JOURNAL FAST IN SOCIAL SCIENCES
ISSN 3123-4453 Vol. 2 No.05, October 2026, pp. 1-18

Developing a Model and Prototype Instrument for Evaluating Management Performance in Baitul Maal wat Tamwil

Miftakhul Khasanah[1*], Julia Noermawati Eka Satyarini [2], Mufti Alam Adha[3]
[1,2] Sharia Economics, University of Muhammadiyah Yogyakarta, Yogyakarta, Indonesia
[2] Islamic Banking, Ahmad Dahlan University
[1]*mifthaulkhasanah@gmail.com; [2] julianoermawati@gmail.com; [3] mufti.alam@pbs.uad.ac.id

* Corresponding author

ARTICLE INFO

Article history
Revised : 15 October 2026
Accepted :28 October 2026

Keywords
development model, performance evaluation, BMT management, assessment instrument prototype, sharia compliance.

ABSTRACT

Evaluating the performance of Baitul Maal wat Tamwil (BMT) requires information regarding the organization's management processes in addition to financial achievements. Clear indicators, verifiable evidence, and defined assessment criteria are essential to ensure that evaluation results are traceable and can be utilized to determine corrective actions. This study aims to develop a preliminary conceptual model for evaluating BMT management performance, serving as a foundation for the subsequent design of a digital application. The study employs a concept-development approach, incorporating a targeted literature review, the adaptation of banking management assessment frameworks, adjustments for the specific characteristics of sharia cooperatives, and the integration of organizational culture. The development process involved mapping domains, formulating indicators, and creating a prototype instrument along with an assessment mechanism. The resulting model consists of ten components: strategy and planning, organizational structure, operational systems and procedures, human resources, leadership, organizational culture, risk management, operational risk, legal risk, and risks related to management and board members. These components are broken down into 40 initial items, supported by evidence guidelines, scoring rubrics, weighted values for each component, result interpretation, and follow-up action plans. The model includes evaluation with the inspection process and the need for improvement in BMT management. As it currently stands, this is a conceptual model and prototype instrument that requires content validation, inter-rater reliability testing, and usability testing. This conceptual framework provides a substantive basis for defining the requirements and evaluation rules for the next phase: the design of a digital application

1. Introduction

Baitul Maal wat Tamwil (BMT) provides Islamic financial services and supports community economic activities, including financing and business assistance for

microenterprises[1]. Sustaining these roles requires sound resource management and responsible governance. Nugraheni and Nasiroh (2023) found that transparency, accountability, responsibility, independence, and fairness positively influenced BMT employee performance[2]. Evaluating BMT performance therefore requires attention to the management processes that support financial outcomes and service delivery. Financial statements describe an institution's financial position and business results, while assessing planning, task allocation, and supervision requires information about its management practices.

Information about management processes becomes particularly useful when boards need to identify areas for improvement. The guidelines for assessing the soundness of Islamic financial cooperatives distinguish between checking the availability of management and operating procedures and examining whether activities follow those procedures. They also require written evidence of organizational structures and clearly defined responsibilities[3]. This distinction suggests that documents should be examined alongside their implementation. An institution may have written procedures, but their consistent application still needs to be assessed. An evaluation model should therefore connect indicators with supporting evidence and the reasons for assigning scores, allowing differences between formal provisions and everyday practices to be traced.

Inayati (2016) illustrates how management assessment can reveal improvement needs that an overall rating alone does not adequately explain[4]. The study of BMT Ben Makmur identified issues concerning long term planning, the position of the Sharia Supervisory Board, and financing monitoring. It assessed five management dimensions based on the guidelines for evaluating the soundness of Islamic financial cooperatives. These findings demonstrate the usefulness of evaluation in identifying management problems and provide a basis for developing a more operational instrument.

This need remains relevant to developments in Islamic microfinance institutions. Mutamimah et al. (2022) found variations in risk management standards and practices among Baitut Tamwil Muhammadiyah institutions. Risk management was implemented partially, with particular attention to financing and operational risks because of their implications for institutional performance and sustainability[5]. Their findings support examining the connections between organizational arrangements, the implementation of controls, and subsequent corrective action.

Digital transformation also changes how organizations manage information. Sudarmanto et al. (2024) discussed opportunities for digitalization in Islamic finance alongside challenges

---

[1] Emile Satia Darma and Akhsyim Afandi, 'The Role of Islamic Corporate Governance and Risk toward Islamic Banks Performance: Evidence from Indonesia', *Journal of Accounting and Investment*, 14 November 2021, pp. 517–38, doi:10.18196/jai.v22i3.12339.

[2] Peni Nugraheni and Candra Dewi Nasiroh, 'Improving the Performance of Baitul Maal Wat Tamwil (BMT) Employees: Do the Principles of Good Corporate Governance (GCG) Matter?', *Journal of Accounting and Investment*, 24.3 (2023), pp. 959–73.

[3] Ministry of Cooperatives and Small and Medium Enterprises, 2007.

[4] Saptaningrum Inayati, 'Evaluasi Kinerja Manajemen Koperasi Syariah Berdasarkan Peraturan Menteri Negara Koperasi Usaha Kecil Dan Menengah Republik Indonesia Nomor 35.3/PER/M. KUKM/X/2007: Studi Kasus BMT Ben Makmur Desa Kenduren Kecamatan Wedung Kabupaten Demak Tahun Buku 2013', *El-Qist: Journal of Islamic Economics and Business (JIEB)*, 6.2 (2016), pp. 1279–97.

[5] Mutamimah Mutamimah, Zaenudin Zaenudin, and Widiyanto Bin Mislan Cokrohadisumarto, 'Risk Management Practices of Islamic Microfinance Institutions to Improve Their Financial Performance and Sustainability: A Study on Baitut Tamwil Muhammadiyah, Indonesia', *Qualitative Research in Financial Markets*, 14.5 (2022), pp. 679–96, doi:10.1108/QRFM-06-2021-0099.

involving data security, regulation, and user acceptance[6]. For evaluation purposes, digital readiness begins with clarity about what is being assessed, what information is required, and how decisions will be made. Unclear assessment rules will remain problematic when transferred into a digital system.

Research on BMT performance has highlighted the importance of governance and human resource management. Nugraheni and Nasiroh (2023) found that transparency, accountability, responsibility, independence, and fairness positively influenced employee performance[7]. At KSPPS BMT Bina Ihsanul Fikri, Tauziat Aqqunduz and Diwanti (2025) also found that human resource quality, compensation, and training influenced employee performance[8]. Both studies provide empirical grounds for identifying management aspects that deserve attention in BMT evaluation. However, their primary purpose was to explain factors influencing employee performance. Further development is needed to translate these aspects into management evaluation procedures at the organizational level.

Attention to strategy and decision making has also developed in research on Islamic cooperatives. Amar et al. (2025) linked strategic management implementation and decision-making styles with organizational performance and employee satisfaction[9]. Meanwhile, Nazaruddin et al. (2024), in a study of higher education institutions, highlighted the role of appraisal systems and feedback in effective performance measurement[10]. Although the organizational context differs, the study offers a conceptual consideration for designing evaluation to support monitoring and improvement, with scores serving as part of that process.

The development of BMT evaluation instruments and systems is not entirely new. Yaya and Putra (2021) developed a performance information system based on maqasid sharia to support the evaluation of individuals and work units at BMT UMY[11]. Their study provides an important foundation for evaluation that considers the characteristics of Islamic financial institutions. The present study builds on this area by developing an internal management evaluation model that connects management functions with operational indicators, implementation evidence, scoring guidelines, and follow up actions. Its proposed contribution lies in clarifying the relationships between what is assessed, the basis for assessment, and the resulting improvement decisions.

Accordingly, this study aims to develop a model and prototype instrument for evaluating BMT management performance. The development process seeks to identify relevant components, formulate indicators that can be examined, and establish procedures for interpreting results and determining subsequent actions. This preliminary study produces an

---

[6] Eko Sudarmanto and others, 'Transformasi Digital Dalam Keuangan Islam: Peluang Dan Tantangan', *Jurnal Ilmiah Ekonomi Islam*, 10.1 (2024), pp. 645–55.
[7] Nugraheni and Nasiroh, 'Improving the Performance of Baitul Maal Wat Tamwil (BMT) Employees'.
[8] Ahmed Tauziat Aqqunduz and Dyah Pikanthi Diwanti, 'Pengaruh Kualitas Sumber Daya Manusia, Kompensasi Dan Pelatihan Kerja Terhadap Kinerja Karyawan Pada Kspps Bmt Bina Ihsanul Fikri Di Yogyakarta', *NUSANTARA: Jurnal Ilmu Pengetahuan Sosial*, 12.8 (2025), pp. 3636–43.
[9] Faozan Amar, Pepi Januar Jelita, and Deni Pandu Nugraha, 'The Relationship Between Strategic Management Implementation and Decision-Making Style in Enhancing Employee Satisfaction: Evidence from Islamic Cooperatives', *Jurnal Manajemen Bisnis*, 16.2 (2025), pp. 508–25.
[10] Ietje Nazaruddin, Hafiez Sofyani, and Tiyas Puji Utami, 'Promoting Performance Measurement System Effectiveness in Higher Education Institution: Antecedents and Consequences', *BAR-Brazilian Administration Review*, 21.2 (2024), p. e230070.
[11] Rizal Yaya and Wahyu Manuhara Putra, 'Sistem Informasi Kinerja Berbasis Maqasid Syariah Untuk Peningkatan Kinerja BMT UMY', *Prosiding Seminar Nasional Program Pengabdian Masyarakat* <https://prosiding.umy.ac.id/semnasppm/index.php/psppm/article/download/438/482> [accessed 27 October 2026].

evaluation model that can be reviewed and refined before being developed into a digital assessment tool.

2. Literature Review

2.1 Management Performance and Evaluation

Management performance concerns the quality of organizational functions in directing and controlling activities to achieve institutional goals. The performance measurement literature emphasizes the connections between strategy, organizational activities, and the indicators used to monitor their implementation[12]. In this study, the evaluation covers planning, allocation of authority, work processes, human resource development, leadership, organizational culture, and risk control. The unit of evaluation is organizational management practice over a specified period, rather than individual employee performance. This scope is intended to clarify how management functions are carried out and which areas require improvement.

Evaluation requires operational criteria so that indicators can be applied consistently. Di Luozzo et al. (2023) emphasized the need for guidelines on implementing activity indicators and connecting them with performance indicators[13]. Although their study does not specifically address BMT, it provides a conceptual basis for examining work processes. In the proposed model, this consideration is translated into connections between indicators, supporting evidence, and scoring justifications. Assessment therefore focuses on practices that can be traced through evidence. Financial results may provide context or prompt further examination, while the instrument primarily addresses internal management processes.

The guidelines for assessing the soundness of Islamic financial cooperatives under Ministerial Regulation No. 35.3/Per/M.KUKM/X/2007 provide a historical foundation for understanding the assessment of general management, institutional arrangements, capital management, asset management, and liquidity management. The guidelines include document examination and verification of implementation, making them relevant to the development of connections between indicators and evidence[14]. Inayati (2016) discussed their application at BMT Ben Makmur. In this study, these sources inform the development of an instrument for internal evaluation and management improvement[15]. The soundness categories in the guidelines are not directly adopted as model categories, and the model's assessment results are not intended to determine an institution's official soundness status.

2.2 Multidimensional Performance Measurement

The Balanced Scorecard approaches performance measurement through financial, customer, internal business process, and learning and growth perspectives (Kaplan & Norton, 1992). This framework broadens attention from final outcomes to the processes and organizational capabilities that support them.

For the BMT management evaluation model, the Balanced Scorecard provides a way to understand the connections between strategic direction, internal processes, and organizational capability development. Its use does not imply that the instrument measures all four

[12] Ivo Hristov, Antonio Chirico, and Riccardo Camilli, 'The Role of Key Performance Indicators as a Performance Management Tool in Implementing Corporate Strategies: A Critical Review of the Literature', *Financial Reporting*, no. 2022/1 (2022) <https://www.francoangeli.it/riviste/Scheda_rivista.aspx?IDArticolo=70951> [accessed 27 October 2026].

[13] Sebastiano Di Luozzo and others, 'Key Activity Indicators: Critical Review and Proposal of Implementation Criteria', *International Journal of Productivity and Performance Management*, 72.10 (2023), pp. 3000–26.

[14] Ministry of Cooperatives and Small and Medium Enterprises.

[15] inayati, 'Evaluasi Kinerja Manajemen Koperasi Syariah Berdasarkan Peraturan Menteri Negara Koperasi Usaha Kecil Dan Menengah Republik Indonesia Nomor 35.3/PER/M. KUKM/X/2007'.

perspectives. Members' experiences and comprehensive financial outcomes require separate outcome indicators. This boundary keeps the model's scope consistent with what the instrument actually assesses.

2.3 Governance and Sharia Compliance

BMT management needs to consider accountability to members and compliance with Sharia Compliance. Both should be reflected in indicators that can be examined, such as clearly defined supervisory functions, actions taken in response to Sharia Supervisory Board recommendations, conflict of interest controls, and transparent accountability. Including Sharia values in an institutional vision is insufficient to demonstrate their implementation.

Wijayanti et al. (2025) examined the integration of maqasid sharia into digital accounting information systems in BMT[16]. Their study is relevant to model development because it highlights the importance of connecting information management with Sharia objectives and accountability. The present study uses these principles as normative considerations in formulating indicators. The resulting instrument is not a validated maqasid index.

2.4 Risk and Management Control

Risk management includes assigning responsibilities, identifying and assessing risks, determining responses, and monitoring their implementation. In smaller institutions, these functions do not necessarily require a separate organizational unit. Assessment should examine whether responsibilities are clear, authority is adequate, and appropriate action is taken. Mutamimah et al. (2022) identified partial risk management practices, reinforcing the need to examine consistency across these processes [17].

The risk management framework should be distinguished from the specific risks being controlled. In this model, risk management is a domain that spans organizational activities, while operational risk, legal risk, and risks associated with board members and managers are examined through separate domains. This distinction helps prevent the same attribute from being scored more than once. Financing and liquidity risks may be included in the institution's risk register, but their outcome ratios are not assessed as separate financial domains in this prototype.

2.5 Regulatory Context and Information Management

The development of legal and accountability indicators needs to consider the BMT's legal form and business activities. Minister of Cooperatives and Small and Medium Enterprises Regulation No. 8 of 2023 provides a regulatory context for the management, supervision, and reporting of cooperative savings and lending activities. Regulation No. 2 of 2024 governs cooperative accounting policies. These regulations help identify relevant categories of evidence. The model's scoring rubric, however, is a research design and does not represent assessment categories established by either regulation.

Compliance items need to be adjusted to the requirements applicable to the institution being assessed. The model does not assume that all BMTs fall under the same regulatory authority. Before applying the instrument, assessors need to establish the institution's legal status and reporting obligations. Information transparency is similarly understood as the availability of information to authorized parties, supported by a clear record of accountability.

2.6 Model Development as a Research Output

---

[16] Provita Wijayanti and others, 'The Integration of Maqasid Sharia Principles on Digital Accounting Information System in Indonesian Islamic Microfinance', *Journal of Islamic Accounting and Finance Research*, 7.1 (2025), pp. 107–26.
[17] Mutamimah, Zaenudin, and Bin Mislan Cokrohadisumarto, 'Risk Management Practices of Islamic Microfinance Institutions to Improve Their Financial Performance and Sustainability'.

An evaluation model represents the elements being assessed and the logical relationships within the assessment process. A prototype instrument translates the model into assessment items, evidence guidelines, and scoring procedures. Both can be developed before being implemented in software.

Tuunanen et al. (2024) explained how design research can be organized into coherent parts that can be evaluated[18]. Their approach supports treating model formulation as a research output with its own evaluation objectives. The appropriateness of indicator content is examined at the model and instrument level, while technical implementation and use of the digital tool are examined in the subsequent development stage.

3. Research Method

3.1 Research Approach

This study is a preliminary investigation into the development of a BMT management performance evaluation model using a conceptual approach. The model is developed through a synthesis of literature to identify relevant management components, clarify their boundaries, and translate them into evaluation indicators. Jaakkola (2020) explained that conceptual research requires a clear design, particularly concerning the selection of concepts and their roles in developing an argument[19]. In this study, literature provides a foundation for constructing the model and serves a purpose beyond describing previous research.

The research outputs comprise the evaluation model structure, component definitions, operational indicators, and a prototype instrument with assessment guidelines. Development is limited to conceptual formulation before expert validation and user testing. This separation recognizes that different development outputs require different forms of evaluation. Tuunanen et al. (2024) provided a basis for organizing design research into coherent parts that can be evaluated[20]. Their principle is used here to clarify the study boundaries, without claiming that the full design research procedure has been implemented.

3.2 Literature Sources

The materials reviewed include publications on organizational performance measurement, BMT management, Islamic cooperative governance, internal control, and risk management. Substantive literature published between 2020 and 2025 is prioritized to understand recent developments, while earlier methodological sources and foundational theories are retained where relevant. A focused narrative review is used to support model development. The review focus and source selection consider the alignment between the review objectives, the materials analyzed, and the intended contribution, as discussed by Kraus et al. (2022)[21].

Sources are selected according to three considerations: their relevance to BMT management or Islamic financial organizations, the clarity of the concepts discussed, and their usefulness in formulating evaluation indicators. Literature from other organizational contexts is used selectively when it offers suitable methodological insights. Its findings are not assumed to represent BMT conditions directly.

---

[18] Tuure Tuunanen, Robert Winter, and Jan vom Brocke, 'Dealing with Complexity in Design Science Research: A Methodology Using Design Echelons', *MIS Quarterly*, 48.2 (2024), pp. 427–58.

[19] Elina Jaakkola, 'Designing Conceptual Articles: Four Approaches', *AMS Review*, 10.1–2 (2020), pp. 18–26, doi:10.1007/s13162-020-00161-0.

[20] Tuunanen, Winter, and Brocke, 'Dealing with Complexity in Design Science Research'.

[21] Sascha Kraus and others, 'Literature Reviews as Independent Studies: Guidelines for Academic Practice', *Review of Managerial Science*, 16.8 (2022), pp. 2577–95, doi:10.1007/s11846-022-00588-8.

The units of analysis comprise management concepts, component definitions, indicators of organizational practices, forms of evidence, and assessment rules. These elements are mapped in a development matrix containing the reference source, the concept adopted, adaptations to the BMT context, the proposed indicator, and examples of supporting evidence. The matrix provides a record of the rationale for selecting and modifying indicators.

3.3 Sources Informing Model Development and Units of Analysis

Model development draws on banking management assessment frameworks, guidelines for assessing the soundness of Islamic financial cooperatives, and literature on organizational culture. Each source is used according to its contribution to the model. The study does not adopt the entire structure or assessment rules of any single source.

The banking management assessment framework is examined with reference to Bank Indonesia Circular Letter No. 6/23/DPNP of 2004, issued to implement Bank Indonesia Regulation No. 6/10/PBI/2004. This framework provides a basis for connecting general management, the implementation of risk management systems, and compliance in the evaluation of institutional management[22].

Adaptation to the institutional characteristics of Islamic cooperatives considers Ministerial Regulation No. 35.3/Per/M.KUKM/X/2007. The guidelines are used to examine the scope of management assessment and forms of verification relevant to Islamic financial cooperatives. The organizational culture component is enriched through The Celestial Management by Amin (2004)[23]. Recent literature supports the discussion of developments in BMT management, while foundational sources are retained to trace the conceptual basis of the indicators.

The literature search follows a focused narrative approach. Sources are selected for their relevance to BMT management, conceptual clarity, and potential to inform attributes that can be examined. The review is not presented as a systematic review covering all available publications.

3.4 Stages of Model Development

Indicator development follows a deductive approach, with items derived from concepts and findings in previous studies. Boateng et al. (2018) explained that item development should begin by defining the domain and its conceptual boundaries[24]. Existing literature and instruments can then inform item formulation. The present study adapts this principle to develop an organizational evaluation instrument, without claiming to have completed the entire process of scale development and validation.

Model development proceeds through four stages.

1. Identification of evaluation needs. This stage defines the objectives, scope, and information required to assess BMT management processes.
2. Mapping and definition of components. Relevant concepts are grouped, operational definitions are established, and boundaries between components are clarified.
3. Formulation of indicators and items. Components are translated into practices that can be examined, accompanied by examples of evidence and instructions for completing the instrument.

---

[22] Bank Indonesia, 'Sistem Penilaian Tingkat Kesehatan Bank Umum.', 2004 <https://www.bi.go.id/id/archive/arsip-peraturan/Documents/0151b17420f84d118de8fdf0c0642730se623dpnp.pdf?utm >.

[23] A. Ridwan Amin, 'The Celestial Management: Zikr, Pikr, Mikr', *Language*, 33.260p (2004), p. 18cm.

[24] Godfred O. Boateng and others, 'Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer', *Frontiers in Public Health*, 6 (2018), p. 149.

4. Development of the evaluation mechanism. This stage establishes the proposed rubric, scoring procedures, presentation of performance profiles, and procedures for subsequent improvement actions.

Analysis involves comparing concepts and synthesizing sources. Indicators with similar meanings are examined to reduce duplication. Items containing several distinct attributes are separated or clarified. Decisions to retain, modify, or combine indicators are documented together with their justifications.

3.5 Design Criteria and Evaluation Boundaries

The model is developed using five design criteria established in this study: relevance to BMT management functions, clarity of definitions, traceability of evidence, proportionality to institutional capacity, and usefulness for improvement. These criteria guide model construction. Their empirical achievement will be examined during subsequent testing.

Equal weights across components are used as an initial assumption for scoring simulations. This choice allows the calculation procedure to be explained transparently when an empirical basis for assigning weights is not yet available. However, equal weighting does not establish that all components have the same importance or contribution. Joint Research Centre guidance emphasizes the need to explain weighting choices because they can influence the results of composite indicators (European Commission, Joint Research Centre, 2020).

The research outputs are limited to a conceptual model and prototype instrument. Content validation, item comprehensibility, agreement between assessors, and the sensitivity of results to weighting choices remain subjects for subsequent testing. Within these boundaries, the model is not yet used to assign institutional performance categories or compare performance across BMTs.

4. Results

4.1 Synthesis of Sources Informing Model Development

This study produces a preliminary model for evaluating BMT management performance by adapting a banking management assessment framework, incorporating the characteristics of Islamic cooperatives, and enriching the organizational culture component. Development focuses on defining the evaluation scope, indicators that can be examined, supporting evidence, and the use of findings to improve institutional management.

The banking assessment framework provides a basis for treating management quality as a distinct object of evaluation. Bank Indonesia Circular Letter No. 6/23/DPNP of 2004, issued to implement Bank Indonesia Regulation No. 6/10/PBI/2004, covers general management, the implementation of risk management systems, and compliance. This framework serves as a conceptual foundation for connecting organizational management with control over business activities[25].

Adaptation to BMT institutional characteristics considers Ministerial Regulation No. 35.3/Per/M.KUKM/X/2007. The guidelines cover general management, institutional arrangements, capital management, asset management, and liquidity management. The instrument also connects assessment questions with document examination and verification of implementation, including work plans, task allocation, and compliance with management and

[25] Bank Indonesia, 'Sistem Penilaian Tingkat Kesehatan Bank Umum.' Bank Indonesia Regulation No. 6/10/PBI/2004

operating procedures. These elements provide a reference for identifying evidence appropriate to Islamic financial cooperatives[26].

The organizational culture component draws on The Celestial Management by A. Riawan Amin[27]. In this development process, the book provides a foundation for enriching organizational values and workplace behavior. Translating these ideas into indicators that can be examined forms part of the research instrument design.

Table 1. Roles of Sources in Model Development

Source	Role in development	Form of adaptation
Banking management assessment framework	Foundation for connecting organizational management, risk, and compliance	Selection of an evaluation scope relevant to BMT
Guidelines for assessing the soundness of Islamic financial cooperatives	Reference for institutional arrangements and assessment practices	Adaptation of terminology, authority, and evidence
The Celestial Management	Enrichment of organizational culture	Formulation of workplace behaviors that reflect organizational values
Research synthesis	Connection between source concepts and evaluation needs	Development of domains, indicators, evidence guidelines, rubrics, and subsequent actions

Source: Synthesis undertaken during model development.

The resulting synthesis does not adopt the complete structure or weighting scheme of any single source. The ten domains represent the organization of the model developed in this study. They are not presented as a complete set of domains specified in one regulation.

4.2 Structure of the BMT Management Performance Evaluation Model

The model comprises ten domains. The first six address organizational direction and management processes, while the remaining four examine the control framework and particular areas of risk. This grouping helps define assessment boundaries, although the domains remain interconnected in practice.

Table 2. Evaluation Domains and Areas of Focus

Domain	Evaluation focus
Strategy and planning	Clarity of direction, target setting, program implementation, and review of achievements
Organizational structure	Allocation of duties and authority, separation of functions, and coordination
Operational systems and procedures	Standardization of work processes, implementation of procedures, and controls over record keeping
Human resources	Suitability of competencies, staff placement, development, and employee appraisal
Leadership	Direction, decision making, delegation, and problem solving
Organizational culture	Application of organizational values in workplace behavior and service delivery
Risk management	Assignment of responsibilities, risk identification, assessment, response, and monitoring
Operational risk	Controls over process errors, misuse of authority, and operational disruptions
Legal risk	Management of issues involving legal status, agreements, and institutional obligations

[26] Ministry of Cooperatives and Small and Medium Enterprises.
[27] Amin, 'The Celestial Management'.

Risks associated with board members and managers	Accountability, conflicts of interest, use of authority, and responses to supervisory findings

Source: Model development in this study.

The distinction between risk management and operational risk requires clear boundaries. Risk management assesses the overall process of managing risks, while operational risk examines controls within particular activities. Similarly, the organizational structure domain evaluates the allocation of authority, whereas the domain concerning board members and managers examines accountability for its use.

Financing and liquidity risks can still be examined through the risk management domain and relevant process indicators. However, the present structure does not treat them as separate domains. The model therefore evaluates internal management processes and does not cover every aspect of BMT institutional soundness.

4.3 Operationalization of Domains into Indicators and Evidence

Each domain is translated into indicators of organizational practice. Indicator formulation connects expected practices with the object of assessment and the evidence supporting an assessor's judgment. Documents are treated as one form of evidence, with further examination of how they are used in institutional activities.

Table 3. Examples of Indicator Operationalization

Domain	Example of a proposed indicator	Examples of evidence
Strategy and planning	Work targets are reviewed periodically, and deviations are addressed	Work plans, achievement reports, and evaluation meeting minutes
Organizational structure	Duties and authority are understood and exercised by those holding the relevant positions	Organizational structure, job descriptions, and confirmation records
Systems and procedures	Sample transactions demonstrate compliance with established procedures and authorization requirements	Standard operating procedures, transaction documents, and examination records
Human resources	Competency development is determined by job requirements	Needs assessments, training records, and evaluations
Leadership	Decisions include a documented rationale, an assigned responsible person, and monitoring arrangements	Meeting minutes, supporting data, and assignment records
Organizational culture	Organizational values are reflected in accurate reporting, fulfillment of responsibilities, and service delivery	Codes of conduct, service records, complaints, and their resolution
Risk management	Major risks are recorded together with controls and responsible persons	Risk registers, mitigation plans, and monitoring results
Operational risk	Process errors are recorded and addressed, and the possibility of recurrence is examined	Incident logs, reconciliation results, and corrective actions
Legal risk	The completeness and validity of legal documents are monitored	Legal documents, registers of obligations, and review findings

Risks associated with board members and managers	Supervisory findings and potential conflicts of interest are addressed	Supervisory reports, declarations of interest, and evidence of resolution

Source: Operational design developed in this study.

Within organizational culture, indicators focus on behaviors that can be examined. A statement that employees demonstrate integrity, for example, requires an explanation of the behaviors used as the basis for assessment. These may include accurate reporting, openness in disclosing errors, and fulfillment of responsibilities. Such examples represent the operationalization adopted in this model. Attributing individual items to specific concepts in The Celestial Management requires comparison with the original text.

One document may support several assessments when the attributes being examined differ. However, the same practice should not be counted repeatedly to increase scores across several domains.

4.4 Design of the Prototype Instrument

The prototype instrument takes the form of an assessment sheet containing domain codes, indicators, assessment criteria, evidence, findings, and improvement notes. Institutional identity, the evaluation period, assessor details, and the instrument version are included so that results can be traced across periods.

The assessment distinguishes three conditions: a practice cannot yet be verified, a practice has been examined but is not implemented, or a practice is implemented with a particular degree of consistency. This distinction is necessary because the absence of evidence during an assessment does not always mean that an activity has never been carried out.

The initial design contains 40 items, with four items in each domain. This number is a provisional design decision and is not taken directly from a source instrument. Its adequacy needs to be examined in terms of content coverage and item clarity. Maintaining the same number of items across domains should not justify retaining overlapping items or excluding necessary attributes.

4.5 Proposed Assessment Mechanism and Presentation of Results

The initial scale ranges from 1 to 10 and is grouped into four levels. This design allows differences in the quality of implementation to be expressed, but it requires specific guidance for each item so that assessors apply comparable criteria. The scale and categories have not yet been validated.

Table 4. Proposed Initial Scoring Rubric

Score range	Category	Description of practice
1 - 3	Unsatisfactory	Examination indicates that the practice is absent or very limited
4 - 5	Adequate	Implementation has begun, but practice remains inconsistent or evidence is limited
6 - 7	Good	The practice is implemented and supported by sufficient evidence, although review or subsequent action requires improvement
8 - 10	Very good	The practice is consistent, documented, reviewed, and followed by appropriate action

Scores within each range require anchors that explain differences in evidence quality and implementation. For the initial demonstration, assessors may use anchor scores of 1, 4, 6, and

8 to represent the four levels until detailed guidance is available. Other scores require descriptions that distinguish them from the anchors, so that numerical precision does not exceed what the evidence can support.

Each component score is calculated as the mean of its four item scores once all items have been verified. For simulation purposes, the overall score is calculated as follows:

$$S = \sum_{j=1}^{10} w_j C_j$$

where S is the overall score, $\mathbf{C_j}$ is the score for component **j**, and $\mathbf{w_j}$ is its assigned weight. Each component receives an equal weight of 0.10, such that:

$$\sum_{j=1}^{10} w_j = 1$$

The overall score is therefore equivalent to the arithmetic mean of the ten component scores:

$$S = \frac{1}{10} \sum_{j=1}^{10} C_j$$

Equal weighting is used as a transparent initial assumption. It does not establish that all domains have the same substantive importance.

The provisional interpretation of the overall score is presented below.

Overall score	Category
$1.00 \leq S < 4.00$	Unsatisfactory
$4.00 \leq S < 6.00$	Adequate
$6.00 \leq S < 8.00$	Good
$8.00 \leq S \leq 10.00$	Very good

Categories are determined before rounding.

These thresholds form a provisional interpretation framework that requires examination through testing and expert judgment.

When evidence is unavailable, an item is recorded as not yet verified rather than assigned a zero score. A practice that has been examined and confirmed as absent may receive a low score according to the rubric. If any item remains unverified, the aggregate score is withheld and evidence completeness is reported. A not applicable status requires justification and reviewer approval. Rules for exclusions and adjustments to calculations must be established before use.

For illustration, four scores of 6, 6, 8, and 4 produce a component mean of 6.00. Results should be interpreted alongside findings and evidence, rather than through categories alone.

The assessment distinguishes the existence of formal arrangements, their implementation, consistency, and subsequent review and improvement. These elements provide the basis for the initial rubric.

Table 5. Basis for Interpreting the Initial Assessment

Level	Description of practice
Not yet adequate	Examination shows that the practice is not implemented or basic controls are unavailable
Initial implementation	The practice has begun but does not yet cover the necessary activities or is not consistently implemented
Implemented	The practice is carried out and supported by sufficient evidence, but evaluation or subsequent action remains limited
Implemented and improved	The practice is consistent, reviewed, and improved using the review findings

Source: Scoring rubric proposed in this study.

The numerical ranges and weights are provisional rules that require further validation. Equal weights can be used to demonstrate calculations, but they do not establish that all domains are equally important. The resulting categories should not be treated as official cooperative soundness classifications.

The main evaluation outputs comprise domain profiles, evidence completeness, findings, and improvement priorities. Where an overall score is used, it is interpreted together with this information. Serious findings involving misuse of authority or compliance issues still require action even when other domains receive high average scores.

4.6 Improvement Actions and the Position of the Preliminary Model

Evaluation findings are connected with an improvement plan specifying the problem, action, responsible person, target completion date, and evidence of completion. Subsequent monitoring examines whether actions have been implemented and whether problems recur. Evaluation therefore provides a basis for assessing changes in management practices.

The research outputs remain a preliminary model and prototype instrument developed before application design.

Expert validation is required to examine domain relevance, indicator coverage, clarity of wording, overlap, and the suitability of evidence. Limited testing is then needed to assess the instrument's usability and the consistency of judgments between assessors. The findings will inform model refinement before its indicators and procedures are translated into a digital assessment tool.

5.Discussion

5.1 Adapting Multiple Sources in Evaluation Model Development

The model brings together a banking management assessment framework, the institutional characteristics of Islamic cooperatives, and an enriched organizational culture component. These three sources serve different purposes. The banking framework provides a basis for understanding the relationships between organizational management, risk, and compliance, while the Islamic cooperative guidelines help adapt assessment practices to the institutional structure of BMT. The organizational culture component extends attention to the behaviors associated with carrying out responsibilities.

This approach is consistent with the scope of management assessment in Bank Indonesia Circular Letter No. 6/23/DPNP of 2004, which covers general management, the implementation of risk management systems, and compliance. Meanwhile, Ministerial Regulation No. 35.3/Per/M.KUKM/X/2007 provides examples of Islamic cooperative management assessment through questions supported by documentary evidence and verification of implementation. The use of evidence in evaluation therefore has a foundation in these sources (Bank Indonesia, 2004; Ministry of Cooperatives and Small and Medium Enterprises, 2007).

The contribution lies in establishing operational connections between the ten selected domains, indicators, evidence, scoring justifications, and subsequent actions. These connections allow evaluation findings to be traced to practices requiring improvement. However, combining sources does not in itself demonstrate that the model is superior to earlier instruments. Any such advantage needs to be established by examining content coverage, assessment clarity, and the usefulness of findings for BMT management.

5.2 Organizational Culture as an Enrichment of Management Evaluation

Treating organizational culture as a separate domain creates an opportunity to examine how organizational values are reflected in everyday work. In this study, *The Celestial Management* by Amin (2004) serves as a source of enrichment[28]. Its operationalization focuses on behaviors that can be examined, including fulfillment of responsibilities, openness in reporting, cooperation, and the resolution of members' complaints. These formulations result from instrument development and still require examination of their alignment with the concepts in the original source.

The relevance of behavior and accountability is supported by Nugraheni and Nasiroh (2023)[29]. Their study of BMT employees in Yogyakarta found that transparency, accountability, responsibility, independence, and fairness positively influenced employee performance. These findings strengthen the rationale for considering governance practices in BMT management evaluation, although they do not directly validate the organizational culture indicators developed in this study.

Organizational culture needs to be distinguished from the existence of formal rules. A code of conduct may describe the values expected by an institution, but its implementation requires additional evidence. Records of error resolution, responses to complaints, and fulfillment of work commitments can help assessors understand how those values are applied. Assessment should also consider whether the evidence is sufficiently representative, so that a single incident does not become the basis for judging the organization's culture as a whole.

5.3 Model Coverage and Connections in Risk Control

The ten domains provide an evaluation scope that connects organizational direction, work implementation, resource capabilities, and control. These connections matter because management problems may involve several functions at once. A transaction error, for example,

[28] Amin, 'The Celestial Management'.
[29] Nugraheni and Nasiroh, 'Improving the Performance of Baitul Maal Wat Tamwil (BMT) Employees'.

may relate to unclear procedures, staff competencies, limited supervision, or the exercise of authority.

Mutamimah et al. (2022) found that Baitut Tamwil Muhammadiyah institutions applied varied risk management standards and methods, with implementation remaining partial[30]. Their study also identified constraints involving human resources and control systems. These findings support examining risks alongside the management processes underlying them, rather than treating risks solely as a list of separate events.

However, the close relationships between domains create potential overlap. The systems and procedures domain should assess the standardization and implementation of work processes, while operational risk examines controls over failures in those processes. The risk management domain assesses the overall framework for identifying, responding to, and monitoring risks. These boundaries need to be maintained in item formulation to prevent the same practice from receiving repeated scores.

The model's coverage should also be interpreted in proportion to its purpose. The instrument focuses on internal management processes and does not measure all financial outcomes, social benefits, or members' experiences. Financing and liquidity risks need adequate representation within relevant indicators because neither is treated as a separate domain. The adequacy of this coverage is therefore an important issue for content validation.

5.4 Sharia Compliance, Evidence Quality, and Score Interpretation

Sharia compliance is define into assessment criteria through clearly defined supervisory functions, responses to Sharia Supervisory Board recommendations, board accountability, and conflict of interest controls. The existence of a supervisory structure needs to be examined alongside the performance of its duties and the resolution of its findings.

Wijayanti et al. (2025) examined the integration of maqasid sharia principles into digital accounting information systems in BMT[31]. Their study demonstrates the relevance of Sharia objectives to institutional information management. For the present model, this implies a need to ensure that evaluation information supports the examination of accountability and the protection of members' interests. Their findings are not used to claim that this instrument measures the achievement of maqasid sharia.

The quality of assessors' judgments depends on the sufficiency and appropriateness of evidence. A status of *not yet verified* is therefore distinguished from a practice confirmed as not implemented. This distinction prevents limited access to documents from being interpreted immediately as evidence of poor implementation. Nevertheless, withholding scores needs to be accompanied by a deadline for completing the evidence, allowing the evaluation to reach a conclusion.

---

[30] Mutamimah, Zaenudin, and Bin Mislan Cokrohadisumarto, 'Risk Management Practices of Islamic Microfinance Institutions to Improve Their Financial Performance and Sustainability'.
[31] Wijayanti and others, 'The Integration of Maqasid Sharia Principles on Digital Accounting Information System in Indonesian Islamic Microfinance'.

Assigning an equal weight of 0.10 provides an initial rule that is easy to trace. Equal weighting nevertheless remains a methodological choice. OECD/European Union/EC-JRC (2008) guidance emphasizes the need to explain weighting and aggregation decisions and examine the robustness of results when assumptions change.

A weighted mean also allows a high score in one domain to offset a low score in another. Domain profiles and serious findings should therefore remain visible alongside the overall score. At this stage, the model has not established thresholds for critical findings that would restrict the final category. The need for such rules should be considered during subsequent validation.

5.5 The Model as a Substantive Foundation for Application Development

The preliminary model identifies information requirements that can inform the development of a digital assessment tool. Domain and item codes provide a structure for recording information, evidence guidelines help determine what needs to be collected, and the rubric and formulas establish calculation rules. The improvement plan introduces further requirements for recording responsible persons, target completion dates, and problem resolution.

Tuunanen et al. (2024) proposed organizing design research into coherent parts to address the complexity of solution development[32]. Their approach supports treating the evaluation model and instrument as outputs with their own assessment objectives before further development. This study draws on that idea to explain the position of its outputs, without claiming to have implemented the entire methodology they proposed.

In subsequent application, automation may support consistent calculations and the traceability of results. However, these benefits depend on appropriate indicators, complete evidence, and a clear rubric. Readiness for application development therefore needs to be assessed in terms of the maturity of the evaluation rules as well as technical readiness.

5.6 Limitations and Directions for Testing

This study produces a conceptual model and prototype instrument that have not yet been empirically tested. The focused literature review may have left relevant approaches or important attributes unrepresented. The 40 items, scale from 1 to 10, anchor scores, equal weights, and category thresholds are initial design decisions that remain open to revision.

Content validation needs to examine relevance, clarity, coverage, potential overlap, and the suitability of evidence. Particular attention is required for the adequacy of financing and liquidity risk indicators, the boundaries between organizational culture and leadership, and the distinction between risk management and operational control.

Subsequent limited testing may involve several assessors examining the same evidence. Comparing their judgments can identify items that lead to different interpretations, while recording assessment time and difficulties can help evaluate the instrument's usability. Changes in weights and category

[32] Tuunanen, Winter, and Brocke, 'Dealing with Complexity in Design Science Research'.

thresholds should also be simulated to determine whether the model's conclusions are sensitive to the selected assessment rules.

Such testing provides a basis for refining the model before it is used to draw conclusions about BMT performance or translated into an application. Each revision needs to be documented so that changes in indicators, rules, and assessment results can be traced across versions.

6. Conclusion

This study develops an initial written formulation of a BMT management performance evaluation model by adapting a banking management assessment framework, incorporating the characteristics of Islamic cooperatives, and enriching the organizational culture component. The model comprises ten components translated into a prototype instrument with 40 assessment items, supported by evidence guidelines, an initial rubric, calculation procedures, interpretation of results, and subsequent improvement actions. This structure provides a basis for connecting the management aspects being assessed with the examination process and institutional improvement needs.

The written model is a preliminary research output developed before the design of a digital application. Its contribution lies in defining the evaluation content and rules that can subsequently inform the requirements of an application for assessing BMT management performance. Expert validation and usability testing are needed to refine the indicators and assessment procedures before wider implementation. The study therefore provides a substantive foundation for the subsequent design and development of a digital application.

7. References

Amar, Faozan, Pepi Januar Jelita, and Deni Pandu Nugraha, 'The Relationship Between Strategic Management Implementation and Decision-Making Style in Enhancing Employee Satisfaction: Evidence from Islamic Cooperatives', *Jurnal Manajemen Bisnis*, 16.2 (2025), pp. 508–25

Amin, A. Ridwan, 'The Celestial Management: Zikr, Pikr, Mikr', *Language*, 33.260p (2004), p. 18cm

Aqqunduz, Ahmed Tauziat, and Dyah Pikanthi Diwanti, 'Pengaruh Kualitas Sumber Daya Manusia, Kompensasi Dan Pelatihan Kerja Terhadap Kinerja Karyawan Pada Kspps Bmt Bina Ihsanul Fikri Di Yogyakarta', *Nusantara: Jurnal Ilmu Pengetahuan Sosial*, 12.8 (2025), pp. 3636–43

Bank Indonesia, 'Sistem Penilaian Tingkat Kesehatan Bank Umum.', 2004 <https://www.bi.go.id/id/archive/arsip-peraturan/Documents/0151b17420f84d118de8fdf0c0642730se623dpnp.pdf?utm_source=chatgpt.com>

Boateng, Godfred O., and others, 'Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer', *Frontiers in Public Health*, 6 (2018), p. 149

Darma, Emile Satia, and Akhsyim Afandi, 'The Role of Islamic Corporate Governance and Risk toward Islamic Banks Performance: Evidence from Indonesia', *Journal of Accounting and Investment*, 14 November 2021, pp. 517–38, doi:10.18196/jai.v22i3.12339

Di Luozzo, Sebastiano, and others, 'Key Activity Indicators: Critical Review and Proposal of Implementation Criteria', *International Journal of Productivity and Performance Management*, 72.10 (2023), pp. 3000–26

Hristov, Ivo, Antonio Chirico, and Riccardo Camilli, 'The Role of Key Performance Indicators as a Performance Management Tool in Implementing Corporate Strategies: A Critical Review of the Literature', *Financial Reporting*, no. 2022/1 (2022) <https://www.francoangeli.it/riviste/Scheda_rivista.aspx?IDArticolo=70951> [accessed 27 October 2026]

Inayati, Saptaningrum, 'Evaluasi Kinerja Manajemen Koperasi Syariah Berdasarkan Peraturan Menteri Negara Koperasi Usaha Kecil Dan Menengah Republik Indonesia Nomor 35.3/PER/M. KUKM/X/2007: Studi Kasus BMT Ben Makmur Desa Kenduren Kecamatan Wedung Kabupaten Demak Tahun Buku 2013', *El-Qist: Journal of Islamic Economics and Business (JIEB)*, 6.2 (2016), pp. 1279–97

Jaakkola, Elina, 'Designing Conceptual Articles: Four Approaches', *AMS Review*, 10.1–2 (2020), pp. 18–26, doi:10.1007/s13162-020-00161-0

Kraus, Sascha, and others, 'Literature Reviews as Independent Studies: Guidelines for Academic Practice', *Review of Managerial Science*, 16.8 (2022), pp. 2577–95, doi:10.1007/s11846-022-00588-8

Mutamimah, Mutamimah, Zaenudin Zaenudin, and Widiyanto Bin Mislan Cokrohadisumarto, 'Risk Management Practices of Islamic Microfinance Institutions to Improve Their Financial Performance and Sustainability: A Study on Baitut Tamwil Muhammadiyah, Indonesia', *Qualitative Research in Financial Markets*, 14.5 (2022), pp. 679–96, doi:10.1108/QRFM-06-2021-0099

Nazaruddin, Ietje, Hafiez Sofyani, and Tiyas Puji Utami, 'Promoting Performance Measurement System Effectiveness in Higher Education Institution: Antecedents and Consequences', *BAR-Brazilian Administration Review*, 21.2 (2024), p. e230070

Nugraheni, Peni, and Candra Dewi Nasiroh, 'Improving the Performance of Baitul Maal Wat Tamwil (BMT) Employees: Do the Principles of Good Corporate Governance (GCG) Matter?', *Journal of Accounting and Investment*, 24.3 (2023), pp. 959–73

Sudarmanto, Eko, and others, 'Transformasi Digital Dalam Keuangan Islam: Peluang Dan Tantangan', *Jurnal Ilmiah Ekonomi Islam*, 10.1 (2024), pp. 645–55

Tuunanen, Tuure, Robert Winter, and Jan vom Brocke, 'Dealing with Complexity in Design Science Research: A Methodology Using Design Echelons', *MIS Quarterly*, 48.2 (2024), pp. 427–58

Wijayanti, Provita, and others, 'The Integration of Maqasid Sharia Principles on Digital Accounting Information System in Indonesian Islamic Microfinance', *Journal of Islamic Accounting and Finance Research*, 7.1 (2025), pp. 107–26

Yaya, Rizal, and Wahyu Manuhara Putra, 'Sistem Informasi Kinerja Berbasis Maqasid Syariah Untuk Peningkatan Kinerja BMT UMY', *Prosiding Seminar Nasional Program Pengabdian Masyarakat* <https://prosiding.umy.ac.id/semnasppm/index.php/psppm/article/download/438/482> [accessed 27 October 2026]