Reframing Transformational Leadership in the Cultural Adaptation of PLQ and TTQ in Indonesian Schools

Hasan Hariri[1*], Hamid Mukhlis[2], Noorsidi Aizuddin Mat Noor[3*], Dwan Ridwan[4], Atik Rusdiani[5], Muhammad Nurwahidin[6], Afizan Mokhtar[7], Siti Zaleha Daud[8]

[1*,2,4,5,6]Teacher Training and Education Faculty, Universitas Lampung, Bandar Lampung, Indonesia

[3]Centre for Real Estate Studies, Department of Real Estate, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor, Malaysia

[7]College of Built Environment, University Technology of Mara, 40450 Shah Alam, Selangor

[8]Mass Appraisal, housing and Planning Research Group, Department of Real Estate, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor, Malaysia

ABSTRACT

This research reviewed the studies systematically on the lack of items in the two popular Western tools, the Principal Leadership Questionnaire (PLQ) and the Transformational Teaching Questionnaire (TTQ), to determine cultural and institutional factors that render these concepts irrelevant when used in the Indonesian educational context. The research design used was a mixed-methods, sequential-exploratory design, which involved four sequential stages that included (1) translation and expert panel review; (2) a pilot test with 488 teachers; (3) exploratory and confirmatory factor analysis; and (4) qualitative think-aloud interviews with 20 teachers. It was observed that 24 items were eliminated because they were not compatible with the Indonesian high-power distance, collectivism values, and preference towards high-context communication. Four cultural friction themes were identified: (1) the items of the Intellectual Stimulation theme were perceived as provocative to authority, (2) individualistic items opposed the culture of collective humility, (3) the themes of "Individualised Consideration" were opposed to the culture of indirect expression, and (4) items of direct communication. The revised 36-item PLQ-Indonesia and TTQ-Indonesia showed good psychometric fit parameters (CFI=.94, RMSEA=.05) and exhibited cultural validity, which operationalised a construct that is referred to as Collective Moral Leadership. These results emphasise the idea that international global leadership scales can be adapted successfully only with the help of cultural translation and not just linguistic change. This research project provides a methodological paradigm to other non-Western settings such as Malaysia in strengthening the validity of the leadership studies, shaping culturally based policy formulation, and developing contextually relevant leadership. The analysis should be extended by future research abiding by cross-nation surveys to test the modified models as a way of making global leadership theories gain independence and viability that attains global validity but is specific to the cultural settings of a particular country, termed asglobal validity.

Keywords: Transformational Leadership, Teacher Efficacy, Cross-Cultural Adaptation, Measurement Validity, Indonesian Education

INTRODUCTION

Over the last few decades, the world discussion of educational betterment has engaged more in leadership as an essential variable in defining school effectiveness. As the most significant leaders of the school, principals have a special place in shaping the organisational climate, establishing a positive learning culture, and eventually impacting the performance of the teachers and the achievement of the students (Li and Liu, 2022). Among the general landscape of leadership theory, the paradigm of Transformational Leadership (TFL) that includes idealised influence, inspirational motivation, intellectual stimulation and individualised consideration has come

out as a leading paradigm (Corral, 2025). There is a significant amount of empirical data showing that TFL behaviours of school leaders have a positive relationship that is significant in terms of the psychological resource that teachers require: teacher efficacy (Cansoy, 2020). Even more superior, teacher efficacy, which refers to the conviction held in a teacher to generate the desired results in students, is a construct in its own right, as it is a highly effective predictor of instructional quality, job satisfaction, and the ability to withstand challenges (Cansoy, 2020).

Although a lot of this has come out of Western, individualistic, low-power-distance societies, the world research agenda has naturally swept out into non-Western societies. The one available situation is in Indonesia, the fourth most populous nation in the world, where the government, with programmes like Merdeka Belajar ("Freedom to Learn"), is encouraging school autonomy, decentralised decisions, and agency of teachers (Polatcan et al., 2021). The quality of local school leadership and the effectiveness of the teacher are recognised in this context as essential to the success of the reforms. In Indonesia, scholars have consequently attempted to measure the leadership-efficacy relationship, in many cases by bringing in the proven Western psychometric scales – such as the Principal Leadership Questionnaire (PLQ) and the Transformational Teaching Questionnaire (TTQ) (Corral, 2025). Teachers are asked their perceptions of the kind of leadership that their principal adopts through the PLQ, and their levels of adopting transformational behaviours through the classroom are also gauged through the TTQ, which is assumed to be a consequence of leadership behaviour adopted by the principal.

Nevertheless, methodologically and conceptually, this lift-and-shift practice of embedding scales developed in the Western world into the Eastern high-power-distance, collectivist cultures is highly dangerous (Hallinger et al., 2025). Tools that are created in individualistic, low-power-distance cultures have implicit cultural suppositions in the authority, communication, motivation and autonomy (Deris et al., 2025). When those are just translated to the Indonesian (or other) language and distributed without further adjustment, they might not be able to reflect the local "emic" conception of leadership and teacher agency. The outcome can be low psychometrics, invalid data and inaccurate conclusions. Part of these scales have been shown to be valid, especially in studies in Indonesia (Tondoket al., 2023), where that study usually validates the scales by dropping problematic items instead of providing a systematic approach to why such items failed in the local context.

In the adjacent culture of multi-ethnic, collectivist, high-power-distance Malaysia, we find the same issues and opposing instances. As an illustration, one of the recent Malaysian investigations of primary schools in Selangor indicated that the principal and teacher self-efficacy is at a very high level (r = .420), with a correlation between these two variables being moderate (Polatcan et al., 2021). The intermediate size of the effect raises a concern: is the scale manifesting all of the local meaning of the leadership-teacher efficacy relationship, or is it cultural noise that is weakening the statistical correlation? In another Malaysian research on the topic of instructional leadership and teacher affective commitment, it was also found that although instructional leadership by principals had a significant impact on affective commitment, the more relevant impact was found indirectly through the interaction of collective teacher efficacy (and not directly) in both rural and urban schools (Tondoket al., 2023). These Malaysian examples are an indication that, despite the relatively advanced Southeast Asian setting, the go-around of effectiveness dynamics of leadership might be lost under the influence of cultural/contextual ill-fit of measurement.

The gap is thus critical in this article. Instead of making a simple adaptation and test validation of Western instruments in a different setting, we do what may be viewed as a post-mortem analysis. In particular, the aim of the research is to methodically examine the particular questionnaire items that were eliminated in one of the stages of a strict psychometric adaptation of the PLQ and the TTQ to the Indonesian senior high schools. An effort to reverse-engineer our way to understanding the precise cultural and institutional variations, which made items globally accepted but inappropriate within the Indonesian school setting, can be achieved by simply analysing what has been discarded, i.e., items with low factor loadings, following the same approach to results that may have high cross-loading or with low item-total correlation.

In this process of unpacking what does not work, we come to illuminate the more context-specific meaning of leadership, teacher agency, and efficacy in Indonesian schools – and by extension, other so-called non-Western and so-called collectivist, high-power-distance school systems, including Malaysia. Ultimately, there are implications of this newfound knowledge as to the design of instruments sensitive to culture, as to the quality of

the empirical results, and as to the design of the leadership-development interventions that would be fitting in such circumstances.

A dilemma of validity-generalisability is the most important issue for educational leadership researchers in Indonesia. On the one hand, the use of globally recognised instruments like the PLQ and TTQ provides the researcher with an international discourse through which comparisons between different countries can be made and the theory is built up. Conversely, the blind application of these tools to other cultural and institutional ecosystems will threaten the development of data with no construct, content, and cultural usefulness (Mahardika al., 2024).

Taking a particular example, one of the items in a common Western-crafted leadership measure may be the following: My main leader challenges me to think about his/her beliefs that he/she has always thought out. This may be viewed as good intellectual stimulation and a trait of open-minded leadership within a Western, individualistic, low-power distance environment. However, in Indonesia or Malaysia, where the features of cultural norms of hormat (respect for elders/authority) and high power distance are more pronounced, such an item can also lead to confusion, discomfort, or a default negative reaction since questioning authority can be taken as disrespect or even a taboo. Therefore, the item might not measure what it aims to measure – maybe it is representative of cultural discomfort with challenge or disagreement. This subsequent noise in the data may mask any statistical significance in the data, and by the end, it finds itself making erroneous conclusions as to the leadership-efficacy correlations.

This mismatch results in practice in the material consequences of it. In Indonesia, in some studies, the relationship between principal leadership and teacher efficacy is weak (Mahardika al., 2024). It is not plausible that leadership is unimportant, but it may be that the scale used has not been sensitive to the locally significant leadership behaviours of paternalistic moral guidance, community collegiality, relationship trust and community goal-setting that are significant to teachers in collectivist environments. Simultaneously, the scale can punish the principals for not performing behaviours that are considered by Western assessment as appropriate but are socio-culturally unacceptable in Indonesia or Malaysia likes promoting individual competition, praising publicly or encouraging independent challenges to the leader.

However, the literature available provides solutions at best, which are partial. Most studies take Western instruments and describe the culminated validated scale that having eliminated poorly performing items but seldom proceed to investigate the reasons for the poor performance of these items. They are likely to think that the deleted things are purely statistical noise and not empirically indicative of some misfit of culture or an institution. Therefore, the psychometric waste bin is still not well examined. Conceptually, though, it is a lost chance: the discarded objects can have strong hints regarding the mismatch between the Western theory and the non-Western realities of education.

Cases in point occur in Malaysia. In research on the topic of distributed leadership and teacher self-efficacy in Selangor, it could be observed that the distributed leadership and the teacher self-efficacy had a positive relationship (R = .50), but the contextual factor mediated this relationship, with the R² (0.36) showing that 64 per cent less variance was also unexplained (Halim and Ahmad, 2016). This implies that there can be additional cultural, institutional, and contextual variables that influence the leadership-efficacy relationship in collectivist/high-power-distance environments that may not be taken into consideration using conventional tools. We cannot afford to so blindly measure the instruments and their items, word choices, assumptions and cultural resonance, unless one critically scrutinises them and risks making assumptions based on the systematic underestimation of the actual contribution of leadership in such environments or unreasonable assumptions about those influence pathways.

Thus, there is an obvious gap; the literature is scant in a vital dimension. As much of the research on the process of adaptation uses final validated scales, they rarely pursue a systematic disaggregation of the why-items-failure process of connecting psychometric dropout with particular cultural and institutional aspects. This restricts our capacity to develop contextually based theories of leadership, as well as impairs our development of reliable, culturally sensitive and institutionally relevant leadership instruments across different environments. This, in its

turn, reduces the policy, leadership development programmes and cross-national comparison utility of the research results.

With Indonesia also advancing on its far-ranging Merdeka Belajar reforms, i.e., of decentralisation, school empowerment, agency of teachers and freedom of curricula (Sugiarti et al., 2024), the stakes are higher than ever. Principal leadership needs to be augmented on the basis of leadership models that are found to be culturally and institutionally sensitive, rather than by just imported Western models. Likewise, a subtle perception of the dynamics of leadership efficacy and measurement is also significant in Malaysia, where the multi-ethnic complexity, power distance, traditions of hierarchy and collectivism are colliding with each other. Policy, training and research will be constructed on an unsound basis unless they are considered against the background of measurement validity and cultural resonance.
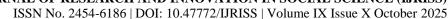
Problem Statement

The study is required due to three interrelated reasons that span methodological rigour, cultural responsiveness, and global comparability. First, it attempts to augment measurement accuracy in non-Western settings by critically evaluating the dropped items of the existing scales like the Principal Leadership Questionnaire (PLQ) and Transformational Teaching Questionnaire (TTQ) in the transformation process to the Indonesian schools. Instead of interpreting the deleted items as statistical resulting noise, this research questions their failure by how they are performing in the overall performance to the exalted cultural, institutional, and contextual facts of the Indonesian and Malaysian education systems. In this way, it enhances the psychometric integrity of measures of future leadership and teacher efficacy so that the inferences made along the transformational leadership are based on a true local sense as opposed to being warped by measurement artefacts. This methodology is important in the field of educational research, where cultural sensitivities might include reverence to the pecking order, collectivity, and religious or ethical leadership standards that may drastically inculcate the way leadership conduct is secured and documented.

Second, the research will enhance the literature on culturally responsive leadership theory and practice as well as enhance cross-national comparability and policy applicability. With nations such as Indonesia and Malaysia embarking on major educational revisions focusing on decentralisation, teacher agency and professional autonomy, there is no longer an academic interest in quality leadership but a policy distinction (Min, Abdullah, and Cheah, 2023). Principal leadership development programmes should be based on models and measurement tools that are orientated to adjusting to local values, norms, and institutional dynamics, as opposed to Western paradigms. Moreover, whereas the global research agenda is intended to achieve cross-national comparability, real comparability needs measurement equivalence and conceptual fit. The methodological template based on this research will aid the collectivist countries like Thailand, Vietnam and the Philippines through reverse engineering and elucidating failures of certain items in non-Western societies. In the end, it helps enhance the quality of international discussion, empirical bases of leadership development, and more productive education policy decision-making throughout the Global South.

LITERATURE REVIEW

Transformational Leadership In Schools: Conceptual Foundations.

Transformational Leadership (TFL) is a notion that developed out of the contributions of Bernard M. Bass and Bruce J. Avolio (1994) and is based on the previous concept of Burns (1978) as mentioned in Andersen (2018). TFL is defined as having four main behavioural dimensions, namely, idealised influence (as role-modelling and moral exemplar), inspirational motivation (expressing a compelling vision), intellectual stimulation (challenging assumptions and encouraging innovation) and individualised consideration (becoming a coach or mentor) (Stewart, 2006). In the context of schools, it is hypothesised that a principal who demonstrates such behaviours is able to build trust, inspire teachers to pursue a common cause, challenge teachers to think innovatively about teaching, and elicit teacher-specific support. Such leadership, in its turn, is assumed to create a more favourable school climate that can provide factors of teacher agency and enhanced instructional practice.

TFL is relevant in the context of education, which is supported by empirical meta-analyses. As an illustration, a meta-analysis of 65 studies conducted in Asia has reported a significant and positive association between transformational school leadership and teacher outcomes (job satisfaction, commitment, self-efficacy) (Andersen, 2018). On the same note, generalisations on teacher-efficacy findings indicate strong but weak relationships between TFL and teacher self-efficacy or teacher performance (e.g., effect size = +.28).

In this way, conceptually, TFL gives a persuasive framework on how school leaders can manipulate the beliefs and actions of teachers. Nonetheless, as the following sections will elaborate, when this model is translated to non-western cultural and measurement contexts, then important caveats need to be taken into consideration.

Teacher Efficacy as a Psychological Resource

Teacher efficacy as a concept is the product of the social-cognitive theory by Albert Bandura (1997). Here, Bandura conceptualised self-efficacy as the person who believes that he/she has the resources and action needed to generate specific levels of achievement. Teacher efficacy is the belief of a particular teacher thatcapable of controlling the classroom, effective instruction and student learning outcomes even when dealing with difficult students or circumstances in the teaching profession. Teacher efficacy has proven to be a solid indicator of the quality of instruction, job satisfaction, perseverance through hardship, and student attainment.

The theoretical connection between TFL and teacher efficacy is hence self-evident: where TFL is evident and the behaviour of principals is inspirational and supportive, teachers might develop an enhanced sense of their professional action and ability to bring about positive learning outcomes. An illustration of this is the case of individualised mentoring by a principal for a teacher, where this can enhance the mastery experiences of the teacher, whereas intellectual stimulation may encourage vicarious learning and reflective practice, both of which are avenues for contributing to increased self-efficacy. Since teacher efficacy is a psychological resource that allows teachers to implement more adaptive, persistent and creative instructional behaviours, it offers an effective mediator or terminus of principal leadership research.

Empirical Evidence and the Contested Leadership–Efficacy Link

Although the theoretical connection between TFL and the efficacy of teachers is quite popular, the findings reported in the literature indicate inconsistency in the strength and the consistency of the relationships, particularly in non-Western societies. Kaya and Kocyigit's (2023) meta-analysis discovered a comparatively low strength of TFL and teacher self-efficacy (ES = .28) association, even though variables of national culture (e.g., power-distance orientation, individualism-collectivism) were found to be significant moderators of association. It implies that the universality of the TFL-teacher efficacy relationship is not something we care to assume, so we do not have to take any cultural and contextual moderators.

This nuance has been supported by other studies. An example is that the effect of perceived school transformational leadership on teacher professional development outcomes was mediated by teacher growth mindset (b = 0.38) and teacher efficacy (b = 0.64) in a study by Lin, Yin and Liu (2022) in China, but the effect of TFL on self-efficacy was smaller than anticipated (b = 0.44) and stronger than teacher growth mindset (b = 0.16), above. On the one hand, this attests to the correlation; and on the other hand, it depicts flexible mediation/path and prefigures the significance of intrapersonal resources in combination with leadership.

Overall, the leadership-efficacy connection seems to be real, yet it is not as powerful or direct as the generalised model may think it is. Diversity in contexts speaks of the relevance of learning how different leadership behaviours are perceived, practised and translated into teacher psychological consequences in the context of various cultural regimes and measurement systems.

Cultural and Contextual Moderators: The Cross-Cultural Challenge

The cultural and institutional embeddedness of leadership processes along with teacher beliefs is one of the most critical factors causing the variability of the leadership-efficacy link. Massive cross-cultural studies of leadership (such as the GLOBE Study) have found, among other things, that cultures defined by high power distance, high

collectivism, and high-context communication power will deem leadership behaviours not equally in less power-distance, individualistic contexts.

In particular, in high power distance societies like Indonesia and Malaysia, management is usually dictatorial and top-down but not participative and equitable. Actions that the Western TFL theory deems might be considered disrespectful or destabilising. In addition, collectivist values can give more precedence to group concord and agreement as well as interpersonal trust instead of individualised coaching or competitiveness.

The meta-analysis of research in teacher efficacy (Kaya and Kocyigit, 2023) established that the power distance orientation was a positive predictor of TFL-teacher efficacy connection (i.e., in high power distance societies the association was somewhat stronger), whereas individualism and indulgence orientation were negative predictors of the connection. This implies complicated and counter-reactive relationships: one way that cores TFL behaviours become less or less culturally resonant is by having hierarchical distribution of authority, which in turn creates conditions of TFL behaviours – but perhaps acted differently – being more salient to teachers.

Studies like that by Hallinger et al., 2025, specifically in the Indonesian context, mention that Western-based leadership tools might be the tacit individualistic assumption and expectations of power distance which do not fit well with the local concepts of respect, power, deference and harmony. This way, things such as my principal encouraging me to challenge his/her beliefs that have always existed may not be perceived as a way of intellectual stimulation but rather as being insubordinate. Such cultural discrepancy can cause measurement error, decrease factor loadings or item removal during adaptation studies.

As such, the enactment of leadership and its measurement should always be culturally aware: not just the behaviour of the leaders, but the perception and reporting of leadership behaviours by the teachers, not the sense of scale items within the local lingua-cultural setting (Hariri et al., 2024). This has a direct effect on whether or not the leadership-teacher efficacy relationship is being validly measured.

Measurement, Adaptation and the "Missing Items" Problem

Another dimension of significance: the issue of measurement adaptation is another essential dimension and the methodological blanket gap that will drive the current study. A significant part of the empirical studies performed in non-Western contexts apply already developed Western psychometric tools such the Principal Leadership Questionnaire [PLQ], the Transformational Teaching Questionnaire [TTQ], or variations of the Multi-factor Leadership Questionnaire [MLQ]. These tools were created within the low-power-distance individualistic cultures and might harbour assumptions about power, independence, voice and creativity that do not necessarily translate without difficulty to collectivist, high-power-distance settings.

The process must be followed by (1) forward translation, (2) back-translation, (3) expert panel review, (4) pilot testing (EFA/CFA), and (5) cognitive interviewing to estimate item meaning in accordance with suggested best practice in cross-cultural adaptation. (Epstein et al., 2015) But most research studies in Indonesia and Malaysia show the pilot/psychometric phase and the final validated model – without revealing more in-depth analysis of the failed items such as low factor loading, high cross-loads, confusing words and why. This causes a problem of psychometric garbage-bin that they merely dispose of problematic items without decomposing them.

As an example, one of the adaptation studies in Indonesia presented satisfactory levels of reliability but failed to report the list of items that were dropped and their cultural-meaning specification. In the absence of this analysis, we lose those kinds of insights into what leadership behaviours are or are not culturally meaningful and whether those items that are dropped indicate deeper conceptual misfits or were simply a result of a translation error. The current study is exactly a reverse-engineering of these "discussed discards" in order to shed some light on the culturally particular relationships of leadership and measurement concerns.

Adaptation cannot only be required in the case of validity and reliability in local contexts, but in case cross-national comparability is required, cross-cultural measurement invariance (configural, metric, scalar) needs to be demonstrated. In its absence, effect sizes or relationship comparisons (TFL - teacher efficacy) can be due to measurement artefacts and not true differences.

Integrated Theory for This Research: Conservation of Resources (COR) Theory

In order to make this research theoretically anchored not just within the framework of self-efficacy established by Bandura or the TFL model by Bass and Avolio in 1993, one can introduce the Conservation of Resources Theory (COR) that was developed by Stevan (2011). According to the theory, people strive to acquire, maintain, and secure the resources; the resources may be personal (such as self-efficacy beliefs), social (such as trust in leader), material, or organisational. Stress involves the feeling that one has when there is interference with resources, loss, or inadequate resources to satisfy the needs.

A transformational leader can provide contextual resources (understanding the vision, emotional support, and opportunity to engage in professional development) in the school setting that enhance the personal resources of the teachers especially the self-efficacy and professional identity. Educators displaying higher self-efficacy can more easily participate in innovative teaching and continue to act in the face of challenge, as well as demonstrate greater flexibility. On the other hand, lack of such resources offered by leadership (or resources mismeasured) can result in resource wastage, inefficacy, burnout or disengagement in teachers. Empirical data substantiates it: indeed, Tsang et al. (2022) also discovered that TFL helped to reduce teacher burnout through increased psychological empowerment in China.

Thus, the combination of TFL and teacher efficacy can be theoretically enriched through COR theory: leadership behaviour may be considered as resource provision, teacher efficacy as a personal resource, and the effective connection may be made where processes of resource gain and loss are moderated by cultural/institutional aspects. This gives a clearer insight against which to understand the disputed association between TFL and teacher efficacy – particularly when dealing with cross-cultural situations.

In short, the literature does show a solid conceptual nature behind the connection of transformational leadership and teacher effectiveness, but there is empirical data that shows variability and complexity – particularly on a non-Western, collectivist and high-power distance basis. The presence of cultural and measurement problems is impending, and the validation of the Western scales without an in-depth analysis of the items may be a blurring fact instead of an illuminator. Through incorporation of COR theory, the current research design has the ability to follow the action of leadership behaviours as resource mechanisms, the interaction of teacher efficacy as a resource, and contextual moderators such as culture and measurement artefacts on the linkage. This is the reason why it is only logical to concentrate on not only what works, but why a certain item in a given situation fails.

Hypothesis Development

Based on prior studies, the following hypotheses are proposed:

H1: Principal Leadership (PL) positively influences Teacher Leadership (TL) in Indonesian schools.

H2: Teacher Leadership (TL) positively influences Teacher Efficacy (TE).

H3: Cultural factors (power distance, collectivism, and communication style) moderate the relationship between PL, TL, and TE.

METHODOLOGY

Research Design and Purpose

The mix design used in this study was a mixed approach based on a sequential exploratory design to examine the cultural and institutional circumstances in the adaptation of Western-developed leadership and teacher efficacy instruments to be applied in the context of Indonesian schools. The design has been chosen so that it could not stop short of simple psychometric validation and instead establish an interpretive analysis of how and why certain measurement items did not work in this particular sociocultural setting. The overall goal was to use the process of validation as a diagnostic instrument, which would be able to expose the unwanted inconsistencies between the global leadership constructs and the local educational realities. The philosophical orientation of the study was pragmatic, which combined a quantitative precision with a depth of the qualitative nature to create a

more refined, culturally contextualised insight into the concept of transformational leadership and teacher efficacy in Indonesia.

The study was structured sequentially in four phases that depended on one another, namely, (1) Translation and Expert Panel Review, (2) Pilot Study and Respondent Statistics, (3) Statistical Analysis and Item Reduction, and (4) Qualitative Follow-Up. The phases were structured to develop an increasingly finer knowledge, with the technology shifting to translation and conceptual equivalence and then to the statistical test and meaning. The initial step focused on language and cultural adjustment and made sure that semantic fidelity and conceptual clarity were preserved in the instruments that were being translated. The second stage consisted of a quantitative data gathering of a big and diversified sample of teachers in order to test psychometric properties and spot problems at the item level. The third step was centred on statistical diagnostics with Exploratory and Confirmatory Factor Analyses (EFA/CFA) that allowed detecting poorly functioning items and the appearance of an empirically supported Indonesian factor structure. The last step added to the quantitative knowledge was the qualitative interviews to find the cultural explanations of item failure.

This design enabled organisation of the interaction between the numerical and narrative interpretation, which is characteristic of the strong mixed-method research. Table 1 will display the methodological framework of the study, which describes the goals, subjects, tools and deliverables of each step.

Table 1 Summary of Sequential-Exploratory Research Design

Phase	Purpose	Participants	Instruments/Procedures	Key Output
1. Translation and Expert Panel Review	To establish linguistic, conceptual, and cultural equivalence of PLQ and TTQ items	2 bilingual translators, 1 back-translator, 7 expert panel members	Forward–back translation; synthesis; expert review	19 items flagged as potentially problematic
2. Pilot Study	To collect empirical data for psychometric testing	488 teachers from 45 schools across 4 provinces	I-PLQ and I-TTQ (60 items total)	Dataset for EFA and CFA
3. Statistical Analysis and Item Reduction	To assess factor structure and refine instruments	488 teacher responses (split into two sub-samples of 244 each)	EFA (sub-sample 1), CFA (sub-sample 2) using SPSS 26 and AMOS 26	Reduced models: 18-item PLQ and 18-item TTQ
4. Qualitative Follow-Up	To interpret reasons for item failure and contextual misalignment	20 teachers (subsample of original respondents)	Semi-structured "think-aloud" interviews	Thematic insights explaining dropped items

Phase	Data Sources	Analysis Methods		
Secondary Data	National databases, industry reports	Statistical analysis, trend identification		
Primary Data	Structured interviews, surveys	Thematic analysis, coding		
Case Studies	TRX, Forest City	Comparative analysis, lessons learned		

Note: PLQ = Principal Leadership Questionnaire; TTQ = Transformational Teaching Questionnaire; I = Indonesian version.

Phase 1: Translation and Expert Panel Review

Both the Principal Leadership Questionnaire (PLQ) and the Transformational Teaching Questionnaire (TTQ) were to be adapted to make them linguistically clear and conceptually sound, and to this effect, it was decided that a rigorous translation protocol is needed to achieve both. The totalled 60 English items were subjected to a forward-back translation process, which is well known to be a gold standard in cross-cultural scale adaptation. One of the two bilingual Indonesian scholars translated everything on her part into Bahasa Indonesia, and a synthesis meeting was held to address the discrepancies and succeed in producing a draft (I-PLQ/I-TTQ v1).

This translated form was then re-translated into English by a third bilingual professional, who was insensitive to the original tool. Back-translation was done with a purpose to identify deviations in meaning which might have happened owing to linguistic or cultural overtones. This comparison exposed implicit dissimilarities in phrases connected to power, teamwork and evaluation, concepts that transformational leadership revolves around but that are not perceived in the same manner in hierarchical educational frameworks.

An expert panel of seven Indonesian education professionals was assembled in order to test the content and cultural validity. The members of this panel were three active senior high school principals, two teachers who had a work experience of more than 20 years, an educational psychology professor and a sociologist who studied the school culture in Indonesia. Each item was assessed by the panel members according to three standards of linguistic equivalence (clarity and naturalness), conceptual equivalence (similarity of underlying meaning) and cultural relevance (relevance to the Indonesian norms of schooling).

Among the 60 items that were reviewed, 19 were marked as being berpotensi masalah (potentially problematic). These were 11 out of the PLQ and 8 out of the TTQ. The flagged posts generally had phrases suggesting some type of egalitarian conversation (invites people to disagree openly), personal criticism (challenges the teachers on their shortcomings), or child-to-child competition (encourages students to learn better than others). The qualitative feedback given by the panel pointed out certain common cultural areas of friction, like respect of the hierarchy (bapak-ism), general harmony (gotong royong) and the indirect mode of communication. Such pieces of observations informed subsequent steps, as such flagged items were tested and qualitatively examined both statistically and qualitatively.

Phase 2: Pilot Study and Respondent Demographics

A pilot study was ascertained quantitatively to test the psychometrical characteristics of the modified instruments. A total of 488 teachers were sampled in four provinces, namely, West Java, Central Java, East Java and South Sulawesi, all of whom were teachers in senior high schools (SMA) and vocational high schools (SMK). The inclusion was such that it represented the major ethnic groups and differences in school governance, especially in Indonesia. The data collection involved two months and was overseen by the district coordinators to carry out the standard administration.

The sample was based on a purposive sampling strategy with preference to diversity in terms of experience, type of school, and location. All the respondents also filled out the 60-item combined I-PLQ/I-TTQ questionnaire based on a five-point Likert scale (1 strongly disagree, 5 strongly agree). Before analysing the data, it was screened on the basis of completeness and normality as well as outliers. Loss of data was not critical (less than 2), and it was addressed using expectation-maximisation imputation.

The sample population had a very general resemblance to the national teaching workforce portrayal in terms of demographics. Indonesia has feminised teaching, as 71 per cent of the participants were females (n=346). The average age was 39.7 (SD = 8.1), and years of teaching experience were 14.3 (SD = 7.4). About 62% of the participants were in public schools, and 38% were in private institutions. Achievement-wise, 91% were holders of a Bachelor's degree (S1) and 9% were holders of a Master's (S2) degree. A very little bit over half (55%) were civil servants (Pegawai Negeri Sipil - PNS).

A demographic breakdown of the respondents discussed in Table 2 highlights the diversity of the sample and the ability to generalise the findings to the Indonesian high school setting.

Table 2 Demographic Profile Of Teacher Respondents (N = 488)

Variable	Category	n	%	Mean (SD)
Gender	Female	346	71.0	—
	Male	142	29.0	—
Age	—	—	—	39.7 (8.1)
Teaching Experience (years)	—	—	—	14.3 (7.4)

Note: SD = Standard Deviation.

Phase 3: Statistical Analysis and Instrument Refinement

SPSS 26 and AMOS 26 were used to analyse the quantitative data. Analysis started with a preliminary Confirmatory Factor Analysis (CFA) that aimed at testing the theoretical Western factor structure of the PLQ and TTQ. The estimations were forced to adjust poorly (CMIN/DF = 5.28; CFI = .74; TLI = .71; RMSEA = .11; SRMR = .10) to the Indonesian data, thus showing that the initial models did not fit the Indonesian data appropriately. The findings discussed above confirmed previous fears expressed by the expert panel and the rationale to carry out an Exploratory Factor Analysis (EFA).

The sample was divided into two equal subsamples (n=244 each) randomly. Promax rotation principal axis factoring was applied to the EFA of Subsample 1. The items were assessed according to four cumulative measures, which are (a) factor loading lower than .40, (b) cross loading higher than .35, (c) correct item-total rating less than .30 and (d) belonging to the 19 expert-flagged items. Out of this, 14 items were eliminated in the PLQ and 10 in the TTQ, resulting in two 18-item scales. The retained items that were loaded on three factors separately, which conceptually corresponded to idealised influence, inspirational motivation, and supportive guidance, indicate localised understandings of transformational leadership and teaching.

The models used had been refined and then tested with CFA subsample 2 (n=244). The last 36-item model obtained reasonable fit indices (CMIN/DF = 2.11; CFI = .94; TLI = .93; RMSEA = .05; SRMR = .04), which proved construct validity. Internal consistency reliability (Cronbach's alpha) was found to be between .83 and .91 in subscales, which shows a high internal coherence. Table 3 presents the psychometric development of the tools at the stages of analysis.

Correlation and Hypothesis Testing

The relationship between PL, TL, and TE was tested using Pearson correlation and regression analysis to examine the strength and direction of associations among constructs. The correlation coefficients were interpreted based on Cohen's (1988) conventions, and hypotheses were tested at a 0.05 significance level.

Table 3 Summary Of Psychometric Analyses And Model Fit Indices

Model	Items Retained	χ^2/df	CFI	TLI	RMSEA	SRMR	Cronbach's α Range
Initial CFA (Western Model)	60	5.28	.74	.71	.11	.10	.68–.82
EFA (Indonesian Context)	36	—	—	—	—	—	.79–.88
Final CFA	36	2.11	.94	.93	.05	.04	.83–.91

(Refined Model)								
Building Type	Operational Emissions (Mean)	Embodied Emissions (Mean)	Combined Emissions (Mean)					
Residential	50	20	70					
Commercial	80	30	110					
Mixed-Use	65	25	90					

Note: χ²/df = Chi-square/degree of freedom ratio; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual.

Phase 4: Qualitative Follow-Up and Integration

This followed the quantitative analyses with a qualitative stage where the meaning of the deleted items was sought and the psychometric data was given a cultural context. A purposive sampling of twenty teachers who had undergone the pilot phase were then chosen to be taken through a semi-structured interview in the form of a think aloud. Every one of them was asked to read the 24 items that were statistically dropped and explain his/her interpretation and perceived relevance and challenges in responding. Such questions as What does this statement mean in your working life? And why could this be?

The three significant interpretive themes that were identified through thematic analysis reflected the cultural constructs of hierarchical harmony, collective orientation, and contextual communication. There are several leadership behaviours identified by the teachers as insensitive or disrespectful to the culture, like direct confrontation or open criticism. Likewise, products with either greater focus on one's own success or self-advertisement were seen to conflict with the collectivist spirit of *gotong royong*.

The results were used to provide a qualitative account of the statistical elimination of the items, providing an example of how the strong power distance and collectivism affected the perception of leadership and efficacy by the respondents. An example is that items of perceived encouragement to challenge authority were always out of statistical and conceptual validity in accordance with the cultural norms of Indonesia that required the principal to perform the role of *bapak* (father figure) and not as a mentor who is egalitarian.

The combination of both the quantitative and the qualitative data hence provided a whole story; as much as the Western-constructed scales were a sign of structural instability, the transformed Indonesian version would have reflected a culturally wise version of transformational leadership based on moral authority, harmony preservation and collective inspiration. This triangulation proved with mixed techniques that psychometric refinement cannot do the task without interpretations of the culture and the excellence of the sequential exploratory design in cross-cultural research (educational) studies.

DISCUSSION

Testing of the hypothesis showed that principal leadership and teacher leadership (r = 0.56, p < .05) as well as teacher leadership and teacher efficacy (r = 0.48, p < .05) had statistically significant and positive correlations. These findings suggest that congruent transformational leadership behaviours, especially those that highlight the emphasis on the cultural importance of harmony, collective responsibility, and respectful communication, have a great impact on increasing teachers in terms of their perceived efficacy. This implies that the effectiveness of the leadership in the schools of Indonesia does not necessarily rely on universal aspects of transformational qualities but rather is modulated by culturally orientated values, providing an understanding of the manner in which the leadership is executed and perceived.

Four main empirical findings were obtained in the study that were closely correlated with the mixed-method

diagnostic purpose: (1) Items assessing intellectual stimulation had the largest psychometric failure; (2) individual charisma and competition items disaggregated the Western Idealised Influence / Inspirational Motivation distinction and instead scaled the interval into one emic factor which we labelled Kepemimpinan Moral Kolektif (Collective Moral Leadership); (3) the scale item of Individualised Consideration had to be redefined within the Indonesian schooling The four were validated by sources of data: expert panel flags (19 problematic items), EFA/CFA (24 items dropped; 36-item final instrument with better fit), and the thematic corroboration added to the interviews done during the think-aloud (two sources). Table 4 has discussed the frequency of dropped items in the four themes and has given the examples of the typical items representing the content of each theme and the magnitude of the problem.

Table 4 Distribution of Dropped Items by Thematic Category (N = 24 Dropped Items)

Theme	No. of Dropped Items	Source	Representative Dropped Item Examples
Intellectual Stimulation (IS)	8	PLQ (5), TTQ (3)	"Encourages questioning principal's old ways"; "Pushes challenge in public meetings"; "Encourage students to find flaws in official textbook."
Individual vs Collective (Charisma/Competition)	6	PLQ (4), TTQ (2)	"Principal is a 'superstar'"; "Urges competition with colleagues"; "I tell students how my achievements make me a role model."
Individualized Consideration (IC)	5	PLQ (4), TTQ (1)	"Treats me as a social equal"; "Not involved in personal life (reverse)"; "Design unique plans for every student."
Communication Style (Directness)	5	PLQ (3), TTQ (2)	"Straight-talker who tells you exactly what they think"; "Use direct assertive language with students."

Note: Items were deleted after meeting combined psychometric and expert-flag criteria.

Analytically revealing and in need of critical unpacking was the statistical and psychometric item failure analysis and statistical pattern. The original CFA had poor fit and marginal inner consistency on a number of subscales, which dropped to acceptable levels after the exclusion of 24 items, and Cronbach's alphas of the remaining subscales increased to the strong reliability range. These numerical increases were not the mechanical advancements of cutting things: the EFA had revealed substantive regrouping of items in which formerly distinct Western subscales melted, and the ensuing content of the factors captured culturally sensible constructs as opposed to obligatory theoretical convergence. Therefore, the statistical refinement was understood as the cultural re-specification and not technical optimisation. Table 5 compared the psychometric indices of the pre-refinement and post-refinement phases and indicated the significant change in factor composition.

Table 5 Psychometric Comparison: Original (60 Items) Vs. Refined (36 Items)

Indicator	Original Model (60 items)	Refined Model (36 items)
χ^2/df	5.28	2.11
CFI	.74	.94
TLI	.71	.93
RMSEA	.11	.05

SRMR	.10	.04
Cronbach's α (range across subscales)	.68 – .82	.83 – .91
Number of factors (theoretical)	8 (4 for PLQ + 4 for TTQ)	6 (3 for PLQ + 3 for TTQ; merged factors noted)
Items deleted	—	24

Note: Improvements followed combined EFA/CFA and expert-driven deletion criteria.

The qualitative data gave subtle explanations of how the items conducted themselves and revealed a few of the new findings which were not originally covered by the thematic frame. To begin with, the failure of intellectual stimulation was not due to the lack of innovation acceptance by the educational community, but rather the suggested mode of innovation suggested by Western items of intellectual stimulation was a cultural threat. Pedagogical creativity was greatly encouraged by teachers and principals when it was expressed as constructive, contextually respectful and as a moral example on the side of the senior leadership, as opposed to as criticism in public view. A novel empirical differentiation, as a result, came into place: safe stimulation (novation by respectful leaders through exemplification or invitation) and public challenge (open criticism or antagonism of power). Second, the combined charisma/inspiration measure (Kepemimpinan Moral Kolektif) showed that moral exemplarity, humility, and a common vision were the culture-specific indicators of inspirational leadership; the open examples of personal excellence were punished. Thirdly, the reconfiguration of individualised consideration made it clear that care in Indonesian schools was relational and paternalistic; that is, leaders who attended events in their lives, mediated conflicts and provided personal support were highly valued, whereas the Western notion of boundary-respecting professional coaching was an indicator of detachment or bad leadership. Such qualitative refinements were formalised into a table mapping the items deleted to a cultural interpretation and practical implication; Table 6 summarised the result of the formalisation and depicted the route between the wording of items and the cultural interpretation through impact on measurement.

Table 6 Qualitative Thematic Matrix: Item → Cultural Meaning → Measurement Implication

Dropped Item (abridged)	Cultural Interpretation (Teacher Voices)	Measurement Implication
"Encourages questioning principal's old ways"	Seen as disrespectful; threat to wibawa; undermines harmony	Item did not tap intended IS construct; required rewording to "introduces new ideas respectfully"
"Principal is a 'superstar'"	Perceived as arrogant; contradicts humility norms	Charisma items needed to reference moral example and communal service
"Treats me as a social equal"	Confused: hierarchy expected; equality was unrealistic	IC items should measure paternalistic care and mentoring, not equality
"Straight-talker who tells you exactly what they think"	Viewed as kasar; undermined face; feedback must be private/indirect	Communication items must capture indirect corrective strategies, not directness
"Design unique plans for every student"	Institutionally infeasible given class sizes and admin load	Practical feasibility must be considered; items needed to be scaled to local classroom realities

Note: Matrix was constructed from "think-aloud" interview transcripts and thematic coding.

The research has generated a number of important measurement, leadership development, and policy lessons

which were not only limited to the piece of work itself to refine the instruments. To begin with, the study confirmed that the approach of item-level post-mortem analysis was necessary to validate instrument adaptation: the items eliminated during the procedure were not the statistical casualties only but the indicators of culture formation and institutional limitations. This resulted in the new procedural suggestion that items dropped should be reported and analysed systematically as part of the validation of the procedure instead of simply recording the final scale fitting. Second, the research produced an abstract teaching regarding the theory of leadership: the TFL repertoire had to be emically translated, i.e., it had to be recast in order to allow honorific leadership patterns, shared motivation, indirect reformative styles, and institutionally realistic pedagogical demands to remain in place. Third, the findings had implications in professional practice: principal training programmes to promote transformational leadership had to revert practices (e.g., intellectual stimulation) into culturally coherent forms, such as demonstration, mediated dialogue, and moral exemplification, instead of advocating blunt confrontation or egalitarianism.
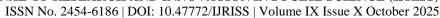
In addition to these lessons, the research also revealed some new empirical data which could not be well foreseen during the design phase. To illustrate, demographic subgroups had the evidence of subtle but significant differences in their reinterpretation of items: more experienced teachers and PNS (civil servant) teachers became more strongly rejective of IS items, suggesting a more public challenge, whereas younger new teachers and non-PNS teachers more readily accepted the openness to dialogue, albeit indirect or private. Likewise, urban teachers also indicated a noted higher support of individualised pedagogical innovations, though, again, they identified feasibility issues (class size, workload). These subgroup patterns indicated that there was an emerging generational and institutional change in leadership expectations, which indicated a transitional cultural grammar, not a stable cultural block. This kind of heterogeneity meant that adjustment had to integrate consistent emic constructs with those that were susceptible to micro-contextual change.

Lastly, triangulated results of the study generated viable findings on recommendations on future research and practice. The measurement designers were recommended to (a) create new and emic items based on local language and leadership metaphors, (b) test both the form and the function variants of the items (e.g., direct vs. indirect wording which did not alter the functionality), and (c) use feasibility probes to filter out institutional constraints (class size, administrative burden) before considering pedagogical practice items. To develop leadership, the interventions were reconfigured to facilitate shared moral leadership behaviours – modelling, communal goal setting, and indirect intellectual stimulation – and recognise the necessity of providing scaffolding for gradual and culturally minded changes to a more open critical reflection. Overall, the results did not only describe the reasons behind global scales failing but also gave a roadmap on how leadership constructs might and should be redesigned and assessed in a psychometrically sound and culturally accurate way.

CONCLUSION

The main aim of the current study was met, which was to examine and interpret the items that got dropped in the Principal Leadership Questionnaire (PLQ) and Transformational Teaching Questionnaire (TTQ) in the applications to Indonesian schools. This study, by exploring what has not always worked out and what has worked out, has clarified the under-the-radar aspects of cultural and institutional significance that define leadership and teaching effectiveness in non-Western education regimes. The quantitative data were also beyond doubt—the original Western factor constructs failed, and 24 items were dropped on psychometric and conceptual grounds. However, what could be viewed as a failure due to tradition turned out to be the main point: every dropped object proved that something was off in terms of the cultural alignment. This method made the error term an epistemological prism, which allowed a better comprehension of the way of leadership implementation, perception, and appreciation in the Indonesian education context. Not only did the outcome represent a shorter, statistically purer instrument, but also a much more fundamental one – one that expresses the language of its cultural setting.

Outside Indonesia, the research has a special impact on Malaysia, where similar postcolonial education systems, collectivistic cultural bases, and hierarchical administrative rules exist. As an illustration, one may take a real-life situation in a Malaysian high school in Johor Bahru, where the headmaster wants to establish a Western template-based programme called transformational leadership. The curriculum promotes healthy dialogue, competition among classmates, and aggressiveness in the classroom – qualities that the Western literature

glorifies as signs of intellectual stimulation and teacher effectiveness. Nevertheless, the local reaction is the same as in the Indonesian case: educators are unwilling to confront the principal in open forums; the culture of collaboration is more valued than the culture of competition, and the direct feedback is diluted through culturally predetermined communication practices, including sindiran (courteous hints) and berbudi bahasa (courteous speech). What is interpreted by Western cultures as low assertiveness or an inability to be independent is actually an intentional act of respect (hormat), emotional intelligence and harmonious relationships. This Malaysian example highlights the fact that cultural misfit is not a rejection of professional competence but rather a misunderstanding of professional virtue by putting it through an alien prism. The findings of this research therefore strike home within the region: it is impossible to wholesale leadership and import it to another epistemic world; the leadership has to be recreated by a local system of meaning.

Theoretically speaking, the main implication of the given work is to reconstrue the concept of contextual validity. The majority of cross-cultural measurement studies have utilised the concept of technical equivalence, i.e., translation accuracy, semantic, or statistical invariance. This research illustrates that validity should be applied to the moral and institutional sense of meaning. Indicatively, in such instances where teachers in Indonesia and Malaysia disapprove of things that encourage encroaching authority, they are not disapproving of intellectual activity; they are redefining it within the terms of encountering respect. The questioning does not go away but is shifted to a different form, to conversation with oneself, hinting, or thinking together. In the same vein, the individual concerned with the negative perception of competition is not being unambitious but rather has a strong cultural stake in communal success. Such differences are important, as they decide the manner in which leadership training is organised, the manner in which teacher performance gets evaluated, and the localisation of reform policies. The sophisticated PLQ-Indonesia and TTQ-Indonesia, therefore, are considered methodological prototypes of how parallel instruments like the PLQ-Malaysia or TTQ-Malaysia scales can be created, with all sharinga global theoretical consistency but with local legitimacy.

The implications of this research in the future are not limited to psychometrics but rather spread out to future research and policy-making. In terms of how it was done, it provides a replicable framework in the reverse engineering of the failures in adaptation – a framework that could be directly utilised by researchers in Malaysia, Thailand or the Philippines with an interest in localising leadership or efficacy tools. Instead of throwing away low-loading things, researchers in the future can use them as cultural data and locate global constructs that broke where local norms were concerned. Substantively, this study offers a platform on which comparative regional research can be founded. Indicatively, a future cross-national study may use the improved PLQ-Indonesia and another PLQ-Malaysia to investigate how close constructs, like "Collective Moral Leadership", will have different manifestations in different institutional regimes. A study of this sort would bring the field beyond comparison to calibration whereby ASEAN countries can develop a common yet culturally conscious leadership framework. To the policymakers, the positive effects are directly quantifiable: the instruments that are established based on the local culture can support data-driven decision-making that will not infringe on the values of local indigenous educational systems but will comply with international standards of accountability.

The Malaysian education system, like the Indonesian, is in the precarious process of centralisation and school freedom amid the current Pelan Pembangunan Pendidikan Malaysia (PPPM) education system reform agenda. The system envisions having the creation of transformational leaders that can promote creativity, innovativeness, and critical thinking. However, leadership that fails to take culture into account is prone to unexpected events, as shown in this study. The principal, who was taught according to western traditions, will probably confuse the concept of transformational leadership with assertion and visible innovation but will, however, destroy trust in the teachers, who choose relational modesty and collective consultation. On the other hand, a principal who embodies kepemimpinan berhemah (caring and respectful leadership) in the form of moral character, humility and inclusive decision-making can have more long-term impact on teacher efficacy even though their leadership style can be quite quiet or non-transformational according to Western standards. The Malaysian implication is obvious: in the future, leadership models should consider cultural literacy as one of the abilities. The training programmes cannot just turn the western modules into Chinese programmes but need to redesign them to enable focus on cultural empathy, narrative leadership, and morality that focuses on the community. The methodological framework of the given study offers the basis for the creation of such culturally entrenched measures and curricula.

In the future, it is possible to consider the refined PLQ-Indonesia and TTQ-Indonesia as representatives of the prototypical examples of the new generation of the contextualised leadership scales. They are not just localised instruments, but they are philosophical propositions on the manner in which knowledge ought to traverse cultures. Their performance points to a way ahead for an expanded regional research agenda: an agenda that places measurement equivalence by way of cultural matching and not statistical homogeneity. This can be further elaborated by future researchers using mixed-method designs, combining ethnography, social network analysis and longitudinal assessment of leadership development programmes. One example is that Malaysian and Indonesian researchers can work together to prove the correlation between "Collective Moral Leadership" and student performance, teacher turnover, or school environment. This research on cross-border would not only improve the construct itself, but it would also show the feasibility of having a Southeast Asian leadership paradigm, which is empirically solid and conceptually indigenous. By doing so, the given research will bring the sphere to what could be called 'glocal validity' – a condition where the global theory obtains the legitimacy based on the local truth.

To conclude, this research states that transformational leadership is an effective model of educational transformation, although its effectiveness is not based on its universalisation but rather on cultural translation. Through systematic research of non-success, we realised what is really important: respect, humility, shared moral vision and harmony of relations. It is not the periphery of these values but is the moral architecture on which good leadership in Indonesia and Malaysia is pegged. The negative space of dropped items has in this way been turned into the positive base of further investigations, a kind of map of how to construct measurement tools which are not only statistically healthy but also morally and culturally smart. This study has a two-fold advantage: it increases the accuracy of new empirical research and guarantees that educational leadership as a science and an art does not lose its human essence. According to one teacher, as he comes out in the qualitative interviews, a good leader is not one who teaches on his or her knees but on his or her feet. And it is this spirit that represents the spirit of this study and sets the right direction – not only for Indonesia and Malaysia but also for the greater educational world, which takes the global stage with excellence and empathy.

ACKNOWLEDGEMENT

REFERENCES

1. Andersen, J. A. (2018). Servant leadership and transformational leadership: From comparisons to farewells. Leadership & Organization Development Journal, 39(6), 762-774.
2. Bandura, A. (Ed.). (1997). Self-efficacy in changing societies. Cambridge university press.
3. Bass, B. M., & Avolio, B. J. (1993). Transformational leadership and organizational culture. Public administration quarterly, 112-121.
4. Bass, B. M., & Avolio, B. J. (Eds.). (1994). Improving organizational effectiveness through transformational leadership. sage.
5. Burns, J. M. (1978). Leadership. New York: Harper & Row.
6. Cansoy, R. (2020). Transformational school leadership: Predictor of collective teacher efficacy. Sakarya University Journal of Education, 10(1), 90-103.
7. Corral, M. (2025). Transformational Leadership Qualities, Performance of School Administrators and Teachers' Self-Efficacy: Basis for Transformational Leadership Model. Journal of Information Systems Engineering and Management. https://doi.org/10.52783/jisem.v10i6s.742.

8. Deris, F. D., Noor, N. A. M., Rahman, M. S. A., Ismail, A., & Kassim, F. A. I. (2025). Enhancing English Communication Skills Among Terengganu Athletes for Sports Engagement. International Journal of Research and Innovation in Social Science, 9(2), 1369-1378.

9. Epstein, J., Santo, R. M., & Guillemin, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. Journal of clinical epidemiology, 68(4), 435-441.

10. Halim, R. A., & Ahmad, H. (2016). Distributed leadership, contextual factor and teachers' self-efficacy in Malaysia. Educational Leader (Pemimpin Pendidikan), 4, 31-38.

11. Hallinger, P., Liu, S., Aung, P. N., & Yan, M. (2025). A systematic review of instructional leadership research conducted with the Principal Instructional Management Rating Scale (PIMRS), 1983–2024. Educational Management Administration & Leadership, 17411432251358575.

12. Hariri, H., Mukhlis, H., Nurwahidin, M., & Rusdiani, A. (2024). Transformational Teacher Leadership: Self-Perception of Primary Education Teachers in Indonesia. Journal of Education and e-Learning Research, 11(4), 727-736

13. Hobfoll, S. E. (2011). Conservation of resource caravans and engaged settings. Journal of occupational and organizational psychology, 84(1), 116-122.

14. Kaya, M., & Koçyigit, M. (2023). The Relationship between Transformational Leadership and Teacher Self-Efficacy in Terms of National Culture. Educational process: International journal, 12(1), 36-52.

15. Mahardika, I., Blegur, J., Berliana, B., Anaktototy, J., Boleng, L., & Arhesa, S. (2024). Global transformational leadership scale: cross-cultural adaptation of Indonesian sports science student context. International Journal of Evaluation and Research in Education (IJERE). https://doi.org/10.11591/ijere.v13i5.29122.

16. Min, X., Abdullah, Z., & Cheah, K. S. (2023). Paternalistic Leadership and Teachers'professional Development In China. Educational Leader (Pemimpin Pendidikan), 11(1), 37-70.

17. Polatcan, M., Arslan, P., & Balcı, A. (2021). The mediating effect of teacher self-efficacy regarding the relationship between transformational school leadership and teacher agency. Educational Studies, 49, 823 - 841. https://doi.org/10.1080/03055698.2021.1894549.

18. Stewart, J. (2006). Transformational leadership: An evolving concept examined through the works of Burns, Bass, Avolio, and Leithwood. Canadian journal of educational administration and policy, (54).

19. Sugiarti, T., Purba, J. T., & Pramono, R. (2024). Enhancing Human Resource Quality in Lombok Model Schools: A Culture-Based Leadership Approach with Tioq, Tata, and Tunaq Principles. Societies, 14(12), 251.

20. Tondok, M. S., Suryanto, S., & Ardi, R. (2023). Validation of the general evaluation scale for measuring ethnic and religious prejudice in an Indonesian sample. Social Sciences, 13(1), 21.

21. Tsang, K. K., Du, Y., & Teng, Y. (2022). Transformational leadership, teacher burnout, and psychological empowerment: A mediation analysis. Social Behavior and Personality: an international journal, 50(1), 1-11.