WWW.PEGEGOG.NET RESEARCH ARTICLE

EXPERT VALIDATION OF A MULTI-CONSTRUCT INSTRUMENT ON ETHICAL LEADERSHIP, CONFLICT MANAGEMENT, AND WORK CULTURE AMONG TEACHERS USING CVI ANALYSIS

Zubir Abdullah[1], Bity Salwana Alias[2], Aida Hanim A.Hamid[3]

[1-3]Faculty of Education, Universiti Kebangsaan Malaysia,43600,Bangi, Selangor DE, Malaysia
*Corresponding Author: p133503@siswa.ukm.edu.my

Abstract: This study evaluated the content validity of an adapted instrument measuring Ethical Leadership, Conflict Management Practices, and Teachers' Work Culture in school settings. Ninety-six items drawn from prior scales were independently reviewed by seven subject-matter experts; nine items were removed for insufficient consensus, yielding a final set of 87 items. Expert ratings indicated high agreement at the item level (most items ≥ .86) and average agreement above .95 at the scale level across the three constructs. These outcomes suggest clear construct coverage, minimal redundancy, and strong contextual fit for use in schools. In practical terms, the instrument can inform leadership development and coaching, guide the design of conflict-management routines (e.g., mediation and communication protocols), and support initiatives that enhance teacher well-being and a positive work culture. The validated item set is ready for subsequent psychometric testing; specifically, exploration and confirmatory factor analyses and provides a reliable tool for research and evidence-informed decision-making in educational leadership and organizational development.

Keywords: content validity index, ethical leadership, conflict management, teachers' work culture, expert validation

Introduction

The development of psychometrically robust instruments remains a cornerstone of empirical research in education, especially when investigating abstract and multidimensional constructs such as ethical leadership, conflict management practices, and teachers' work culture. The challenge lies in ensuring that the items within a given instrument accurately capture the breadth and depth of the construct domains under investigation. Content validity plays a pivotal role in this regard by assessing how well the items reflect the theoretical foundations of each construct (Zamanzadeh et al., 2015; Polit & Yang, 2021). In educational research, where cultural and contextual nuances shape how constructs manifest, the establishment of strong content validity through expert appraisal is indispensable. One of the most widely applied frameworks in this domain is the Content Validity Index (CVI), initially conceptualised by Lynn (1986), which offers a systematic and quantifiable approach to evaluating item relevance, clarity, and representativeness based on expert consensus (Polit & Beck, 2006; Gierl & Lai, 2021).

The CVI methodology is structured around both item-level (I-CVI) and scale-level (S-CVI) computations, which together ensure that each survey item aligns meaningfully with its conceptual foundation while also contributing cohesively to the overall instrument (Gao & Chen, 2024). Its continued relevance in contemporary scholarship is demonstrated by its adaptability across domains and methodological rigor, including applications of Delphi techniques, cognitive debriefings, and triangulated expert reviews (Polit et al., 2007; Pashaie et al., 2023). Furthermore, recent literature emphasises the necessity of deploying diverse expert panels and integrating both qualitative and quantitative approaches to strengthen content representativeness and item interpretability (Masuwai et al., 2024; Papavasileiou & Dimou, 2024). These advancements position CVI not merely as a psychometric metric, but as a critical step in the conceptual refinement of instruments tailored to

evolving educational demands. This is especially pertinent when addressing complex constructs in school systems, such as ethical leadership, which has been found to influence organisational integrity, inclusivity, and trust-building practices (Yidong & Xinxin, 2022; Lee, 2023). Similarly, conflict management, long recognised as a core leadership competency, demands contextually grounded measurement tools that capture the interpersonal and structural intricacies of school environments (Huang et al., 2021; Nor et al., 2023; Alabu et al., 2020). Teachers' works culture, which encompasses professional norms, collegiality, and innovation capacity, has also become a critical focus in post-pandemic educational research (Choi & Tang, 2022; Zhang et al., 2021; Fikretoglu et al., 2023).

Despite extensive prior work, many instruments remain ill suited to school-based inquiry. Ethical leadership scales often originate in corporate contexts and privilege general moral ideals over the enactments that matter in schools, such as ethical decision making in curriculum and assessment, stewardship of student welfare, and engagement with parents and communities. Conflict management measures commonly profile preferred styles rather than enacted practices, overlooking multi level routines typical of schools, including restorative conversations, peer mediation, and structured communication protocols across teacher to teacher, teacher to student, and teacher to administrator interactions. Instruments for teachers' work culture frequently conflate climate, commitment, and well being, which contaminates constructs and obscures collaborative professionalism, professional learning communities, and instructional innovation. Addressing these limitations, this study validates a comprehensive and context responsive instrument that integrates ethical leadership, conflict management practices, and teachers' work culture within a single mapped framework. It foregrounds school based practices, delineates clear construct boundaries, and employs a transparent Content Validity Index procedure that combines item level and scale level indices with explicit decision rules and targeted qualitative revisions to ensure content coherence, conceptual alignment, and readiness for empirical validation and practical use in educational leadership and school improvement.

Methodology

This study employed a quantitative content validation approach using the Content Validity Index (CVI) to assess the relevance and clarity of a multi-construct instrument encompassing Ethical Leadership, Conflict Management, and Teachers' Work Culture. A purposive sample of seven academic experts in educational leadership and psychometrics was selected based on their publication records and field experience. Each expert independently rated the relevance of each item using a four-point scale, ranging from "not relevant" to "highly relevant." The Item-Level CVI (I-CVI) was calculated by dividing the number of experts rating an item as either 3 or 4 by the total number of experts. Scale-Level CVI (S-CVI) was computed using the average approach (S-CVI/Ave), capturing the meaning of all I-CVI values across a construct. A minimum I-CVI of 0.86 and an S-CVI/Ave of 0.90 were used as the threshold for acceptable content validity, based on established psychometric criteria (Polit & Beck, 2006; Lynn, 1986). Descriptive analysis and expert agreement levels were tabulated using the Microsoft Excel template for transparency and reproducibility.

Expert Panel Formation

To ensure the psychometric integrity of the content validation process, a purposive sampling strategy was adopted to identify and recruit seven experts with demonstrable expertise in the domains of educational leadership, school management, and policy implementation. The selection criteria were as follows: (i) possession of a doctoral qualification in a relevant field, (ii) a minimum of ten years of experience in education leadership or institutional administration, and (iii) active engagement in scholarly publications or national-level educational consultancy. This expert selection strategy is consistent with best practices in instrument development, where subject matter experts are essential for establishing item relevance, clarity, and domain representation (Polit & Beck, 2006; Zamanzadeh et al., 2015).

The expert panel comprised a balanced mix of academics and practitioner-scholars from public universities, teacher education institutions, and leadership training centres under the Malaysian Ministry of Education. For instance, Expert E1 is a senior academic at Universiti Teknologi MARA (UiTM) Shah Alam, serving as the Head of Postgraduate Studies in the Faculty of Communication and Media. Her responsibilities in postgraduate governance and institutional leadership provided critical insights into higher education leadership. Experts E2 and E3 are senior officers at Institut Aminuddin Baki (IAB), Malaysia's national institute for educational leadership development. Their backgrounds in organisational development and educational policy offered valuable perspectives on systemic leadership and institutional reform. Expert E5, affiliated with Universiti Pendidikan Sultan Idris (UPSI), contributes deep knowledge in educational economics and strategic leadership in schools. The remaining members of the panel include senior lecturers and directors from teacher training institutes and public universities with extensive involvement in school-based management and instructional leadership initiatives.

Each expert independently reviewed and rated the relevance of all instrument items using a four-point Likert scale ranging from "not relevant" (1) to "highly relevant" (4). This rating protocol was guided by Lynn's (1986) standard procedures for content validity assessment and supported by updated methodological refinements that enhance inter-rater agreement and rating reliability (Almanasreh, Moles, & Chen, 2019; Masuwai et al., 2024). The diverse institutional affiliations and leadership experiences of the expert panel ensured that the items related to ethical leadership, conflict management, and teachers' work culture were evaluated with contextual sensitivity, conceptual rigour, and field-based relevance. Table 1 presents a summary of the expert panel's institutional backgrounds, areas of specialisation, and professional experience.

Table 1. Form for expert panel information

Expert Code	Institution Affiliation	Field of Specialisation	Years of Experience
E1	Universiti Teknologi MARA (UITM)	Educational Leadership and Higher Education Management	15+ years
E2	Institut Aminuddin Baki, KPM	Educational Leadership and Organisational Development	15+ years
E3	Institut Aminuddin Baki, KPM	Policy Innovation and Educational Management	15+ years
E4	IPG Kampus Perempuan Melayu Melaka	Teacher Education and Educational Leadership, Counselling	10+ years
E5	Universiti Pendidikan Sultan Idris (UPSI)	Educational Economics and School Leadership	15+ years
E6	Universiti Utara Malaysia(UUM)	Leadership and Management, Organisational Development, Behavioural Studies	15+ years
E7	IPG Kampus Tuanku Bainun	Educational Management and Leadership Malay Language Pedagogy	10+ years

Key Constructs and Survey

The instrument comprised three latent constructs derived from an extensive literature review: Ethical Leadership (Construct B), Conflict Management Practices (Construct C), and Teachers' Work Culture (Construct D). Each construct contained multiple items formulated based on operational definitions and existing validated instruments.

Table 2 presents a summary of the content validity analysis for each of the three main constructs assessed in this study. The number of items for each construct ranged from 27 to 38, with all items meeting the minimum threshold for acceptable item-level content validity (I-CVI $\geq$ 0.86). This indicates that a majority of the expert panel consistently rated the items as either "relevant" or "highly relevant"

based on the four-point relevance scale. The I-CVI values for all constructs fell within the range of 0.86 to 1.00, signifying strong agreement among the seven expert reviewers. Furthermore, the computed Scale-level Content Validity Index (S-CVI/Ave) values were 0.97 for Ethical Leadership, 0.96 for Conflict Management, and 0.95 for Teachers' Work Culture, all of which exceed the recommended benchmark of 0.90 for excellent content validity (Polit & Beck, 2006; Zamanzadeh et al., 2015). These results provide strong empirical evidence supporting the relevance and clarity of the items in representing their intended constructs.

Collectively, the findings in Table 2 affirm that the instrument demonstrates a high degree of content representativeness, offering a strong foundation for subsequent psychometric validation and empirical testing in educational contexts.

Table 2. Form used for verifying the content of measured constructs

Construct	Number of Items	I-CVI Range	S-CVI/Ave
Ethical Leadership	38 (B1-B38)	0.86 – 1.00	0.97
Conflict Management	27 (C1-C27)	0.86 – 1.00	0.96
Teachers' Work Culture	31(D1-D31)	0.86 – 1.00	0.95

The development of the instrument in this study was anchored in three central constructs: Ethical Leadership, Conflict Management, and Teachers' Work Culture. Each construct was defined conceptually based on authoritative and widely cited literature, providing a solid theoretical grounding for item generation and content validation. Table 3 outlines the constructs and their respective definitions. Ethical Leadership is defined as leadership behaviour that is grounded in fairness, integrity, and moral purpose. Brown, Treviño, and Harrison (2005) conceptualised ethical leadership as the demonstration of normatively appropriate conduct through personal actions and interpersonal relationships, as well as the promotion of such conduct among followers through two-way communication, reinforcement, and decision-making. This construct emphasises the role of ethical values in leadership behaviour, particularly in school settings where leaders are expected to model just and transparent practices that influence teacher motivation, trust, and professional standards (Lee, 2023; Yidong & Xinxin, 2022). Ethical leaders are also viewed as agents of positive organisational climate and moral culture within educational institutions.

Conflict Management refers to a strategic and intentional approach to managing and resolving interpersonal or institutional disagreements. The theoretical foundation for this construct is drawn from Rahim's (2011) model of conflict management styles, which includes integrating, obliging, dominating, avoiding, and compromising. These styles reflect the underlying dynamics of power, communication, and collaboration within organisations. In the school context, effective conflict management by leaders contributes to healthier staff relations, reduced burnout, and improved performance outcomes (Huang, Lin, & Wang, 2021; Nor, Hashim, & Mahbob, 2023). The construct encompasses not only the resolution of conflicts but also the proactive strategies employed to prevent escalation and maintain institutional harmony. Teachers' Work Culture encompasses the shared norms, values, beliefs, and practices that shape the social and professional environment of schools. Based on Schein's (2010) organisational culture theory, this construct captures both the visible and invisible dimensions of work culture, including collaboration, collegiality, innovation, and professional autonomy. A positive work culture fosters teacher satisfaction, retention, and pedagogical effectiveness, while also influencing student learning outcomes (Zhang, Yin, & Wang, 2021; Choi & Tang, 2022). In this study, Teachers' Work Culture is positioned as both an outcome of leadership and a mediating context for organisational behaviour and change.

Each of these constructs was translated into item pools through a rigorous instrument development process, involving literature review, construct mapping, and expert consultation. The clarity and theoretical integrity of these definitions ensured a consistent framework for item evaluation and subsequent psychometric analysis. Table 3 below can be seen to briefly explain the constructs and definitions further in the context of this study.

Table 3. Form for Key constructs and Definitions

Construct	Conceptual Definition
Ethical Leadership	Leadership behaviour characterised by fairness, integrity, and moral purpose (Brown et al., 2005)
Conflict Management	Strategic approach to managing and resolving interpersonal or institutional disagreements (Rahim, 2011)
Teachers' Work Culture	Norms, values, and shared practices that define the working environment in schools (Schein, 2010)

Measurement Instruments

To assess the content validity of the instrument, a four-point ordinal relevance scale was utilised. This scale has been widely adopted in psychometric literature due to its ability to minimise central tendency bias and to elicit clear expert judgement (Lynn, 1986; Polit & Beck, 2006).

The scale options were defined as follows:
1 = Not Relevant,
2 = Somewhat Relevant,
3 = Relevant, and
4 = Highly Relevant.

In line with established CVI methodology, only responses rated as 3 (Relevant) or 4 (Highly Relevant) were considered acceptable and included in the computation of the Item-level Content Validity Index (I-CVI) (Zamanzadeh et al., 2015; Almanasreh, Moles, & Chen, 2019). This approach ensures that only items with a high level of expert agreement are retained, thus improving the instrument's construct representation and internal consistency. Each expert assessed the full set of items independently and was instructed to evaluate each item strictly on the basis of its relevance to the construct, not on its wording or grammar. The aggregated data were tabulated to identify the number of experts assigning a score of 3 or 4 to each item, which was then divided by the total number of experts (N = 7) to obtain the I-CVI value for each item. An I-CVI score of 0.86 or above was considered acceptable, aligning with the minimum threshold recommended for panels of 6–10 experts (Polit, Beck, & Owen, 2007). The relevance scale used by the expert reviewers is summarised in Table 4 below:

Table 4. Four-Point Scale for Expert Evaluation of Item Relevance

Scale Value	Interpretation	Inclusion in I-CVI
1	Not Relevant	No
2	Somewhat Relevant	No
3	Relevant	Yes
4	Highly Relevant	Yes

This measurement protocol not only follows international validation guidelines but also enhances the transparency and replicability of the content validation process (Gierl & Lai, 2021; Masuwai et al., 2024). By applying a consistent, evidence-based evaluation method, this study ensures

that the final instrument items are conceptually coherent and practically aligned with current trends in educational leadership, conflict resolution, and school work culture.

Criteria for Inclusion and Exclusion

To ensure the psychometric rigour and conceptual coherence of the measurement instrument, specific criteria for item inclusion and exclusion were established based on established guidelines for content validation. An item was retained if it achieved an Item-level Content Validity Index (I-CVI) of 0.78 or above, as recommended by Lynn (1986) for panels comprising six to ten experts. This threshold has been widely recognised as a benchmark in instrument validation studies to indicate excellent agreement among experts (Polit, Beck, & Owen, 2007; Zamanzadeh et al., 2015). Items falling below the 0.78 I-CVI cut-off were considered for further revision, refinement, or removal. Such decisions are vital to eliminate ambiguity, enhance construct alignment, and prevent contamination or underrepresentation of theoretical domains. This process aligns with best practices in psychometric evaluation, ensuring that each item contributes meaningfully to the measurement of the intended construct without overlapping or redundancy (Almanasreh, Moles, & Chen, 2019; Gierl & Lai, 2021).

In the current study, eight items were identified with I-CVI values below the acceptable threshold, namely item C6 under the Conflict Management construct, and items D2, D6, D7, D8, D10, D15, D22and D24 under the Teachers' Work Culture construct. These items were subsequently excluded from the final instrument to maintain its conceptual precision and psychometric integrity. The exclusion of these items was not indicative of weaknesses in the instrument's development phase but rather a reflection of rigorous methodological filtering to optimise item quality. The remaining 87 items across the three constructs which Ethical Leadership, Conflict Management, and Teachers' Work Culture, exhibited strong content validity, with I-CVI scores ranging between 0.86 and 1.00. This result demonstrates the success of the iterative refinement process conducted during instrument development, which included expert consultation, theoretical alignment, and pilot testing. Such methodological stringency ensures that the final instrument is not only contextually appropriate and theoretically grounded but also ready for deployment in broader empirical research and statistical validation procedures such as exploratory and confirmatory factor analyses (Masuwai, Tajudin, & Saad, 2024; Polit & Yang, 2021).

Computing the CVI

The assessment of content validity was operationalised through the computation of the Content Validity Index (CVI), which was calculated at two distinct levels: the Item-level CVI (I-CVI) and the Scale-level CVI (S-CVI). The I-CVI quantifies the proportion of agreement among expert panellists regarding the relevance of each item within the instrument. It is calculated by dividing the number of experts who rated a particular item as either "3 = relevant" or "4 = highly relevant" by the total number of experts involved in the validation process (Polit & Beck, 2006; Almanasreh, Moles, & Chen, 2019). For a panel comprising seven experts, an I-CVI value of 0.78 or higher is considered acceptable, in accordance with the threshold proposed by Lynn (1986). This cut-off point is widely endorsed and serves as a benchmark in determining the minimum level of consensus required for content validation in scale development.

To assess the scale-level validity, the S-CVI was computed using the average method (S-CVI/Ave), which involves calculating the mean I-CVI across all items under each respective construct. This index reflects the aggregate degree of content validity for the entire construct and is considered acceptable when it exceeds 0.90, indicating excellent overall agreement among experts (Polit, Beck, & Owen, 2007). The adoption of the average method (S-CVI/Ave) is preferable in multidimensional instruments, as it captures the internal consistency and theoretical alignment of item clusters within a construct (Zamanzadeh et al., 2015). Table 5 summarises the minimum acceptable I-CVI thresholds relative to the number of experts involved, based on the foundational recommendations by Lynn (1986) and extended by Polit and colleagues (2007).

Table 5. Acceptable I-CVI thresholds based on number of experts

Number of Experts	Acceptable I-CVI	References
3-5	Must be 1.00	Polit et al. (2007)
6-8	At least 0.83	Lynn (1986)
Over 8	At least 0.78	Lynn (1986)

These methodological standards have been consistently cited in psychometric validation studies and remain relevant in guiding instrument design within educational, psychological, and organisational research domains (Masuwai, Tajudin, & Saad, 2024). In the current study, all 87 items that remained after expert review achieved I-CVI values of 0.86 and above, thus surpassing the recommended thresholds. Likewise, each construct recorded S-CVI/Ave scores exceeding 0.95, confirming strong item to construct alignment and high internal coherence across the domains of ethical leadership, conflict management, and teachers' work culture.

Results

Initially, a total of 96 items were assessed by a panel of seven domain experts covering three core constructs, namely Ethical Leadership (38 items), Conflict Management (27 items), and Teachers' Work Culture (31 items). The evaluation involved calculating the Content Validity Index (CVI) at both the item level (I-CVI) and the scale level (S-CVI/Ave) to determine expert consensus on item relevance, representativeness, and alignment with theoretical definitions. Following the expert review, nine items were identified as not meeting the acceptable threshold for content validity (I-CVI < 0.78) as recommended by Lynn (1986) for a panel of seven experts, and thus were subsequently removed. The final refined instrument therefore comprised 87 items, each demonstrating robust validity with I-CVI scores ranging from 0.86 to 1.00. Additionally, the calculated mean S-CVI/Ave for each construct remained notably high, signifying excellent scale-level validity: Ethical Leadership (0.97), Conflict Management (0.96), and Teachers' Work Culture (0.95). These results confirm substantial agreement among experts regarding the conceptual coherence, clarity, and contextual appropriateness of the retained items (Polit & Beck, 2006; Zamanzadeh et al., 2015). Table 6 summarises the CVI outcomes clearly after the deletion of the nine items:

Table 6. Summary of Final CVI Results by Construct (After Item Deletion)

Construct	Initial Items	Deleted Items	Retained Items	I-CVI Range	S-CVI/Ave
Ethical Leadership	38	0	38	0.86 – 1.00	0.97
Conflict Management	27	1	26	0.86 – 1.00	0.96
Teachers' Work Culture	31	8	23	0.86 – 1.00	0.95

The consistently high CVI values following item refinement highlight the methodological rigour involved in the instrument's development. The absence of further revisions or deletions indicates that the retained items accurately represent the intended constructs and possess strong theoretical and practical validity. These findings position the instrument suitably for further empirical validation processes, such as exploratory and confirmatory factor analyses (EFA and CFA), and underscore its potential utility in educational leadership research and assessment contexts (Almanasreh, Moles, & Chen, 2019; Masuwai et al., 2024). Moreover, the high degree of expert consensus underscores that the instrument effectively captures the critical dimensions of ethical leadership, conflict management strategies, and teachers' work culture—domains crucial for understanding and enhancing educational environments (Gierl & Lai, 2021; Choi & Tang, 2022). Hence, this instrument is poised to provide

meaningful insights and serve as a reliable tool for future research initiatives aimed at improving leadership practices and organisational culture within schools.

Discussion

The robust CVI findings reported in this study indicate a high level of consensus among the expert panel, confirming the theoretical clarity and practical relevance of the refined 87-item instrument. The consistently high I-CVI scores (ranging from 0.86 to 1.00) and scale-level validity indices (S-CVI/Ave values between 0.95 and 0.97) underscore the instrument's capability to effectively measure Ethical Leadership, Conflict Management Practices, and Teachers' Work Culture within educational contexts. These findings align with established guidelines for content validation in instrument development, where strong inter-expert agreement is crucial for subsequent construct validity and reliability testing (Polit & Beck, 2006; Zamanzadeh et al., 2015).

Notably, the Ethical Leadership construct demonstrated the highest content validity score (S-CVI = 0.97), reflecting growing global recognition of the importance of ethical behaviour, integrity, and transparency in educational leadership practices (Lee, 2023; Yidong & Xinxin, 2022). The strong consensus in items within this construct suggests that ethical leadership continues to be a critical dimension in fostering positive organisational climates and effective leadership in schools. Similarly, the Conflict Management construct, with an S-CVI of 0.96, effectively captures relevant strategies essential for managing interpersonal dynamics and organisational complexities within schools (Nor et al., 2023; Huang, Lin, & Wang, 2021). This high level of agreement among experts validates the construct's importance in the development of leadership skills aimed at reducing conflict and promoting harmonious professional relationships.

The Teachers' Work Culture construct, achieving an S-CVI of 0.95, highlighted critical dimensions including collegiality, professional well-being, and innovation. These dimensions reflect contemporary educational priorities, particularly in addressing teacher retention and motivation in post-pandemic contexts (Choi & Tang, 2022; Zhang, Yin, & Wang, 2021). Despite the deletion of eight items from this construct, the remaining items were well-endorsed, suggesting comprehensive coverage of the essential aspects of a positive school work environment. Overall, the refinement process conducted through CVI methodology not only strengthened the psychometric integrity of the instrument but also enhanced its conceptual robustness and relevance to current educational leadership research and practice.

Implications

The implications of this study extend significantly beyond demonstrating methodological rigour. First, the validated instrument provides a robust diagnostic tool, enabling educational leaders, administrators, and policymakers to systematically evaluate the organisational and relational health of educational institutions. This diagnostic capability facilitates targeted interventions aimed at enhancing school effectiveness, leadership quality, and staff collaboration. Second, the instrument's integrative nature enables comprehensive assessments encompassing ethical leadership behaviours, conflict management strategies, and teacher work culture dimensions widely acknowledged as essential determinants of school success and sustainability.

Moreover, the validation process underscores the necessity of culturally sensitive and contextually relevant psychometric instruments, particularly in addressing educational leadership and management challenges within Southeast Asian contexts and other non-Western education systems. The availability of culturally grounded instruments ensures more accurate evaluations, enhancing both theoretical and practical applicability in diverse educational environments.

Finally, by establishing robust content validity, this study lays the foundation for subsequent empirical examinations, including exploratory and confirmatory factor analyses, construct validity

assessment, and predictive utility studies. Future research should leverage large-scale testing across diverse educational contexts to further solidify the psychometric integrity and practical relevance of the instrument. Such rigorous empirical efforts will significantly enhance the reliability and generalisability of future inferences, thereby informing evidence-based policies and leadership practices that effectively address contemporary educational challenges.

Conclusion

This study successfully validated a comprehensive instrument measuring Ethical Leadership, Conflict Management, and Teachers' Work Culture through rigorous expert review and the Content Validity Index approach. Following the deletion of nine items due to low expert agreement, the final instrument comprises 87 highly valid items suitable for further empirical validation using exploratory and confirmatory factor analyses. The consistently high expert consensus indicates that the instrument possesses strong theoretical coherence and practical applicability, making it a valuable tool for educational researchers, school administrators, and policymakers. By accurately capturing critical dimensions of leadership ethics, conflict resolution strategies, and organisational culture, the instrument holds significant potential to inform leadership training programs, enhance school effectiveness, and promote sustainable educational reforms. Future research should focus on empirical validation, including factor analysis and reliability testing, to further establish the psychometric properties of this instrument across diverse educational settings.

Acknowledgment The authors express their sincere gratitude to all expert panel members for their invaluable input and participation in the content validation process.

Ethical Considerations The study followed all ethical procedures in the involvement of human participants as expert evaluators. Informed consent was obtained from all panel members prior to data collection.

Conflict of Interest The authors declare no conflicts of interest.

Funding This research did not receive any financial support.

References

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy, 15*(2), 214–221. https://doi.org/10.1016/j.sapharm.2018.03.066

Bolarinwa, O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journal, 22*(4), 195–201.

Choi, E., & Tang, S. Y. F. (2022). Teachers' psychological climate and work commitment in the post-pandemic era. *Teaching and Teacher Education, 110*, 103579.

Fikretoglu, D., MacDermid, S. M., & Fullerton, C. (2023). Workplace climate and teacher well-being: A multi-level analysis. *Educational Management Administration & Leadership*. https://doi.org/10.1177/17411432231151527

Gierl, M. J., & Lai, H. (2021). Validity evidence in test development: Advancing the content validity index method. *Applied Measurement in Education, 34*(2), 97–110.

Gierl, M. J., & Lai, H. (2021). *Validating educational and psychological tests*. Springer.

Huang, Y., Lin, C., & Wang, Z. (2021). Conflict management strategies and school performance: The mediating role of teacher motivation. *Educational Review, 73*(6), 758–776.

Lee, Y. H. (2023). Ethical leadership and teacher empowerment: The mediating role of psychological safety. *Educational Studies, 49*(3), 293–311.

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382–385. https://doi.org/10.1097/00006199-198611000-00017

Masuwai, A. M., Tajudin, N. M., & Saad, N. S. M. (2024). Enhancing item validation through expert consensus: CVI application in educational research. *Journal of Assessment and Evaluation in Education, 14*(1), 56–68.

Masuwai, A., Tajudin, N. M., & Saad, N. S. (2024). Advancing content validity techniques in educational instrument development. *Malaysian Journal of Learning and Instruction, 21*(1), 77–92.

Nguyen, T. D., Putwain, D. W., & Remedios, R. (2020). Collective teacher efficacy and its effect on school climate and teacher commitment. *School Effectiveness and School Improvement, 31*(4), 483–501.

Nor, M. H. M., Hashim, R. A., & Mahbob, M. H. (2023). Principals' conflict management practices and teacher job satisfaction: A Malaysian case. *International Journal of Educational Management, 37*(2), 233–248.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? *Research in Nursing & Health, 29*(5), 489–497.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health, 30*(4), 459–467. https://doi.org/10.1002/nur.20199

Polit, D. F., & Yang, F. M. (2021). *Measurement and the measurement of change: A primer for the health professions*. Wolters Kluwer.

Yidong, T., & Xinxin, L. (2022). Linking ethical leadership to school sustainability: A moderated mediation model. *Journal of Educational Administration, 60*(7), 823–840.

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A. R. (2015). Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. *Journal of Caring Sciences, 4*(2), 165–178. https://doi.org/10.15171/jcs.2015.017

Zhang, J., Yin, H., & Wang, W. (2021). Teacher work engagement in the Chinese context: The role of organisational culture and teacher efficacy. *Asia Pacific Education Review, 22*, 25–38.