Prioritising Staff Support in International Schools Using Importance–Performance Analysis

Issah Quarmson*., Maria Sonia Justin Raj

Westford University College and Universidad Católica San Antonio de Murcia (UCAM)

***Correspondence Author**

DOI: https://doi.org/10.47772/IJRISS.2026.101100010

Received: 03 November 2026; Accepted: 09 November 2026; Published: 19 November 2026

ABSTRACT

School leaders need a way to identify staff concerns without treating every low survey rating as an instruction to spend. This article applies importance–performance analysis to a survey of 52 expatriate educators in Guangzhou's international high schools. Performance is the mean rating of each of six job satisfaction facets, and derived importance is the facet's correlation with affective commitment. The analysis works from the summary statistics of the original study, since the respondent-level file is no longer held; new calculations are confined to sensitivity checks on the cut-offs and the importance measure, analytic confidence intervals, algebraic reconstruction of selected item correlations and the rescaling of reported variance shares. Leadership, development and progression, and compensation combine below-average ratings with above-average associations with commitment under both mean and median cut-offs, although compensation sits close to both dividing lines. Every facet mean lies above the midpoint of the six-point scale, so 'low' here means low relative to the other facets. At item level, the two lowest-rated statements concern clear career progression (M = 3.67) and the clarity of leadership communication (M = 3.75), while immediate supervisors were rated favourably (M = 4.40). The contrast suggests a particular concern about how leadership decisions are communicated to staff. The map is a diagnostic shortlist, not a causal ranking of interventions or an estimate of financial return. Schools should test such maps against staff consultation, measurement checks and small, evaluated changes before committing resources.

Keywords: importance–performance analysis; international schools; expatriate educators; job satisfaction; affective commitment; human resource management

INTRODUCTION

Staff surveys tell school leaders which aspects of employment are rated favourably and which attract concern. They say much less about what to do next. A low score may point to a real problem, but it does not show whether a particular response would strengthen attachment, reduce departures or justify its cost. Leaders need a way of organising survey evidence that keeps the limits of that evidence in view. Importance–performance analysis places attributes on two dimensions, their performance and their importance (Martilla and James, 1977). Applied to employment conditions, the display separates relatively low-rated conditions that are strongly associated with an organisational outcome from conditions that are already rated well. The quadrants are appealing because they reduce a complicated set of comparisons to a picture. What the picture means for management, however, depends on how importance is defined and where the dividing lines are drawn.

In international schools the question touches both human resource management and educational continuity. Supportive leadership has been linked to teacher retention in international schools (Mancuso, Roberts and White, 2010). Odland and Ruzicka (2009) found leadership, compensation and personal circumstances in teachers' accounts of leaving, and Mo and Morris (2024) describe dissatisfaction with leadership changes, workload, employment packages and professional development among international high school teachers in China. These studies show where to look. They cannot say what a given intervention will achieve in a particular school.

This article focuses on expatriate educators and on a practical question: where should school leaders look first when investigating staff-support needs? It uses summary statistics from the same survey of 52 expatriate educators analysed by Quarmson and Justin Raj (2026), with affective commitment as the outcome for the present importance–performance analysis. Three questions guide the analysis. Which facets combine relatively low ratings with relatively strong associations with commitment? How sensitive is that classification to the decision rule? And what action can reasonably follow from the map? The contribution is an exploratory diagnostic application whose conclusions are bounded by the evidence available.

Background

Satisfaction and attachment

Job satisfaction concerns employees' evaluations of their work, whereas affective commitment concerns their emotional attachment to the organisation (Spector, 1997; Meyer and Allen, 1997). The difference matters because teachers can be content with their classroom work and still feel little attachment to the school that employs them. Commitment is chosen here as a relevant organisational outcome. It is not treated as equivalent to retention, and the analysis does not assume that satisfaction affects departure only through commitment. Research on teachers places satisfaction within working conditions, belonging and professional experience (Skaalvik and Skaalvik, 2011; Toropova, Myrberg and Johansson, 2021). For expatriate educators, the six facets examined here are compensation and benefits, professional development and progression, leadership and supervision, meaningful work and autonomy, adjustment support, and work–life balance. Each offers a different point of organisational action. An overall satisfaction score can hide a specific concern, such as unclear progression within an otherwise acceptable job.

What derived importance means

Direct importance asks staff how much they value an attribute; derived importance estimates its association with a chosen outcome. This paper uses the Pearson correlation between each facet and affective commitment. A higher correlation indicates a stronger linear association in this sample. It does not show that staff value the attribute more, that its effect is independent of the other facets, or that changing it would raise commitment. The distinction between conditions of employment and sources of motivation also shapes interpretation. Herzberg, Mausner and Snyderman (1959) argued that some aspects of work relate more to dissatisfaction and others to satisfaction. The present data do not test that argument. A high mean paired with a modest correlation could reflect restricted variation, measurement limitations, overlap with other facets or other processes, and a ceiling effect can only be judged from the distribution, not from the mean.

From association to a decision

Correlated predictors share explanatory information. Relative-weight analysis divides a model's explained variance among its predictors and gives a different description from their separate correlations (Johnson, 2000). A predictor's rescaled weight is a percentage of the model R^2, not a percentage of the outcome's total variance, and it is not an intervention effect. Neither correlations nor variance shares take account of cost, contractual obligations or feasibility. Those considerations enter once the map has identified what to investigate.

METHOD

Setting, sample and records

The data come from an anonymous online survey of expatriate educators working in international high schools in Guangzhou, collected between 14 September and 16 October 2026 as the quantitative phase of a mixed-methods doctoral study. Quarmson and Justin Raj (2026) report a separate analysis of the same sample. A population sampling frame was unavailable, so participants were recruited through professional networks and teachers' social-media groups, with a snowball invitation to eligible colleagues. Because the invitation travelled through open networks, the number of people who received it is unknown and no response rate can be calculated.

Fifty-two responses were received and all passed an embedded attention check. All respondents provided complete responses to the substantive items used in this analysis, so estimates rest on n = 52 unless stated otherwise. Respondents worked in fully international schools (n = 21, 40.4%), Chinese internationalised schools (n = 16, 30.8%) and international private schools (n = 12, 23.1%); three did not state their school type. To protect anonymity the survey did not ask respondents to name their school, so the number of schools represented cannot be counted. The original study summaries record that most respondents were aged 26 to 45 (69.2%), 57.7% were female, half had been at their current school for less than three years, and 23 countries of origin were represented. The study's sub-Saharan African category, which excluded South Africa, accounted for 51.9% of respondents.

The sample size was set by the parent study rather than by a prospective power calculation for this facet analysis. Its planning considered the mediation simulations of Fritz and MacKinnon (2007), in which a bias-corrected bootstrap test achieved 80% power with approximately 54 respondents when the first path was large and the second medium. That benchmark does not establish adequacy for the present six-facet analysis. Here the main issue is precision: under a Fisher z approximation, n = 52 provides about 80% power to detect a population correlation of .38 against zero at $\alpha = .05$, two-tailed (Cohen, 1988). This is a sensitivity calculation, not a post hoc demonstration of adequate power. The observed differences among facet correlations are interpreted descriptively rather than as established differences in their population strengths.

Non-probability recruitment may have favoured educators active in professional networks and excluded some of those who had already left. Selection could affect both means and correlations, but its direction and magnitude cannot be established from these records. Neither the means nor the ranking of facets can be assumed to hold for the wider population.

This article uses retained summary records from the original survey. Published analyses of the same sample report construct-level descriptive statistics and correlations (Quarmson and Justin Raj, 2026) and comparisons with meta-analytic benchmarks (Quarmson, 2026). The facet-level and item-level summaries used here are reproduced in the tables and appendix. The facet correlations, bootstrap intervals and relative-weight results were calculated from the original survey responses; they are retained results, not new analyses of reconstructed participant data. The respondent-level dataset is no longer available, so the original analyses could not be independently recomputed. New calculations are limited to quantities derivable from the retained summaries.

Measures and scoring

Job satisfaction was measured with twelve items adapted from the facet structure of Spector's (1997) Job Satisfaction Survey. Eleven items form six facets: compensation and benefits (B1, B2), professional development and progression (B3, B4), leadership and supervision (B5, B6), meaningful work and autonomy (B7, B8), adjustment support (B9) and work–life balance (B10, B11). The twelfth item (B12) is a global rating of overall satisfaction and belongs to no facet. The global item contributes to the overall satisfaction score but is excluded from the facet analysis. Facet scores are the means of their items. The full wording of every item is given in Appendix A.

Affective commitment was measured with three items from Meyer and Allen's (1997) three-component scales (C1 to C3; for example, 'I feel a strong sense of belonging at this school'), and turnover intention with the three-item scale of the Michigan Organizational Assessment Questionnaire (Cammann et al., 1979; D1 to D3; for example, 'I often think about leaving this school'). All items used the same six-point agreement scale: 1 = strongly disagree, 2 = disagree, 3 = slightly disagree, 4 = slightly agree, 5 = agree, 6 = strongly agree. Every item is positively worded with respect to its construct, so no item was reverse-scored; higher turnover intention scores indicate a stronger intention to leave. Construct scores are item means. Cronbach's alpha was .89 for the twelve-item satisfaction scale, .89 for affective commitment and .84 for turnover intention (n = 52; Quarmson and Justin Raj, 2026). Affective commitment averaged 4.08 (SD = 1.15) and turnover intention 3.58 (SD = 1.27). Facet-level internal-consistency coefficients were not reported; facets are used descriptively.

The short facet measures were a trade-off between coverage and respondent burden. The original study records describe pre-testing with experienced international educators and a questionnaire taking approximately twelve to fifteen minutes. Adaptation from Spector (1997), context-specific wording and pre-testing support content

relevance, but do not establish the validity of the modified facet structure. The reported twelve-item alpha of .89 describes internal consistency across the satisfaction items; it does not establish unidimensionality or the reliability of each facet. Two-item measures can be assessed using the Spearman–Brown coefficient (Eisinga, te Grotenhuis and Pelzer, 2013), but facet-level reliability was not directly verified against original item-level records here. Single items can be adequate for narrow, concrete constructs (Wanous, Reichers and Hudy, 1997; Fisher, Matthews and Gibbons, 2016), but one adjustment-support item provides limited coverage of daily living, social interaction and work. Measurement limitations may therefore contribute to its observed association with commitment.

Analysis

Facet performance is the facet mean, and importance is the facet's correlation with affective commitment. The facet correlations, bootstrap intervals (reported as based on 2,000 resamples) and relative-weight decompositions were produced in the original survey analysis and are reproduced from retained summaries. Original software settings and respondent-level data were unavailable for independent verification. This article also reports approximate 95% confidence intervals calculated from the reported correlations and n = 52 using Fisher's (1921) z transformation. Quarmson (2026) describes the application and limitations of this procedure for small-sample correlations. These intervals assume independent observations and approximate bivariate normality; they do not independently validate the original bootstrap results.

Where a two-item facet was scored as the unweighted mean of its items, an unreported item–outcome correlation was reconstructed algebraically from the facet standard deviation, facet–outcome correlation, item standard deviations and the other item–outcome correlation. This calculation assumes identical respondents and scoring across the summaries. Because the inputs were rounded to two decimal places, reconstructed correlations are approximate and are distinguished from directly reported values in Table 4. No new significance tests were conducted for these reconstructed values.

$$r_{missing,Y} = \frac{2s_{facet}r_{facet,Y} - s_{known}r_{known,Y}}{s_{missing}}$$

Here Y denotes the outcome, s a standard deviation, "known" identifies the item with a reported outcome correlation, and "missing" identifies its paired item. Rounding-sensitivity ranges were calculated by varying each input independently by ±.005; these are arithmetic ranges, not confidence intervals.

Observations are treated as independent. Respondents in the same school may resemble one another, but because school identity was not collected, that dependence could not be estimated or modelled. Positive within-school dependence could lead to underestimated uncertainty; its magnitude could not be assessed without school identifiers.

The performance cut-off is the unweighted mean of the six facet means, 4.2317, and the importance cut-off is the mean of the six correlations, 0.5367. A facet is classed as lower-performing and higher-associated when its mean falls below the first and its correlation above the second. Four sensitivity checks follow. The first replaces both cut-offs with medians (4.20 and 0.52). The second uses the scale midpoint, 3.50, for performance while keeping the mean importance cut-off. The third replaces commitment with turnover intention as the outcome, taking the sign-reversed correlation of each facet with turnover intention as importance (mean cut-off 0.1767). The fourth uses each facet's relative-weight share for commitment as importance (mean cut-off 16.7%), which allows for the overlap among facets. Facet correlations with normative or overall commitment are not reported in the available sources, so these could not serve as alternative outcomes; affective commitment is, in any case, the component most closely related to job satisfaction (Meyer et al., 2002; r = 0.76 in Table A2). Because the original data cannot be re-analysed, the Fisher intervals also serve as an alternative analytic comparison based on the same reported correlations. These checks examine the decision rule; they are not estimates of sampling stability. Finally, reported relative-weight percentages are multiplied by the reported R^2 to show their contribution to total outcome variance. No new regression, mediation or significance test was carried out.

RESULTS

Facet ratings and associations

Table 1 Facet summaries (n = 52)

Facet	Mean	SD	Agree %	r with AC	Bootstrap 95% CI	Fisher 95% CI	r with TI
Leadership and supervision	4.08	1.16	73	0.69	[0.48, 0.82]	[0.51, 0.81]	−0.31
Development and progression	3.88	1.29	62	0.67	[0.47, 0.81]	[0.49, 0.80]	−0.30
Compensation and benefits	4.15	1.08	77	0.55	[0.33, 0.73]	[0.33, 0.72]	−0.23
Meaningful work and autonomy	4.66	0.84	85	0.49	[0.22, 0.73]	[0.25, 0.67]	−0.17
Adjustment support	4.37	1.24	83	0.48	[0.19, 0.69]	[0.24, 0.67]	−0.16
Work–life balance	4.25	1.08	77	0.34	[−0.03, 0.64]	[0.07, 0.56]	0.11

Note. AC = affective commitment; TI = turnover intention. Means, SDs, agreement rates, correlations and bootstrap intervals are retained original-study summaries. Agreement is the percentage with a facet mean of 4 or above. Fisher intervals were calculated for this article from the reported r and n = 52. Original-study summaries reported the facet correlations with TI of −0.30 and −0.31 as significant at $p < 0.05$. The original analyses were not independently recomputed for this article.

Meaningful work and autonomy has the highest mean (4.66), followed by adjustment support (4.37), work–life balance (4.25), compensation (4.15), leadership (4.08) and development and progression (3.88). Every mean lies above the scale midpoint of 3.50. ‘Lower performance’ in the map therefore describes position within this set of facets, not average disagreement or failure against an external standard.

Leadership and development have the strongest associations with commitment (.69 and .67), followed by compensation (.55), meaningful work and autonomy (.49), adjustment support (.48) and work–life balance (.34). These are point-estimate rankings. The two sets of intervals agree closely for the five strongest facets. For work–life balance they differ: the bootstrap interval includes zero, whereas the Fisher interval (.07 to .56) does not. The difference is a reminder that, with 52 respondents, the weakest association is estimated with little precision. Every interval spans the mean importance cut-off, and the intervals cannot by themselves show whether two correlations differ.

The priority map and threshold sensitivity

Leadership, development and progression, and compensation all fall below the mean performance cut-off and above the mean importance cut-off (Figure 1). Compensation meets the rule, but only just: its mean is 0.08 below the performance cut-off and its correlation 0.013 above the importance cut-off. A small change in either estimate could move it to another quadrant.

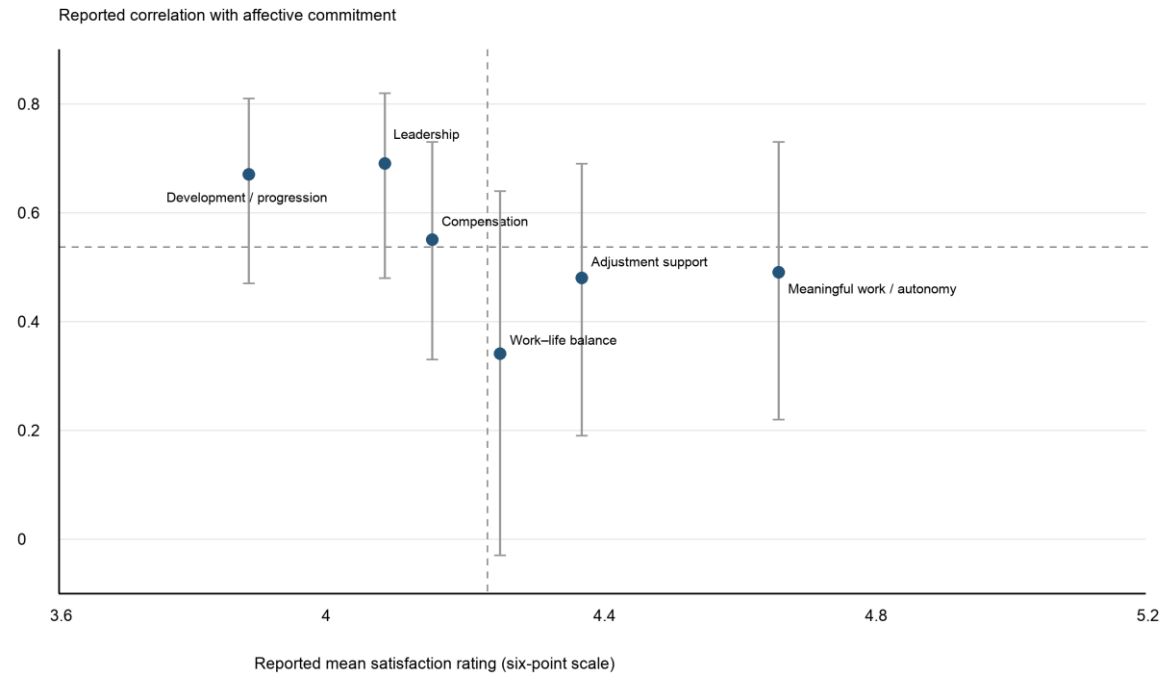


Figure 1 Facet map. Points are facet means and their correlations with affective commitment; vertical bars are the bootstrap intervals from the original analysis. Dashed lines mark the mean cut-offs, 4.2317 and .5367.

Table 2 Sensitivity of the priority classification to the decision rule and importance measure

Rule	**Performance cut-off**	**Importance cut-off**	**Lower rating, higher association**
Means of the six facets	4.2317	0.5367	Leadership; development; compensation
Medians of the six facets	4.2000	0.5200	Leadership; development; compensation
Scale midpoint for performance	3.5000	0.5367	None
Turnover intention as outcome	4.2317	0.1767	Leadership; development; compensation
Relative-weight share as importance	4.2317	16.7%	Leadership; development

The median rule picks out the same three facets. The midpoint rule picks out none, because every facet mean exceeds 3.50. The two kinds of rule answer different questions. A relative rule finds the least favourable ratings within the observed set; a scale rule asks whether ratings fall on the disagreeing side of the response range. Neither is an established standard of adequate employment conditions, and neither comparison shows that the classification would hold in another sample. The same three facets are selected when turnover intention replaces affective commitment in this sensitivity check. Using relative-weight shares as importance retains leadership and development but drops compensation, whose share (16%) sits just below the mean (16.7%). Compensation is therefore the facet whose priority status depends most on the analytic choices.

Model fit and variance shares

In the original analysis the six facets together accounted for 60% of the variance in affective commitment (R^2 = .60) and 19% of the variance in turnover intention (R^2 = .19). Table 3 converts the reported commitment shares into contributions to total variance. Leadership's 32% share of an R^2 of .60, for example, corresponds to about 19% of the total variance in commitment, not 32%.

Table 3 Relative weights for affective commitment and their contribution to total variance

Facet	Share of R^2 (%)	Contribution to total variance
Leadership and supervision	32	0.192
Development and progression	26	0.156
Compensation and benefits	16	0.096
Meaningful work and autonomy	13	0.078
Adjustment support	8	0.048
Work–life balance	5	0.030
Total	100	0.600

Note. Shares are the rescaled relative weights (Johnson, 2000) reported by the original analysis; contributions are the products of the rounded shares and $R^2 = .60$.

Leadership and development together carry 58% of the explained variance in commitment, and adding compensation raises this to 74%. For turnover intention the original analysis reported shares of 30% for leadership, 23% for development and 24% for work–life balance, leaving 23% across the other three facets. Because relative weights carry no sign, the 24% for work–life balance does not mean that better work–life conditions raise intentions to leave; its simple correlation with turnover intention was small (.11) and not significant. None of these figures describes what an intervention would achieve.

Item-level results

Table 4 Satisfaction items (n = 52 for every item)

Item	Content	Facet	Mean	SD	Agree %	r AC	r TI
B1	Fairly paid for the work	Compensation	4.31	1.09	79	0.3	−0.11†
B2	Benefits and package adequate	Compensation	4.00	1.33	73	0.5	−0.28
B3	Meaningful professional development	Development	4.10	1.43	71	0.51†	−0.18†
B4	Clear career progression opportunities	Development	3.67	1.40	65	0.71	−0.37
B5	Supervisor competent and fair	Leadership	4.40	1.24	85	0.51†	−0.27
B6	Leadership communicates decisions clearly	Leadership	3.75	1.44	65	0.67	0.13
B7	Work personally meaningful	Meaningful work	4.63	1.03	90	0.58†	0.17
B8	Appropriate professional autonomy	Meaningful work	4.69	1.00	87	0.23	0.15
B9	Adjustment to Guangzhou supported	Adjustment	4.37	1.24	83	0.48‡	−0.16

B10	Workload manageable within contracted hours	Work–life	4.21	1.29	73	0.22	0.10 / 0.09
B11	Personal and family time respected	Work–life	4.29	1.11	81	0.41†	0.10 / 0.09
B12	Overall satisfied with the job	Global	4.48	1.24	85	0.34	0.9 / 0.11

Note. Means, SDs, agreement (scores of 4 to 6) and unmarked correlations are retained original-study summaries; full wording is in Appendix A. AC = affective commitment; TI = turnover intention. † Approximately reconstructed from rounded summary statistics using the two-item mean-score identity in Section 3.3, not independently estimated from respondent-level data. ‡ B9 is the sole adjustment-support item, so its correlation equals the facet correlation in Table 1. The two work–life items correlated .10 and .09 with TI, but the retained source does not identify which value belongs to which item. n.r. = not reported and not recoverable from the available summaries.

Approximate rounding-sensitivity ranges for reconstructed correlations are: B1–TI [−0.134, −0.094]; B3–AC [0.491, 0.537]; B3–TI [−0.197, −0.161]; B5–AC [0.488, 0.539]; B7–AC [0.554, 0.598]; and B11–AC [0.385, 0.427]. These ranges describe input rounding only and are not confidence intervals.

Two items stand out. Career progression (B4) is the lowest-rated statement (M = 3.67), the satisfaction item with the strongest reported association with turnover intention (r = −0.37), and, in the original analysis, strongly associated with commitment (r = 0.71). Leadership communication (B6) is the second-lowest (M = 3.75; r with commitment = 0.67). About a third of respondents disagreed to some degree with each (Figure 2).

The item detail sharpens the facet picture in two ways. Within leadership, immediate supervisors were rated favourably (B5, M = 4.40, 85% agreeing), whereas communication of decisions received a lower rating (B6, M = 3.75, 65%). This contrast suggests a particular concern about communication, without establishing that it exhausts teachers' concerns about leadership. Within development, access to meaningful professional development (B3, M = 4.10) was rated better than a visible route to promotion (B4, M = 3.67). At the other end, the highest-rated items were professional autonomy (B8, M = 4.69) and meaningful work (B7, M = 4.63), and the work–life items drew majority agreement (73% and 81%).

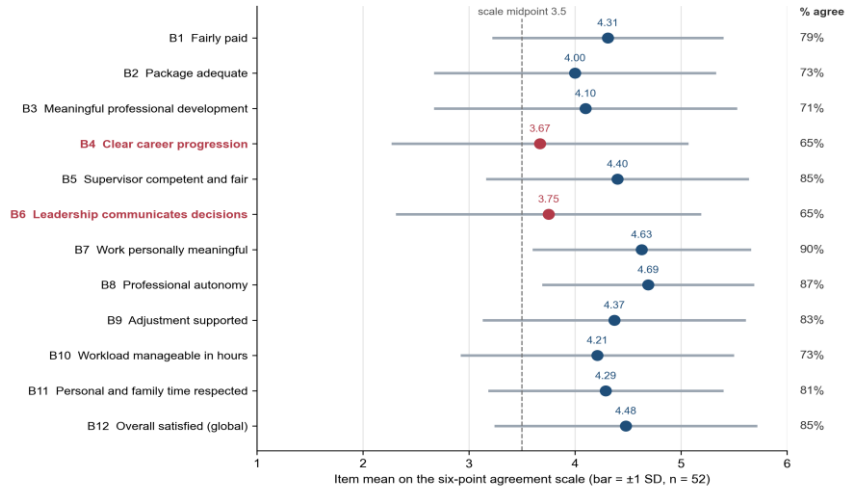


Figure 2 Mean rating of each satisfaction item with ±1 SD and the percentage agreeing (n = 52). Source: retained original-study item summaries reproduced in Table 4. Highlighted items are the two lowest-rated statements.

DISCUSSION

A shortlist for enquiry

The map directs attention to leadership, development and compensation. Leadership and development combine relatively low ratings with the strongest associations with commitment, and compensation meets the same rule, if narrowly. Including compensation corrects a reading that would select only the two most visually striking facets. The sensible response is to investigate these concerns locally, not to fix a spending order.

The item results turn the facet shortlist into sharper questions. For leadership, the contrast between supervisors and senior communication suggests asking staff which decisions are poorly explained, whether explanations arrive in time, and how uncertainty affects their work. For development, the gap between training (4.10) and progression (3.67) suggests separating access to development from transparent promotion criteria, recognition and access to responsibility. The gap motivates that distinction; it does not show that clearer progression would achieve more than additional training.

Compensation deserves the same care. Its position near the cross-hairs makes a yes-or-no instruction unhelpful. The package item (B2, M = 4.00) was rated lower than fairness of pay (B1, M = 4.31) and was associated with turnover intention ($r = -0.28$), which suggests looking at the completeness and consistency of the package, including housing and allowances, alongside salary itself. Staff accounts would be needed to decide whether the concern is the amount, the components or how changes are communicated.

Protecting favourable conditions

Meaningful work and autonomy, adjustment support and work–life balance sit in the higher-rating, lower-association region under the mean rule. Calling that region ‘possible overkill’, as quadrant labels often do, would imply knowledge about spending that was never collected. These conditions may be necessary, valued or unevenly experienced even when their averages are favourable. The result supports keeping them in place while the less favourable areas are investigated; it gives no basis for reducing them.

A high mean is also not a ranking of what teachers value. Agreement with a statement about one’s experience is different from a judgement of what matters most. Restricted variation could weaken an observed association, and the standard deviations of the two highest-rated items (1.03 and 1.00) are the smallest among the satisfaction items, which is consistent with some restriction. Without the full response distributions, however, a ceiling explanation cannot be confirmed.

Workload and retention

The work–life results call for restraint. The correlation with commitment is the weakest of the six, its intervals are wide, and the two work–life items had small observed correlations with turnover intention (0.10 and 0.09). These estimates do not establish that workload is irrelevant or that favourable conditions encourage exit. The items asked about hours and personal time, while workload can also involve intensity and administrative demands. Further item development and staff consultation would be needed to determine which aspects of workload matter in a particular school.

Commitment, intention to leave and actual departure should be kept apart. Staff may feel attached and still leave for family reasons or a better opportunity; others may stay despite weak attachment because moving is difficult. The larger R^2 for commitment than for turnover intention does not, on its own, establish that commitment carries the effect of satisfaction on leaving.

Comparison with other samples and settings

The prominence of leadership is consistent with evidence from other countries and types of school. In international schools, supportive leadership has been linked to retention in the Near East and South Asia region (Mancuso, Roberts and White, 2010) and featured prominently in teachers’ accounts of leaving (Odland and Ruzicka, 2009). In public schools, perceptions of school leadership were the working condition most predictive

of planned and actual departure in North Carolina (Ladd, 2011), principal leadership and collegial relationships mattered more than physical resources in Massachusetts (Johnson, Kraft and Papay, 2012), inadequate administrative support recurs in national analyses of teacher turnover in the United States (Ingersoll, 2001), and belonging and working conditions shape satisfaction in Norway and Sweden (Skaalvik and Skaalvik, 2011; Toropova, Myrberg and Johansson, 2021). Compensation also warrants attention in this setting because expatriate packages can combine salary, housing and allowances, and dissatisfaction with employment conditions has been reported among international high school teachers in China (Mo and Morris, 2024). These comparisons concern direction rather than size, because the studies differ in measures, outcomes and populations; replication with the same instrument in other countries and types of institution is needed to show whether the shortlist travels.

Implications For School Management

The practical value of the map lies in structuring a decision. A school should first define the decision it faces and check that its survey items describe conditions staff actually experience. It should then look at participation, missing responses, scoring and uncertainty before naming candidate concerns. Conversations with staff and a review of relevant policies can then establish what an unfavourable rating means. Cost, feasibility, professional obligations and possible side effects should guide the choice of response.

A feasible pilot might publish clear progression criteria or introduce a consistent routine for explaining major decisions: what was decided, why, and what it means for staff. These are examples to test, not interventions this study shows to work. Before acting, leaders should state what they expect to change, when it will be measured and who is responsible. Repeated measurement and a comparison group, where feasible, would strengthen any evaluation, and a change in ratings would still need to be read alongside changes in who responded. Table 5 translates the shortlist into candidate actions, each paired with indicators of staff morale, commitment and retention that a school can collect itself.

Table 5 Candidate actions and indicators for evaluating them

Priority	**Example action to pilot**	**Indicators**	**When to measure**
Leadership communication (B6)	Explain each major decision: what was decided, why, and what it means for staff	B6 mean and % agreeing	Baseline, then each term
Career progression (B4)	Publish promotion criteria; hold an annual career conversation with each teacher	B4 mean and % agreeing; share of posts filled internally	Baseline and after 12 months
Compensation package (B2)	Review package components and explain them in writing	B1 and B2 means; package reasons in exit interviews	Before contract renewal
Morale and commitment	Annual survey with the same items and a mid-year pulse check	Global satisfaction (B12) for morale; affective commitment (C1–C3)	Annually and mid-year
Retention	Track leaving intentions and departures	Turnover intention (D1–D3); voluntary turnover and contract renewal rates	Annually

Note. Item codes refer to Appendix A. The actions are examples to pilot and evaluate, not interventions shown to work. Voluntary turnover rate = voluntary leavers in a year divided by average staff numbers. Leaders should establish a baseline in their own school and specify the change expected. The 65% agreement with B6 in this study is a pooled sample percentage, not a baseline for any individual school.

The level of analysis must match the decision. A map pooled across schools may reflect differences between schools rather than patterns within any one of them, and it cannot set priorities for an individual school. No respondent count makes correlations reliable in itself; adequacy depends on the precision needed, the quality of measurement and the dependence among observations. Where the evidence is thin, the right output is a clearly bounded question for further enquiry.

Limitations

The analysis rests on retained summary statistics. The respondent-level file is no longer held, so the original facet correlations with commitment, bootstrap intervals and relative weights could not be independently recomputed. Full regression coefficient tables, the predictor correlation matrix and model diagnostics were unavailable. Four item-level correlations with affective commitment were directly reported; four more were approximately reconstructed, and the adjustment-support item correlation equals its single-item facet correlation. These reconstructions depend on the accuracy and consistency of the original summaries and provide no independent verification. Fisher intervals likewise depend on the reported correlations and assume independent observations. The article and the two related publications use the same survey sample, so they do not constitute independent replications (Quarmson and Justin Raj, 2026; Quarmson, 2026).

The sample is small, recruited through networks and cross-sectional. It cannot be assumed to represent all expatriate educators in Guangzhou, and its regional composition, with a majority classified in the study's sub-Saharan African category, differs from that of many international-school workforces. Because all variables were measured once, the associations cannot establish that satisfaction influences commitment; commitment may equally shape how staff evaluate their conditions. Shared self-report measurement, omitted variables and dependence within schools may also contribute to the associations. School identity was not collected, so clustering could not be assessed.

Most facets have two items and adjustment support has one, limiting measurement coverage. Facet-level reliability coefficients were not reported and could not be independently verified from the original item-level records. Internal consistency cannot be estimated for the single adjustment-support item. A single indicator offers limited coverage of a multidimensional experience and is more exposed to idiosyncratic readings of its wording, so adjustment support should be interpreted with particular caution. The intervals do not establish a firm ranking of facets, and the threshold checks examine decisions based on rounded point estimates rather than sampling stability. Costs, later departures and the effects of interventions were not observed, so neither an optimal allocation of resources nor a causal claim about retention is supported.

CONCLUSION

Leadership, development and progression, and compensation combine relatively low ratings with relatively strong associations with affective commitment under both mean and median rules. At item level, unclear career progression and poorly explained senior decisions are the most specific concerns, while immediate supervision, meaningful work and autonomy are rated well. Favourable ratings do not show that a condition is unimportant or over-resourced. The defensible contribution is a transparent diagnostic map that makes both the priorities and the decision rules visible. Using it well requires sound measurement, consultation with staff and honest evaluation of whatever changes follow.

Declarations

Ethics approval and consent. The study was reviewed on behalf of the DBA ethics committee of Westford University College, delivered with Universidad Católica San Antonio de Murcia (UCAM), and approved on 1 July 2026 as a Category 3 project, normal primary data collection with minimal risk (reference WU-UCAM-DBA-2026-047), before any participant was approached. Participants received a plain-language information sheet and gave informed consent through a statement at the start of the anonymous questionnaire. The survey collected no names or school names.

Relationship to other work. This article uses the same survey sample of 52 expatriate educators analysed by Quarmson and Justin Raj (2026) and Quarmson (2026). Those publications examine constraint and turnover intention, and the benchmarking of small-sample correlations, respectively. The present article examines importance–performance mapping of employment facets and the sensitivity of the resulting priority classifications. The shared sample and overlapping descriptive statistics mean that these articles are not independent replications. The same survey also underlies companion manuscripts on career progression, affective commitment and retention measurement.

Funding. None.

Conflict of interest. The first author is a teacher and examiner in Guangzhou's international school sector; this position was disclosed to participants. The authors declare no other interests.

Data availability. The summary statistics used in this analysis are presented in the article and appendix. Related construct-level results are reported by Quarmson and Justin Raj (2026), and benchmarking calculations are reported by Quarmson (2026). The respondent-level dataset is no longer available and cannot be shared. Algebraically reconstructed item correlations are identified in Table 4.

REFERENCES

1. Cammann, C., Fichman, M., Jenkins, D. and Klesh, J. (1979) The Michigan Organizational Assessment Questionnaire. Unpublished manuscript. Ann Arbor, MI: University of Michigan.
2. Cohen, J. (1988) Statistical Power Analysis for the Behavioral Sciences. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum.
3. Eisinga, R., te Grotenhuis, M. and Pelzer, B. (2013) 'The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown?', International Journal of Public Health, 58(4), pp. 637–642. https://doi.org/10.1007/s00038-012-0416-3
4. Fisher, G.G., Matthews, R.A. and Gibbons, A.M. (2016) 'Developing and investigating the use of single-item measures in organizational research', Journal of Occupational Health Psychology, 21(1), pp. 3–23. https://doi.org/10.1037/a0039139
5. Fisher, R.A. (1921) 'On the "probable error" of a coefficient of correlation deduced from a small sample', Metron, 1, pp. 3–32.
6. Fritz, M.S. and MacKinnon, D.P. (2007) 'Required sample size to detect the mediated effect', Psychological Science, 18(3), pp. 233–239. https://doi.org/10.1111/j.1467-9280.2007.01882.x
7. Herzberg, F., Mausner, B. and Snyderman, B.B. (1959) The Motivation to Work. 2nd edn. New York: Wiley.
8. Ingersoll, R.M. (2001) 'Teacher turnover and teacher shortages: an organizational analysis', American Educational Research Journal, 38(3), pp. 499–534. https://doi.org/10.3102/00028312038003499
9. Johnson, J.W. (2000) 'A heuristic method for estimating the relative weight of predictor variables in multiple regression', Multivariate Behavioral Research, 35(1), pp. 1–19. https://doi.org/10.1207/S15327906MBR3501_1
10. Johnson, S.M., Kraft, M.A. and Papay, J.P. (2012) 'How context matters in high-need schools: the effects of teachers' working conditions on their professional satisfaction and their students' achievement', Teachers College Record, 114(10), pp. 1–39.
11. Ladd, H.F. (2011) 'Teachers' perceptions of their working conditions: how predictive of planned and actual teacher movement?', Educational Evaluation and Policy Analysis, 33(2), pp. 235–261. https://doi.org/10.3102/0162373711398128
12. Mancuso, S.V., Roberts, L. and White, G.P. (2010) 'Teacher retention in international schools: the key role of school leadership', Journal of Research in International Education, 9(3), pp. 306–323. https://doi.org/10.1177/1475240910388928
13. Martilla, J.A. and James, J.C. (1977) 'Importance-performance analysis', Journal of Marketing, 41(1), pp. 77–79. https://doi.org/10.1177/002224297704100112
14. Meyer, J.P. and Allen, N.J. (1997) Commitment in the Workplace: Theory, Research, and Application. Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781452231556

15. Meyer, J.P., Stanley, D.J., Herscovitch, L. and Topolnytsky, L. (2002) 'Affective, continuance, and normative commitment to the organization: a meta-analysis of antecedents, correlates, and consequences', Journal of Vocational Behavior, 61(1), pp. 20–52. https://doi.org/10.1006/jvbe.2001.1842
16. Mo, J. and Morris, G. (2024) 'Investigating the employment motivation, job satisfaction, and dissatisfaction of international high school teachers in China: the impact of the COVID-19 pandemic', Frontiers in Psychology, 15, 1271604. https://doi.org/10.3389/fpsyg.2024.1271604
17. Odland, G. and Ruzicka, M. (2009) 'An investigation into teacher turnover in international schools', Journal of Research in International Education, 8(1), pp. 5–29. https://doi.org/10.1177/1475240908100679
18. Quarmson, I. (2026) 'Is our result unusual? A practical procedure for benchmarking small-sample correlations against meta-analytic evidence', International Journal for Multidisciplinary Research, 8(6), pp. 1–15. https://doi.org/10.36948/ijfmr.2026.v08i06.91207
19. Quarmson, I. and Justin Raj, M.S. (2026) 'When satisfaction loses its grip: Constraint, mobility and turnover intention among expatriate educators in China', International Journal of Science and Research Archive, 21(2), pp. 10–21. https://doi.org/10.30574/ijsra.2026.21.2.1961
20. Skaalvik, E.M. and Skaalvik, S. (2011) 'Teacher job satisfaction and motivation to leave the teaching profession: relations with school context, feeling of belonging, and emotional exhaustion', Teaching and Teacher Education, 27(6), pp. 1029–1038. https://doi.org/10.1016/j.tate.2011.04.001
21. Spector, P.E. (1997) Job Satisfaction: Application, Assessment, Causes and Consequences. Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781452231549
22. Toropova, A., Myrberg, E. and Johansson, S. (2021) 'Teacher job satisfaction: the importance of school working conditions and teacher characteristics', Educational Review, 73(1), pp. 71–97. https://doi.org/10.1080/00131911.2019.1705247
23. Wanous, J.P., Reichers, A.E. and Hudy, M.J. (1997) 'Overall job satisfaction: how good are single-item measures?', Journal of Applied Psychology, 82(2), pp. 247–252. https://doi.org/10.1037/0021-9010.82.2.247

APPENDIX A SURVEY ITEMS

All items used a six-point agreement scale (1 = strongly disagree, 2 = disagree, 3 = slightly disagree, 4 = slightly agree, 5 = agree, 6 = strongly agree). No item was reverse-scored. Scores are item means.

Table A1 Item wording and scoring

Item	Wording	Scale or facet
B1	I am fairly paid for the work I do at this school.	Compensation and benefits
B2	The benefits and expatriate package I receive are adequate.	Compensation and benefits
B3	This school offers me meaningful professional development.	Development and progression
B4	There are clear career progression opportunities at this school.	Development and progression
B5	My supervisor is competent and fair.	Leadership and supervision
B6	School leadership communicates decisions clearly.	Leadership and supervision
B7	I find my work at this school personally meaningful.	Meaningful work and autonomy
B8	I have appropriate professional autonomy in my teaching.	Meaningful work and autonomy
B9	The school supports my adjustment to living and working in Guangzhou.	Adjustment support

B10	My workload is manageable within contracted hours.	Work–life balance
B11	The school respects my personal and family time.	Work–life balance
B12	Overall, I am satisfied with my job at this school.	Global item (not in a facet)
C1	I feel a strong sense of belonging at this school.	Affective commitment (M = 4.23, SD = 1.13)
C2	I feel emotionally attached to this school.	Affective commitment (M = 4.02, SD = 1.28)
C3	This school has a great deal of personal meaning for me.	Affective commitment (M = 4.00, SD = 1.40)
D1	I often think about leaving this school.	Turnover intention (M = 3.44, SD = 1.39)
D2	I will probably look for a new job with a different employer in the next year.	Turnover intention (M = 3.48, SD = 1.49)
D3	As soon as I can find a better position, I will leave this school.	Turnover intention (M = 3.81, SD = 1.50)

Source: survey items used in the original study. Satisfaction items adapted from Spector (1997); commitment items from Meyer and Allen (1997); turnover intention items from Cammann et al. (1979). Item-level descriptive statistics are retained original-study summaries.

Table A2 Correlations among the study constructs (n = 52)

	JS	**AC**	**CC**	**NC**	**OC**
Affective commitment (AC)	.76				
Continuance commitment (CC)	.27	.19			
Normative commitment (NC)	.68	.71	.41		
Organisational commitment (OC)	.73	.82	.65	.90	
Turnover intention (TI)	−.26	−.42	−.33	−.35	−.46

Note. JS = twelve-item job satisfaction score. Source: Quarmson and Justin Raj (2026, Table 2 and accompanying note), supplemented by retained original-study summaries for the correlations between overall organisational commitment and its three components. The original summaries reported JS–CC ($p = .052$), JS–TI ($p = .066$) and AC–CC as not significant; all other correlations shown were reported as significant at $p < .05$.