



Vol. 4 No. 6 (Jun) (2026)

Assessment Techniques and the Revised Bloom's Taxonomy: A Mixed-Methods Analysis of Undergraduate Examinations at Abdul Wali Khan University Mardan

Sonia Begum

MPhil Scholar, Department of Education, Abdul Wali Khan University Mardan, Pakistan

Abdul Ghaffar*

Department of Education, Abdul Wali Khan University Mardan, Pakistan

Azmat Ali Shah

Department of Education, Abdul Wali Khan University Mardan, Pakistan

ABSTRACT

Examinations shape what students learn to value as much as any lecture does, yet whether undergraduate assessment in Pakistani universities actually cultivates higher-order thinking remains empirically under-examined. This study analysed the alignment of BS-level assessment techniques at Abdul Wali Khan University Mardan (AWKUM) with the cognitive levels of the Revised Bloom's Taxonomy (RBT), using a sequential explanatory mixed-methods design. A systematic document analysis coded 600 individual question items drawn from 60 examination papers across four departments; a structured survey captured the perceptions of 150 BS students and 30 faculty members; and semi-structured interviews with faculty explored the practical barriers to RBT-aligned assessment design. Document analysis showed that lower-order cognitive levels (Remembering and Understanding) accounted for 68% of all examination items, while higher-order levels (Analyzing, Evaluating, and Creating) accounted for only 15%, with departmental alignment to higher-order thinking skills (HOTS) ranging from a low of 12% in Chemistry to a moderate 28% in Computer Science. Only 20% of faculty reported being very familiar with RBT, and 66.7% reported only rarely or never intentionally aligning assessment items to specific cognitive levels. Faculty identified large class sizes (73%), a rigid, centrally controlled examination system (67%), lack of RBT-specific training (67%), and heavy teaching workloads (60%) as the principal barriers to higher-order assessment design. These findings indicate that the gap between assessment practice and cognitive taxonomy is not primarily a matter of individual instructor effort but a structural one, rooted in institutional workload, policy rigidity, and insufficient faculty development. Implications for assessment policy, faculty training, and institutional reform are discussed.

Keywords: *Revised Bloom's Taxonomy, Higher-Order Thinking Skills, Assessment Design, Higher Education, Examination Reform, Pakistan*

Introduction

What a university chooses to test is, in practice, what it teaches students to value. If examinations reward recall over reasoning, students will rationally direct their study effort toward memorisation rather than analysis, evaluation, or creation — regardless of



Vol. 4 No. 6 (Jun) (2026)

how ambitious the stated learning outcomes of a course might be. This gap between the cognitive aspirations of a curriculum and the cognitive demands of its assessments is a persistent concern in higher education globally, and one with particular urgency in resource-constrained public universities across Pakistan, where large enrolments and limited faculty capacity make sustained, higher-order assessment design especially difficult to sustain.

This article examines that gap directly, analysing the extent to which undergraduate (BS-level) assessment practices at Abdul Wali Khan University Mardan align with the six cognitive levels of the Revised Bloom's Taxonomy, and identifying the institutional barriers faculty face in closing it. As with the corresponding author's other recent work extracted for publication, the literature review that follows draws exclusively on the corresponding author's own published research — spanning activity-based teaching, classroom and conflict management, teacher job stress, student discipline, reflective teaching practice, principal support, and teachers' social capital — together with the wider scholarship cited within those studies, in order to situate the present assessment-focused findings within a coherent and self-contained body of work.

Literature Review

Practical, Activity-Based Engagement and Higher-Order Cognitive Demand

The present study's finding that assessment at AWKUM is dominated by multiple-choice and short-answer formats testing Remembering and Understanding, while higher-order formats such as case studies and research projects remain underused, mirrors a pattern documented directly in Khitab, Zaman, Ghaffar and Jan's (2015) experimental study of activity-based teaching, which found that structured, hands-on engagement produced significantly stronger student attitudes and deeper engagement than passive, recall-oriented instruction. That study's central argument — that cognitive engagement is built through active, applied practice rather than passive exposure — applies with equal force to assessment design: a Remembering-level question invites passive recall in the same way a lecture-only classroom does, while an evaluative or creative task demands the same active engagement that activity-based teaching was shown to cultivate. Hussain, Anwar and Majoka (2011) similarly found that activity-based, peer-engaged learning produced stronger academic outcomes than passive instruction, and Malik et al. (2010) found that active, problem-solving teaching strategies produced measurably more positive student attitudes than conventional, lecture-based approaches. Bandura's (1986) account of self-efficacy adds a further dimension: faculty members who lack confidence in designing and grading open-ended, higher-order tasks — precisely the concern raised by AWKUM faculty regarding large classes — are, by Bandura's logic, less likely to attempt such tasks even where institutional constraints are not absolute, underscoring the importance of the training interventions this study recommends.

Faculty Workload, Time Constraints, and the Feasibility of Higher-Order Assessment

The finding that 60% of AWKUM faculty cited excessive time constraints and heavy teaching loads — often exceeding 18 to 20 contact hours per week — as a barrier to higher-order assessment design is directly consistent with Mahmood, Ghaffar, Ullah and Zaman's (2022) finding that job stress among Punjab's school teachers is driven substantially by excessive workload and role overload, with job stress showing a moderate positive correlation with turnover intentions. Karasek's (1979) job demands–control model, which underlies that study, holds that strain and reduced performance



Vol. 4 No. 6 (Jun) (2026)

quality emerge specifically when the demands of a role exceed the time and latitude available to meet them — precisely the condition AWKUM faculty describe when explaining their reliance on “convenient,” machine-gradable multiple-choice and short-answer formats over labour-intensive, higher-order alternatives. Kokkinos (2007) similarly identifies time constraints as a leading predictor of occupational stress among educators, while Michie (2002) and Ganster and Schaubroeck (1991) both note that chronic work overload erodes the capacity for the kind of sustained, deliberate task design that higher-order assessment construction requires. Read together, this literature suggests that faculty's reliance on lower-order assessment formats at AWKUM is not simply a matter of preference or oversight, but a predictable consequence of workload conditions closely resembling those documented among school teachers in the same province.

Reflective Practice and Faculty Development in Assessment Design

This study's finding that only 20% of AWKUM faculty felt very familiar with the Revised Bloom's Taxonomy, and that intentional alignment of assessment items to specific cognitive levels was rare, points to a faculty-development gap that Imran, Zaman and Ghaffar's (2020) study of reflective teaching practice speaks to directly. That study found that primary school teachers who engaged in regular reflective practice — whether individually or with colleagues — were better able to identify gaps in their own practice and adjust it accordingly, describing reflection as a means of continually recalibrating their teaching in light of its effects on students. The same logic extends naturally to assessment design: without structured opportunities to reflect on whether one's own test items actually engage the cognitive levels they are intended to, faculty are likely to default to familiar, low-effort formats, exactly as this study's document analysis found. Schön's (1983) account of the reflective practitioner holds that meaningful professional improvement depends on continually interrogating the gap between intended practice and actual outcomes — the same gap this study's alignment analysis exposes between AWKUM's stated cognitive aspirations and its actual examination content. Sellars (2012) identifies such reflective capacity as among the most durable engines of professional change, and Killen (2007) links structured reflection directly to more coherent, outcomes-based curriculum and assessment design, reinforcing this study's recommendation that RBT training be embedded as an ongoing, reflective faculty practice rather than a one-time workshop.

Peer Collegiality, Social Capital, and Departmental Assessment Review

This study's recommendation that departments establish structured peer-review processes for assessment items — building collegial accountability into examination design without requiring major new resources — aligns closely with the logic of Rafiq, Ghaffar and Zaman's (2021) study of teachers' social capital, which found that professionals embedded in trusting, collaborative relationships with colleagues reported stronger performance and greater confidence navigating professional challenges than those working in relative isolation. That study's participants described colleague feedback as a primary source of professional growth, particularly for less experienced staff — a dynamic directly applicable to junior AWKUM faculty currently designing RBT-aligned assessments with little formal guidance. Putnam's (2000) foundational account of social capital frames such collegial trust and reciprocity as a resource institutions can deliberately cultivate, Pil and Leana (2009) found that teachers embedded in richer professional networks had significantly better access to the knowledge needed to



Vol. 4 No. 6 (Jun) (2026)

improve their practice, and Hargreaves and Fullan (2012) argue that this kind of professional capital, built through structured peer relationships, is as important to sustaining quality practice as formal training itself. Applied to the present findings, this literature suggests that a departmental peer-review process for assessment items — rather than leaving cognitive-level alignment to individual faculty discretion — could meaningfully narrow the alignment gap this study documents.

Institutional Leadership and Policy Commitment to Assessment Reform

The finding that 67% of faculty regarded AWKUM's centralised, standardised examination system as poorly suited to higher-order assessment design underscores that the gap this study identifies cannot be closed by individual faculty effort alone. Ali, Ghaffar, Zaman and Quraishi (2021) found that institutional leadership — specifically, whether principals were perceived as supportive and responsive versus distant and authoritative — had a substantial effect on staff morale and willingness to engage with improvement initiatives, a dynamic equally relevant to whether faculty invest discretionary effort in redesigning assessment items against institutional resistance. This is consistent with Abazaoglu and Aztekin's (2016) finding that institutional atmosphere and administrative support are among the strongest predictors of staff engagement with reform, and with Leithwood, Harris and Hopkins's (2020) argument that sustained institutional change depends on leadership that builds a shared culture of expectation across staff, rather than relying on isolated individual initiative. Applied to the present case, this literature reinforces this study's conclusion that meaningful RBT alignment requires examination-policy reform and sustained leadership commitment, not merely faculty goodwill operating within an unchanged institutional structure.

Institutional Consistency and the Politics of Assessment Reform

This study's finding that RBT alignment at AWKUM is presently incidental rather than institutionally mandated echoes Khan, Naz, Iqbal and Ghaffar's (2019) finding that student discipline outcomes improved most reliably where schools adopted a clearly articulated, consistently applied, institution-wide code of conduct rather than leaving behavioural expectations to individual teacher discretion; the same logic suggests that cognitive-level alignment in assessment is unlikely to become consistent practice unless it, too, is formalised as institutional policy rather than left to vary faculty by faculty. Achieving that kind of institution-wide policy change, however, typically requires navigating competing interests and resistance among stakeholders accustomed to existing practice — the situation Ghaffar and Khan (2012) describe in their account of conflict-management styles in schools, where integrative, collaborative approaches to disagreement produce more durable institutional outcomes than either avoidance or unilateral imposition. Saddler (1998) similarly found that collaborative, consultative approaches to organisational change produce more sustained buy-in than top-down mandates alone. Read together, this literature suggests that AWKUM's move toward RBT-aligned assessment policy will depend not only on the technical recommendations this study offers, but on a genuinely collaborative process of negotiating that change with the faculty who must implement it.

Objectives of the Study

This article addresses the three research objectives of the parent thesis: (1) to determine the extent to which current assessment practices at AWKUM promote higher-order thinking skills; (2) to identify the problems faculty encounter in applying the Revised



Vol. 4 No. 6 (Jun) (2026)

Bloom's Taxonomy to BS-level assessment; and (3) to evaluate the degree of alignment between BS-level assessment techniques and the cognitive levels of the Revised Bloom's Taxonomy.

Methodology

Research Design

A sequential explanatory mixed-methods design was used, in which quantitative data (document analysis and survey) were collected and analysed first, followed by qualitative interviews to contextualise and explain the quantitative findings. This design was chosen because assessing the distribution of cognitive levels across examination items is inherently a quantitative task, while understanding the institutional and professional barriers faculty face in achieving RBT alignment required rich, contextual qualitative data.

Population and Sample

The population comprised BS-level students and faculty at Abdul Wali Khan University Mardan. A systematic document analysis coded 600 individual question items drawn from 60 examination papers purposively sampled across four departments (Computer Science, Education, English, and Chemistry). A structured survey was completed by 150 BS students (stratified by academic year and department) and 30 faculty members, with a subset of faculty subsequently participating in semi-structured interviews using purposive and snowball sampling to reach thematic saturation.

Data Collection Instruments

A Document Analysis Protocol, based on the action-verb descriptors of Anderson and Krathwohl's (2001) taxonomy table, was used to code each examination item to its highest required cognitive level; inter-rater reliability on a 15% subsample reached a Cohen's Kappa of 0.75, indicating substantial agreement. A four-section student questionnaire captured demographic information and Likert-scale perceptions of assessment cognitive demand, while a twelve-item semi-structured faculty interview guide explored RBT awareness, assessment design practice, and institutional barriers.

Data Analysis

Quantitative data were analysed using descriptive statistics to establish the distribution of cognitive levels across examination items and departments. Qualitative interview data were analysed thematically to identify recurring institutional and professional barriers, which were then integrated with the quantitative findings during the interpretation phase.

Results

Distribution of Examination Items Across Cognitive Levels

Table 1 presents the distribution of the 600 coded examination items across the six RBT cognitive levels. Lower-order cognitive skills (LOTS) accounted for 85% of all items, while higher-order cognitive skills (HOTS) accounted for only 15%.

RBT Cognitive Level	Category	Number of Items	Percentage
Remembering	LOTS	210	35.0%
Understanding	LOTS	198	33.0%



Vol. 4 No. 6 (Jun) (2026)

RBT Cognitive Level	Category	Number of Items	Percentage
Applying	LOTS	102	17.0%
Analyzing	HOTS	54	9.0%
Evaluating	HOTS	24	4.0%
Creating	HOTS	12	2.0%
Total	—	600	100%

Remembering and Understanding together accounted for 68% of all items, while Creating — the taxonomy's highest cognitive level — accounted for just 2%. This concentration at the lower end of the taxonomy indicates a substantial mismatch between AWKUM's assessment practice and the balanced cognitive engagement the Revised Bloom's Taxonomy is intended to promote.

Departmental Alignment with Higher-Order Cognitive Levels

Table 2 presents the departmental breakdown of HOTS-level items, expressed as an alignment index.

Department	Total Items	HOTS Items (%)	Alignment Level
Computer Science	150	28%	Moderate
Education	150	22%	Low-Moderate
English	150	18%	Low
Chemistry	150	12%	Low
Overall Average	600	20%	Low

Computer Science showed the strongest — though still modest — alignment with higher-order cognitive levels (28%), reflecting its emphasis on problem-solving and applied tasks, while Chemistry showed the weakest alignment (12%), reflecting greater reliance on procedural recall. The overall average of 20% indicates that, across disciplines, AWKUM's assessment practice remains concentrated well below a balanced distribution across the six cognitive levels.

Faculty-Reported Barriers to RBT-Aligned Assessment

Table 3 summarises the institutional and professional barriers faculty identified as limiting their ability to design higher-order assessments.

Challenge Identified	Faculty Reporting (%)
Large class sizes	73%
Rigid examination/assessment system	67%
Lack of RBT training / professional development	67%
Excessive time constraints / heavy workload	60%
Insufficient assessment resources and tools	43%

Large class sizes were the most frequently cited barrier (73%), with faculty describing



Vol. 4 No. 6 (Jun) (2026)

the practical impossibility of designing and grading performance-based or open-ended tasks for classes of 80–100 students. A rigid, centrally controlled examination system (67%) and lack of RBT-specific training (67%) were reported at equal frequency, followed by excessive workload (60%) and insufficient assessment resources (43%). Only 20% of faculty described themselves as very familiar with the Revised Bloom's Taxonomy, and 66.7% reported rarely or never intentionally aligning their assessment items to specific cognitive levels, despite widespread awareness that current practice under-serves higher-order thinking.

Discussion

The pattern of findings reported here is best understood as a set of structural, rather than individual, constraints on higher-order assessment — each of which finds a clear parallel in the corresponding author's own prior research. The concentration of assessment items at the Remembering and Understanding levels mirrors the passive, recall-oriented instruction that Khitab et al. (2015) found produces weaker engagement than activity-based teaching; just as that study showed hands-on practice cultivates deeper engagement than passive exposure, the present findings suggest that recall-oriented testing cultivates recall-oriented study habits, and that closing this gap requires assessment tasks that actively demand analysis, evaluation, and creation rather than simply testing for their presence in a syllabus. The workload and class-size barriers faculty identified echo Mahmood et al.'s (2022) account of how excessive occupational demands erode capacity for sustained, quality-focused professional work; just as that study found job stress compounding among teachers carrying unmanageable workloads, AWKUM faculty describe a comparable dynamic in which the labour intensity of higher-order assessment design makes lower-order, machine-gradable formats the path of least resistance under similar workload pressure. At the same time, the faculty-development gap identified here — low RBT familiarity and rare intentional alignment — is consistent with Imran et al.'s (2020) finding that reflective professional engagement, not one-off training, is what sustains meaningful practice change, while the isolation implicit in faculty designing assessments individually, with limited peer input, represents exactly the missed opportunity for professional social capital that Rafiq, Ghaffar and Zaman (2021) identify as consequential for professional performance. Finally, the finding that faculty regard the centralised examination system itself as a structural barrier reinforces Ali et al.'s (2021) and Leithwood et al.'s (2020) shared conclusion that meaningful reform depends on institutional leadership and policy commitment, achieved through the kind of collaborative, integrative negotiation with faculty stakeholders that Ghaffar and Khan (2012) and Saddler (1998) both identify as more durable than top-down mandate alone.

Conclusion

This study demonstrates that undergraduate assessment at Abdul Wali Khan University Mardan remains heavily concentrated at the lower cognitive levels of the Revised Bloom's Taxonomy, with only 15% of examination items engaging higher-order thinking skills and departmental alignment ranging from low to moderate at best. The evidence indicates that this gap is not primarily a matter of individual faculty capability or motivation, but a structural consequence of large class sizes, heavy workloads, limited RBT-specific training, and a rigid, centrally controlled examination system. Read alongside the corresponding author's prior work on activity-based teaching, teacher workload and stress, reflective practice, social capital, institutional leadership, and



Vol. 4 No. 6 (Jun) (2026)

conflict management, these findings suggest that closing the alignment gap will require simultaneous investment in faculty development, peer collegial review, and institutional policy reform — achieved through genuine negotiation with faculty rather than administrative mandate alone.

Recommendations

Introduce compulsory, discipline-specific Revised Bloom's Taxonomy training for all teaching faculty, framed as an ongoing reflective practice rather than a one-time workshop (Imran, Zaman & Ghaffar, 2020; Killen, 2007).

Establish departmental peer-review processes for assessment items prior to finalisation, drawing on the demonstrated value of collegial professional networks (Rafiq, Ghaffar & Zaman, 2021; Hargreaves & Fullan, 2012).

Reform the centralised examination system to permit a wider range of assessment formats — portfolios, case studies, and project-based work — recognising that individual faculty effort cannot substitute for institutional policy change (Ali, Ghaffar, Zaman & Quraishi, 2021).

Reduce the practical burden of higher-order assessment in large classes through technology-assisted tools such as digital portfolios and structured peer-assessment frameworks, rather than relying on unsustainable increases in individual faculty workload (Khitab et al., 2015; Mahmood et al., 2022).

Develop formal, institution-wide assessment blueprints that specify minimum cognitive-level requirements per course, applying the same logic of consistent, institution-wide policy shown to support behavioural consistency in the code-of-conduct literature (Khan et al., 2019).

Pursue assessment-policy reform through collaborative negotiation with faculty stakeholders rather than top-down mandate, consistent with evidence that integrative approaches to institutional disagreement produce more durable change (Ghaffar & Khan, 2012).

References

- Abazaoglu, I., & Aztekin, S. (2016). The role of teacher morale and motivation on students' science and math achievement: Findings from Singapore, Japan, Finland and Turkey. *Universal Journal of Educational Research*, 4(11), 2606–2617.
- Ali, M. Z., Ghaffar, A., Zaman, A., & Quraishi, M. (2021). Investigating school principals' impact on teachers' morale at secondary level. *Webology*, 18(6), 682–693.
- Bandura, A. (1986). Cognitive processes mediating behavioral change. *Journal of Personality and Social Psychology*, 35, 125–139.
- Ganster, D. C., & Schaubroeck, J. (1991). Work stress and employee health. *Journal of Management*, 17(2), 235–271.
- Ghaffar, A., & Khan, U. A. (2012). Nature of conflict and conflict management strategies. *International Journal of English and Education*, 1(2), 332–340.
- Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. Teachers College Press.
- Hussain, S., Anwar, S., & Majoka, M. I. (2011). Effect of peer group activity-based learning on students' academic achievement in physics at secondary level. *International Journal of Academic Research*, 3(1), 940–944.
- Imran, S., Zaman, A., & Ghaffar, A. (2020). Primary school in-service teachers' perceptions about reflective teaching practice. *Global Regional Review*, 5(1), 23–30.
- Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, 24(2), 285–308.
- Khan, A., Naz, B. A., Iqbal, J., & Ghaffar, A. (2019). Maintenance of students' code of conduct in developing progressive school discipline. *Journal of Education*,



Vol. 4 No. 6 (Jun) (2026)

- Society and Behavioural Science, 30(1), 1–6.
- Khitab, U., Zaman, A., Ghaffar, A., & Jan, Q. (2015). Impact of low cost teaching material on students' attitude towards science at secondary level in Khyber Pakhtunkhwa Pakistan. *International Journal of Innovation in Teaching and Learning*, 1(2), 1–14.
- Killen, R. (2007). Teaching strategies for outcomes-based education. Juta.
- Kokkinos, C. M. (2007). Job stressors, personality and burnout in primary school teachers. *British Journal of Educational Psychology*, 77(1), 229–243.
- Leithwood, K., Harris, A., & Hopkins, D. (2020). Seven strong claims about successful school leadership revisited. *School Leadership & Management*, 40(1), 5–22.
- Mahmood, K., Ghaffar, A., Ullah, M. I., & Zaman, A. (2022). Job stress and turnover intentions among school teachers of Punjab, Pakistan. *Multicultural Education*, 8(9), 78–90.
- Malik, A. M., Shah, A. Z., Iqbal, Z., & Rauf, M. (2010). Effect of problem solving teaching strategy on 8th grade students' attitude towards science. *Journal of Education and Practice*, 1(3), 16–27.
- Michie, S. (2002). Causes and management of stress at work. *Occupational and Environmental Medicine*, 59(1), 67–72.
- Pil, F. K., & Leana, C. (2009). Applying organizational research to public school reform: The effects of teacher human and social capital on student performance. *Academy of Management Journal*, 52(6), 1101–1124.
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon & Schuster.
- Rafiq, M., Ghaffar, A., & Zaman, A. (2021). Impact of teachers' social capital on their professional performance: A narrative inquiry. *Liberal Arts & Social Sciences International Journal*, 5(1), 1–15.
- Saddler, P. (1998). *Conflict management and leadership*. Coopers and Lybrand.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
- Sellars, M. (2012). Teachers and change: The role of reflective practice. *Procedia – Social and Behavioral Sciences*, 55, 461–469.