is not a minor demographic distinction. It reflects different institutional locations and professional experiences.

The DIF findings add further nuance to the invariance results. Although many items showed statistically significant DIF, most effects were negligible in magnitude. This distinction is important because large samples can make small differences statistically significant. The use of effect-size classification therefore provided a more informative basis for interpretation than significance testing alone, consistent with the logic of DIF evaluation (Zwick & Ercikan, 1989). The small number of items with medium DIF should still be considered carefully. These items were located in Autonomy Satisfaction, Relatedness Satisfaction, and Competence Frustration. This pattern may indicate that some item wording carries slightly different implications for pre-service and in-service teachers. For instance, autonomy-related items may be interpreted differently by students in teacher education programmes and teachers working in schools. Similarly, competence frustration may be shaped by different sources of pressure across preparation and employment contexts. These interpretations remain tentative, but they identify useful directions for further scale refinement.

A further contribution of the study lies in its adaptation process. The scale was not only translated but also contextually adapted for university and school settings. This approach is consistent with cross-cultural adaptation principles, which emphasise semantic and psychological equivalence rather than literal translation alone (Hambleton et al., 2005). At the same time, contextual adaptation may introduce small differences in item interpretation across groups. The partial scalar invariance and DIF findings are therefore not surprising. They show why validation should be treated as an empirical process rather than a procedural formality. For future use, the adapted Arabic BPNSFS appears suitable for teacher motivation research, but researchers should remain attentive to whether items function similarly across subgroups.

Several practical implications follow from these findings. First, the Arabic BPNSFS can provide researchers with a theoretically grounded instrument for examining teacher need satisfaction and frustration in Kuwait. Second, the distinction between satisfaction and frustration may help researchers move beyond general measures of motivation and identify more specific motivational risks. For example, low competence satisfaction and high competence frustration may have different implications for teacher development. Third, the scale may be useful in studies of teacher commitment, self-efficacy, and intentions to leave the profession. However, its use should remain research-oriented. The present evidence does not support high-stakes individual diagnosis or administrative evaluation of teachers.

The study has several limitations. First, the data were cross-sectional. The observed associations with occupational commitment, self-efficacy, and intentions to quit cannot establish directionality. Longitudinal research is needed to examine whether need satisfaction and frustration predict changes in teacher motivation and retention-related outcomes over time. Second, the study relied on self-report data. Although this is appropriate for assessing psychological need experiences, future work could include additional sources of evidence, such as interviews, institutional indicators, or longitudinal professional outcomes. Third, one item was removed after the initial CFA because of a weak loading. The resulting 23-item version showed

acceptable psychometric properties, but it should be cross-validated in an independent teacher sample. Fourth, although the sample was large and included both pre-service and in-service teachers, it was limited to Kuwait. Further validation is needed across other Arabic-speaking educational contexts.

In conclusion, the present study provides initial evidence for the validity and reliability of a 23-item Arabic BPNSFS among pre-service and in-service teachers in Kuwait. The findings support the six-factor structure of the scale and show broadly coherent associations with relevant professional outcomes. They also indicate that the scale functions similarly across career stage at the configural and metric levels, with partial scalar invariance after freeing selected intercepts. Most DIF effects were negligible, although a small number of items require further attention. The Arabic BPNSFS therefore appears to be a useful measure for teacher motivation research in Kuwait, provided that researchers acknowledge the qualified evidence for discriminant validity and cross-group comparability.

Declarations

Ethics Approval:

This study received ethics approval from the Institutional Review Board (IRB) at Kuwait University, College of Education. No specific approval number was provided. The research was conducted following the ethical guidelines of the Declaration of Helsinki.

Consent to Participate:

All participants provided their informed consent prior to their participation in this study. To protect anonymity, the questionnaire did not request names, student IDs, or contact details, and age was the only non-identifying background item included.

Consent to publish:

Not Applicable

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

AI Use:

An AI-assisted tool (QuillBot) was used to support language editing (e.g., improving clarity, grammar, and concision) and to assist with formatting consistency. No AI tool was used to generate or alter the study data, statistical outputs, or substantive interpretations. All AI-assisted edits were reviewed and verified by the author, who takes full responsibility for the final content.

Author Contribution

F.A. and Y.A. contributed equally to the conception and design of the study. F.A. and Y.A. adapted the instrument, collected the data, conducted the statistical analyses, interpreted the findings, and wrote the main manuscript text. Both authors reviewed and approved the final manuscript.

Data Availability

Data are available from the corresponding author upon request.

References

1. Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, A., Ryan, R. M., Sheldon, K. M., Soenens, B., Van Petegem, S., & Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion*, *39*(2), 216–236. https://doi.org/10.1007/s11031-014-9450-1
2. Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834
3. 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
4. Hackett, R. D., Lapierre, L. M., & Hausdorf, P. A. (2001). Understanding the links between work commitment constructs. *Journal of Vocational Behavior*, *58*(3), 392–413. https://doi.org/10.1006/jvbe.2000.1779
5. Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (Eds.). (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Lawrence Erlbaum. https://doi.org/10.4324/9781410611758
6. 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
7. Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118
8. Olafsen, A. H., Halvari, H., & Frølund, C. W. (2021). The Basic Psychological Need Satisfaction and Need Frustration at Work Scale: A validation study. *Frontiers in Psychology*, *12*, 697306. https://doi.org/10.3389/fpsyg.2021.697306
9. Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860
10. Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. https://doi.org/10.1016/S0742-051X(01)00036-1
11. Vansteenkiste, M., Ryan, R. M., & Soenens, B. (2020). Basic psychological need theory: Advancements, critical themes, and future directions. *Motivation and Emotion*, *44*, 1–31. https://doi.org/10.1007/s11031-019-09818-1
12. Zayed, K. N., Omara, E. N., Al-Rawahi, N. Y., Al-Shamli, A. K., Al-Atiyah, A. A., Al-Haramleh, A. A., Azab, M. S., Al-Khasawneh, G. M., & Hassan, M. A. (2021). Psychometric properties of the Arabic version of the Basic Psychological Needs Satisfaction-Frustration Scale (BPNSFS). *BMC Psychology*, *9*. https://doi.org/10.1186/s40359-020-00506-1. Article 15.
13. Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the NAEP history assessment. *Journal of Educational Measurement*, *26*(1), 55–66. https://doi.org/10.1111/j.1745-3984.1989.tb00318.x