INTERNATIONAL JOURNAL OF ADVANCED RESEARCH
International Peer-Reviewed Research Journal
ISSN (Online): 3106-7980 | ISSN (Print): 3106-7972
Volume 3 • Issue 7 • November 2026

SUSTAINABLE PURCHASING AND SUPPLIER SELECTION THROUGH IMPLEMENTATION AND PERFORMANCE EVIDENCE A CRITICAL INTEGRATIVE REVIEW

Muhammad Usman[*1], Dr. Saima Hassan[2]

[*1]*PhD Scholar in Business Administration, NCBA&E, Lahore, Director Purchase, Government College University, Lahore, Pakistan,*
[2]*Associate Professor in Business Administration, NCBA&E, Lahore, Pakistan,*

[*1]dr.usmanbuttar@gmail.com, [2]saimahassan@ncbae.edu.pk

DOI: https://doi.org/10.5281/zenodo.24102377

Keywords
Sustainable purchasing; supplier selection; supplier development; sustainability performance; emerging economies; integrative review

Article History
Received: 05 August 2026
Accepted: 19 October 2026
Published: 10 November 2026

Copyright @Author
Corresponding Author: *
Muhammad Usman

Abstract
Sustainable purchasing involves organizations needing to capture supplier assessment requirements in actions that drive positive environmental, social and business results. This article critically reviews the published literature related to supplier selection, supplier management and supplier performance taking into consideration manufacturing supply chains and implications for Pakistan. There is a purposeful analysis of credible journal sources between decision models, observational studies, qualitative case studies and previous syntheses. The analysis finds a separation of evidence when looking at developing a supplier scorecard and making a procurement decision and further improvement. The selected literature gives considerable focus to environmental outcomes and the social outcomes are explicit with benefits apportioned among parties – buyers and suppliers. Pakistan evidence should lead to further research on green purchasing and organizational capabilities, but far from an overall causal conclusion, there are variations depending on outcomes and research designs. The article builds an integrative framework of the criteria, selection, implementation/verification, and suggests non-compensatory criteria, supplier development and context. The main takeaway for purchasing managers is to integrate proportionate evaluation with believable information, realistic improvement programs and monitoring the results. The results of the review can be called analytical propositions, as it does not provide an exhaustive census of the field.

1. INTRODUCTION

Purchasing decisions determine who have access to a buyer's orders, which capabilities to support and which production practices to tolerate in a supply relationship. Sustainability thus gets into purchasing via concrete decisions on specification, supplier eligibility, supplier evaluation and ongoing supplier relationships. A supplier can provide cheap prices and good delivery, but may cause environmental or social issues that are not included in the purchasing scorecard. On the other hand, a supplier with feasible improvement prospects can find it hard to keep up with the documentation requirements of the buyer. These decisions need to be understood in the context of what is being chosen, how the relationship is handled and whose results are then evaluated."

The concepts and principles of Sustainable Supply Chain Management go beyond just

compliance with the environment. Carter & Rogers (2008) relate environmental, social and economic issues in a supply chain. Seuring and Müller (2008) differentiate between supplier management related to risk and performance and supply chain management related to sustainable products. These contributions show that sustainability can relate not just to the nature of a product, but also to the way that it is produced. In the case of research, the implication is that purchasing a supplier can't be judged without considering the relationship and production arrangements by which the purchased product or service is delivered.

There are also published reviews of supplier-selection techniques that offer significant coverage. Govindan et al. (2015) explore the criteria for green evaluation and the decision methods, and Karakoç et al. (2024) categorize the sustainable supplier-selection research based on the method and extensions. This work aids researchers pinpoint the means for combining competing criteria. When it results in an "elected" supplier, it leaves another question for an integrative review – what can be inferred about sustainability? Ranking is the result of a decision procedure. How orders are ordered, and whether suppliers change their practices, and whether there are improvements in relevant outcomes are further questions that need different evidence.

The difference is important for the understanding of the research in Pakistan. According to Rubab et al. (2024) there is a positive association between green purchasing and environmental performance. Khan and Qianli (2017), however, find for their sample in Pakistan that the relationship between green purchasing and organizational performance is statistically nonsignificant. It is important to not conclude that green purchasing is either effective or ineffective from these outcomes. Environmental performance and organizational performance are two different constructs and studies have different samples and analytical specifications. However, their juxtaposition does not automatically imply their direct opposition, but rather a careful examination of the outcome definitions.

The supplier relationship gives a valid link between a purchasing decision and an outcome thereafter. Gimenez and Tachizawa (2012) discuss two processes (assessment and collaboration) to spread sustainability to suppliers. Zimmer et al. (2016) explicitly take into account selection and monitoring and development. So it is not the purpose of the present article to claim for the first time the introduction of these connections. It makes its contribution by interpreting selected work from these research streams with the aid of an explicit distinction between decision evidence, implementation evidence and outcome evidence. This attitude helps a business administration viewpoint where the credibility of performance claims, together with organizational feasibility, is taken into account in deciding the quality of a decision method.

The review consists of four questions. Firstly, what is the representation of the economic, environment and social aspects within supplier evaluation? Secondly, what do the various types of validation mean in terms of their supplier-selection methods usefulness in practice? Third, how might capabilities become outcomes via purchasing and supplier-management practices? Fourth, what are the boundaries of the implications which can be reasonably drawn with respect to Pakistan in view of context and research design. These questions will help you to critically synthesize, not to statistically aggregate. The primary contexts are manufacturing supply chains, and applicable wider theoretical and relational studies explain mechanisms pertinent to purchasing.

The concepts and theoretical lenses are then defined, the source-selection discussed, and the analytical approach described, along with a mapping of the functions of various types of evidence. Thematic synthesis based on criteria, validation, implementation and outcomes. Propositions for future research are then developed by using an integrative framework and implications for purchasing managers and a research agenda are presented. Given throughout, reported findings are underlined, and articles own analytical deductions are indicated. The aim is to ensure that any claims made about sustainability are more useful by identifying what has been proven and what is not.

2. Concepts and theoretical foundations

2.1 Sustainable purchasing and the unit of analysis

The communication boundary search is a wide boundary: information that a buyer can gain from a supplier on his/her environmental or social characteristics, including claims, disclosure, certification, reputation or remedial action. These are not necessarily credible signals, but are signals to consider. Vinayavekhin et al. (2026) point to observability, meaning and cost as factors to consider when justifying a signalling mechanism. Costs would be costs incurred in the production or maintenance of evidence or consequences of misrepresentation and not merely the product of a claim. The conceptual analysis then dissects communicated information, substantiated attributes and perceived credibility of the buyer. The wide eligibility range is kept in order to be able to consider unsupported claims and greenwashing as well.

The starting point in analysing information asymmetry and receiver's interpretation of a signal is in signalling theory (Connelly et al., 2011). For the purpose of this review the relevant receiver is a buyer in an organisation. There may be more information about his practices available with the supplier and it will be the buyer's judgement as to whether this information is credible and relevant. Therefore, a separation of supplier attribute, the communication and the buyer's perception is done in the analysis. The survey measurement of perceived sustainability in no way reflects the level of sustainability of that supplier.

2.2 Capabilities and interfirm relationships

Hart's (1995) natural-resource-based view provides an understanding of how pollution prevention, product stewardship and sustainable development are related to organizational resources and competitive advantage. It gives a motivation to look at capacities that underlie environmental performance instead of external costs, which are often the needs. Reuter et al. (2010) apply the dynamic-capabilities perspective to sustainable supplier management on a global level in the chemical industry via case studies. All of these add up to more than just the instrument for scoring; for the companies to have purchasing capabilities, they must also notice the changes that are relevant, form routines and change how they manage their suppliers.

Dyer and Singh (1998) identify assets, knowledge-sharing routines, complementary capabilities and governance as possible sources of competitiveness. Such a relational perspective can be used to understand why the pre-existing attributes of a supplier may not be enough to explain the results of the specific purchasing relationship. The buyer's technical knowledge may complement the supply's production capability and conditions of the relationship between them influence the possibility of knowledge utilization. When applied to the current review, this is a theoretical explanation that helps to analyze cooperation and resource commitments as well as selection criteria. It does not, in itself, constitute proof of the fact that cooperation always leads to better sustainability.

Pagell and Wu (2009) build a model of sustainable supply chain management based on ten exemplar cases that comprise of supply chain practices as well as less conventional behaviours. Their evidence suggests that consideration should be given to organizational structure arrangements that are not necessarily reflected in a simple efficiency comparison. The cases are about exemplars and not about the numbers or average returns of those arrangements. Such case-based theorizing is employed here to determine mechanisms and boundary conditions, and leaving comparisons of performance to studies that actually report evidence of relevant outcomes.

2.3 Social sustainability and accountability

Zorzini et al. (2015) identify socially responsible sourcing as a body of work in relation to upstream social issues and the underpinning theory. Klassen and Vereecke (2012) analyse social competence based on examples of monitoring, collaboration and innovation. These contributions however, call for the explicit addressing of social issues and do not take an environmental improvement as synonymous with sustainability. Some questions that are relevant to purchasing are who is liable for the cost of the adjustments, how workplace issues are surfaced and if those impacted by the supplier practices can provide information.

Answers can vary from those that a buyer would get from a supplier reliability assessment only.

Gualandis et al. (2015) differentiate between the processes of evaluation and verification, and present the issues of inclusivity, scope and disclosure of sustainability information. Their model focuses on the evidence reliability in supplier selection. Accountability is thus considered as a characteristic of the assessment process itself and its relation with the stakeholders concerned in the present article. That a large amount of information is reported does not necessarily mean that it is material, reliable or complete. Together, the theoretical lenses provide analysis of capability, relationship and evidentiary quality, which each pertain to a different aspect of the purchasing process.

3. Review approach

3.1 Design and scope

The article has critical integrative review design. Snyder (2019) describes how literature reviews can contribute to the synthesis and conceptual development if they are aligned with their research purpose. In this context the intent is to make comparisons between the types of conclusions that can be drawn from different types of sources and to create an explanatory frame. The sources were carefully selected not comprehensive. It comprises journal articles on sustainable purchasing, supplier selection, supplier governance and performance, foundational theory and a methodological reference. It is not possible to make estimates on the percentage of all studies that follow a specific procedure or report a specific outcome.

There are 27 journal publications altogether including the methodological source in the final bibliography. The range of these references covers basic theory to research published online as of 2026. Earlier theoretical work and reviews have been provided to place the argument historically and to situate it within a context of more recent empirical and methodological work. The basic empirical context is the manufacturing sector. The Swedish workwear purchasing case serves as contextual comparator due to an explanation of relationships between a manufacturer, an intermediary and SME customers. Interpretation is done through a broader review of the supply chain and not necessarily as a study just about purchasing.

3.2 Identification and source verification

We performed source discovery consisting of (a) targeted web search of publisher and university domains; (b) search in LJB archives; (c) SciSpace semantic literature search. The five searches were: search for families concerned sustainable purchasing and supplier selection, assessment, collaboration and development, social sustainability, performance in Pakistan and emerging economies and existing reviews of selection methods. Publications found in discovery were then checked using exact-title and DOI search. Sources retrieved provided additional leads in the form of relevant references. The publication connections, material available for each publication and bibliographic discrepancies are catalogued on the source record that accompanies the document.

The criteria for selection were that the publication must have been in a journal that was verifiable and that there must be a substantive link with at least one of the review questions. Preprints and unverifiable records were not considered published evidence. To increase the reference count, no broad sustainability articles, in which there was no obvious purchasing, supplier-management and/or theoretical contribution were added. Publisher records and/or institutional repositories were consulted to ascertain publication information. When an online date was used, but it was different from the journal issue date, the publisher's recommended date or the issue year was used. Records have been distinguished by means of the DOI identifier, avoiding the distinction of different hosting locations as different publications.

The available content was different from source to source. Some offered full-text portions of method, findings and limitations, and others offered bibliographic records and/or publisher or author-repository abstracts. For what is said there, only abstract-only sources are used, without reconstructing unobserved coefficients, measurement items or research procedures. This difference in access is registered as publication verification and full methodological assessment

are two different processes. No subscription database export, comprehensive citation search, independent dual screening or interrater reliability assessment was done. For this reason, the article does not state that PRISMA is met or that an invented screening flow is presented.

3.3 Analysis and interpretation

The synthesis is performed regarding sources with respect to the following six dimensions: purchasing activity analyzed, dimension of sustainability, organizational context, research design, form of validation and outcome boundary. This comparison is arranged in three analytical categories. Decision evidence relates to the building or enacting of an evaluation procedure. Implementation evidence is for the purchasing and supplier-management practices, whether they are in place and how they are implemented. Outcome evidence relates to environmental, social, operational or financial outcomes of those practices. A study may have more than one category, but evidence in one category does not necessarily count as evidence in another.

Appraisal is research design specific and qualitative. Decision models are analysed with regard to their data and validation claims. Surveys are analysed, self-reporting is taken into consideration, the unit of analysis is noted and temporal or causal inference is considered. Cases provide information on mechanism and context, but not on population level impacts. Previous reviews plot set-up arguments and findings and are not regarded as independent replications on the primary studies from which they are derived. Pooled effect size, vote count of significant findings or numerical quality score is not calculated. The comparison results in the conceptual framework and the propositions which are still to be tested empirically.

4. Structure of the selected evidence

The chosen sources are used in various analytical capacities. Foundational contributions: define sustainability and explain capabilities/relationships. Review articles structure existing methods and discover what the conceptual space already explored are. Decision-model applications are examples of how criteria are turned into an ordering of alternatives. Empirical purchasing and performance studies are studies that investigate relationships or offer descriptions of organizational processes. This classification does not assume that methodological diversity means that each publication was addressing the same effectiveness question. This explains also the inclusion of different dated sources and the different levels of analysis.

Selected earlier reviews are compared with the article in Table 1. As a comparison, it shows that the selection methods, supplier development and performance have been attended before. The present synthesis thus attempts a more limited interpretive endeavor: that of clarifying the inference which connects one phase of the purchasing process with another. It poses the question: Is it based on evidence of implementation or on correlations between reported practices and outcomes or on characteristics of a proposed selection procedure? This separation may indicate where a convincing managerial recommendation might still need further empirical reinforcement.

Table 1: Positioning the review against selected prior syntheses

Source	Established coverage	Role in this article
Miemczyk et al. (2012)	Definitions and measures across dyads, chains and networks.	Clarifies the organizational boundary of a performance claim.
Govindan et al. (2015)	Green supplier criteria and decision approaches.	Provides a baseline for interpreting selection methods.
Zimmer et al. (2016)	Models for selection, monitoring and development.	Establishes that integration of these activities is not a new idea.

Source	Established coverage	Role in this article
Geng et al. (2017)	Meta-analysis of green supply chain practices and performance.	Provides aggregate evidence while preserving distinctions between practices.
Schramm et al. (2020)	Sustainable selection approaches and compensation effects.	Motivates examination of trade-offs and minimum requirements.
Karakoç et al. (2024)	Selection methods in 101 papers covering 2018–2022.	Locates the contribution beyond a catalogue of techniques.
Gidiagba et al. (2025)	Machine learning applications in sustainable selection.	Situates digital tools within an existing research stream.

Table 2 shows the selected primary studies' roles and limitations. It contains manufacturing applications as well as the relational comparator mentioned in the scope. Entries are not meant to be an exhaustive evidence table for the field. They are selected to bring the methodological comparisons which underpin the argument to the fore. Specifically, the use of an industrial application can prove a method's usability, but not necessarily that its recommendations can lead to better conditions over time. Similarly, when the relationship between observed practices and performance is observed, this can lend weight to a hypothesis, but other explanations and reverse causation remain.

Table 2: Selected primary studies and the interpretation of their evidence

Source	Design and setting	Evidentiary contribution and boundary
Luthra et al. (2017)	Industrial supplier-selection application using AHP and VIKOR.	Demonstrates an evaluation framework; application does not itself establish subsequent sustainability gains.
Tong et al. (2022)	Extended PROMETHEE II applied to a Chinese SME.	Illustrates selection under SME constraints; broader outcome claims require follow-up evidence.
Sancha et al. (2016)	Survey of 120 Spanish manufacturers.	Distinguishes buyer and supplier social performance; association does not establish the full causal sequence.
Khan and Qianli (2017)	Study of 218 Pakistani manufacturing firms.	Reports a nonsignificant green-purchasing relationship with organizational performance; this does not establish harm.
Rubab et al. (2024)	Pakistan survey with 214 usable responses.	Reports green purchasing linked to environmental performance; manufacturing framing includes a mixed reported sector profile.
Bashir et al. (2024)	Environmental-management study of 243 Pakistani textile companies.	Provides contextual performance evidence; it does not directly compare supplier-selection choices.

Source	Design and setting	Evidentiary contribution and boundary
Junejo et al. (2025)	Cross-sectional survey of 469 SME employees in Pakistan.	Examines capability-related associations; employee responses should not be equated with distinct firms.
Vellesalu et al. (2026)	Swedish workwear case involving a manufacturer, intermediary and four SME customers.	Explains relational purchasing processes; transfer to other settings requires contextual evaluation.

The geographical range of the chosen studies allows for a comparison in the context, but not for a ranking of the countries according to the maturity in sustainable purchasing. There are differences in the studies in terms of sample, constructs, and purpose. Results concerning SMEs are also reported separately from those gathered from larger firms or with a mix of firms. This review by Sánchez-Flores et al. (2020) highlights the importance of studying context and provides the following avenues for further research: collaboration, sourcing and supplier development. This helps to explain local conditions without using a description of a country label used as an after-thought in the analysis.

5 Thematic synthesis

5.1 Criteria and the meaning of a sustainable supplier

Criteria reveal what a decision maker sees of a supplier. A scorecard comprising resource use, worker safety and delivery performance is a different decision problem to one that focuses on commercial continuity. While both may be termed sustainable, the same term doesn't make them equivalent. The defensible synthesis needs to check the dimensions that were actually assessed, the evidence that represented them, and the implications of the weighting process. The present review views sustainability as a characteristic of the criteria and outcome measures of the model and not solely as a characteristic of the name of a study.

In the paper of Schramm et al. (2020), 82 papers related to sustainable supplier selection are analyzed and the compensation effects that supplier selection methods can create are explicitly discussed. This is a practical consequence of this issue. A fully compensatory assessment is one in which high scores on some attributes compensate for low scores on others. That is a substance purchasing decision rather than a mathematical choice. For instance, maybe the buyer is willing to accept some flexibility in delivery for a lower price and decides to make an important safety factor a requirement for eligibility. In the example, this is a design decision, not a statement that all published methods allow for any kind of exchange between the criteria.

This article therefore makes a distinction between minimum conditions and comparative attributes. The minimum conditions are used to identify whether a supplier is considered or if corrective action needs to occur prior to an award. Suppliers with the same attributes are distinguished by comparative attributes. Improvement potential is what a supplier is able to realistically do with support. These categories should be identified prior to making a final ranking as they address different questions. If not, a single aggregate score might otherwise mask whether a supplier is eligible, desirable or capable of improving. The categories are suggested here as a means of analyzing the data and are not intended to be a universal category.

Boundaries are also a part of the design of the environmental criteria. The energy intensity per unit, the total energy use and the environmental characteristics of a purchased material are different measures. One might make a purchasing decision that will make one better but not the others. If criteria are derived from certification, or written procedures, the relationship assumed between the procedure and desired outcome should be stated in the assessment. This produces a question for monitoring which can be tested: “Do the characteristics used at selection predict the performance the buyer wanted to achieve?” It also enables the organisations to alter a criterion

where it is found to be weak but useful from an administrative point of view.

Social criteria need to be similar in specificity, and with an added focus on whose account is represented. There are differences between the information sources, namely: buyer perceptions of ethical performance and supplier documentation and worker experiences. Tapping them as a single entity would make the disagreement and possible blind spots invisible. The social sourcing literature Zorzini et al. (2015) discovered provides a foundation for making human and community concerns substantive purchasing questions. The present synthesis, continuing that distinction, suggests a social criterion to determine the affected group, the condition under evaluation and the means by which the condition is manifest.

5.2 Decision methods and the limits of validation

Decision methods can be used for making explicit complex comparisons. Luthra et al. (2017) merge the analytic hierarchy process with VIKOR in an application with 22 evaluation criteria. Tong et al. (2022) propose an extension of PROMETHEE II for an SME case. The following studies demonstrate the use of structured preferences and supplier information to assist with a choice. They should be evaluated based on that decision task. The fact that an industrial case exists does not, in itself, demonstrate that the selected supplier later made an improvement in environmental or social performance compared to an alternative supplier that was deemed to be viable.

The difference between validation tasks is therefore of key importance. Computational checks are checks that ensure a procedure functions as intended. Sensitivity analysis is a study which considers how conclusions may vary if other assumptions are made. Based on expert assessment, a problem representation may be determined to be meaningful to practitioners. Implementation evidence indicates if the decision makers utilize the resulting recommendation. Outcome evaluation looks at the post-purchase consequences. These types of validation can be combined, yet they set up various things. Even if a strong ranking is achieved based on a small number of weight adjustments, it may still be poorly supported when the supplier data are flawed or the criteria are missing a material concern.

Karakoç et al. (2024) explain a wide variety of traditional and hybrid selection methods. Because of the diversity, a technique cannot be recommended just because it has been used and published many times. A purchasing organization must appreciate the requirements of a method with regard to the demands it puts on information, judgement and maintenance. The analytical implication developed here is to link in some way methodological complexity to the decision and evidence available. Other calculations should help clarify a reasonable ambiguity or enhance a useful decision characteristic and not be a standalone measure of research quality.

Machine learning introduces related questions. This is already a topic of review as Gidiagba et al. (2025) review its applications to sustainable supplier selection. Even if the model in the framework developed here is predictive, it would still need to have an explanation for the nature of the variable it is predicting. The prediction of a past supplier rating is not the same as an independently measured sustainability improvement. A model can simulate past patterns of assessment but not say that these are good patterns. Research on the effectiveness of these tools should therefore differentiate between predictive performance, usefulness in the purchase decision and the impact of the purchase decision made by the tool.

In the case of SMEs, the resources needed to maintain a method (or process) should be included in the comparison. A model might have a well-built research sample and might be great with that data, but it might require data that the purchasing team isn't able to get on a regular basis. However, a simpler solution might be more sustainable and less complicated, and leave out critical uncertainty. It is a trade-off that is researchable instead of a call for simplicity in all settings. Data sources, frequencies of data updating and treatment of missing data in the studies should be disclosed. These details will make a difference in whether or not a proposed method will become a continuing organizational practice.

5.3 Implementation through assessment and supplier development

The level of access to orders are dependent on supplier selection and the terms of orders determine the feasibility of subsequent improvement. A buyer can ask for a commitment to environmental investing, and at the same time, set a short term of commitment; or a buyer can ask for an improved working environment and at the same time, introduce an unpredictable variation in demand. Here are some analytical examples of possible inconsistent purchasing signals. It would be observable to have the specifications, the commercial arrangements and the improvement requirements evaluated as consistent. Whether or not the relationship provides the supplier with a viable chance to provide the performance on which it was evaluated.

In particular Sancha et al. (2016) offer a useful stratification of the 'where' of social performance. In their study, they believe that the supplier assessment is linked to the buyer's social performance, and they believe that the collaboration with supplier social performance is linked to supplier assessment. Results highlight the need to be wary of using a positive buyer-level metric to reflect the conditions inside the supplier. They also suggest reasons for analysing assessment and development separately. One can enhance the buyer's capacity for determining and dealing with exposure, whereas the other deals with activities nearer to changing supplier practices. The causal order of these activities and outcomes is not conclusively determined by the observational design.

The act of cooperation is to be critically examined. Lund-Thomsen and Lindgreen (2014) doubt the effectiveness of cooperative corporate-responsibility strategies in changing the power dynamics to achieve long-term improvement in developing-country value chains. The bottom line for the current scheme is that if collaboration is to be described then it must be done operationally. Different interventions: joint meetings, technical assistance, financing and changes in purchasing commitments. The term collaborative gives little indication as to who can make decisions or who pays for the adjustment. A meaningful account includes information about what resources are being exchanged, what discretion is left and what happens if agreed improvement does not take place.

The purchase arrangements of a Swedish workwear manufacturer and the distributor and SME customers are discussed by Vellesalu et al. (2026). Their case connects relationship organization and user involvement with work to prolong the lifetime of products. It offers a mechanism-oriented comparator to look at the intermediaries' purchasing activities coordination. Its single case design and its context limit its applicability to manufacturers in Pakistan. The important question is not whether the arrangement would be desirable for copying, but which coordination functions could conceivably be required anywhere else and which organizations could reasonably provide them given different commercial circumstances.

Based on these observations, supplier development is considered as a process with well-defined content, and therefore, as an intervention. In a study, the first constraint, the support offered, the actor who would make the change and the timing of the change should be identified. This information could make it possible to compare between, for example, technical training and improvements in order planning. It would also help to determine if a perceived lack of progress is due to an inappropriate supplier or to a lack of support or mismatch between purchasing needs and the supplier's approach. The proposed distinction moves the spotlight of analysis from the existence of the supplier-development programme to the changes it makes to the relationship.

5.4 Performance measures and interpretation of associations

There's no one single end of performance. Environmental outcome could be related to resource consumption, emissions or waste. Social outcomes can be related to either safety or working conditions, or distribution of opportunity. Operational results are those related to aspects like quality and reliability and financial results can involve costs and revenues or profitability for various time periods. With these dimensions together may be helpful for a well-established objective, but additionally needs

to be a policy for how to deal with contradictory modifications. The review thus sees the results at the level of those outcomes reported, rather than turns the results into a general judgements of success.

Geng et al. (2017) review 50 studies related to green supply chain practices in the context of Asian emerging economies and conclude that green supply chain practices positively relate to various aspects of performance. They also look at contextual moderators. Their meta-analysis provides evidence on a broader level than a single case; however, the results of their analysis should not be interpreted as the impact of a particular supplier-selection intervention. The practices that are part of a broad green supply chain can be implemented in production, design or logistics and also in purchasing. An outcome of the combined construct does not necessarily indicate an outcome of supplier selection within it.

Relevant evidence from the Pakistani textile sector presented by Bashir et al. (2024) is related to environmental management systems (EMS), environmental performance (EP), customer satisfaction and business performance (CP). This study contributes to a picture of potential organizational relationships between environmental performance and commercial results. It is not a direct measure of the effectiveness of a supplier selection decision. Likewise, Junejo et al. (2025) study on green supply chain management, knowledge sharing, innovation and data related capabilities in SMEs of Pakistan. Their cross-sectional, self-reported evidence is supportive of the investigation of capability-related explanations, but the temporal ordering and causal contribution of individual purchasing practices remain unresolved.

These variations in research designs give rise to several inferential problems. Firms with more resources may be more capable with respect to implementing sustainability practices and may be more likely to report better performance. A purchasing decision might be affected by an external change in the market and so might an outcome. Respondents can also judge practices and outcomes via a shared positive or negative first impression. These are suggested alternative explanations – not that any particular finding is incorrect. A good empirical design must consider the alternatives in its context, and the consequences for the explanations that remain open should be stated. That task can't be done with statistical significance alone.

Another source of no comparability is time. An assessment can have an immediate administrative cost, changes in equipment can have an impact on results, as can learning. The relationship estimated in a study conducted shortly after adoption does not necessarily be the relationship estimated in a study conducted several years after adoption. Financial results may also be a result of who pays for an intervention, and not the whole value of the supply chain. This analytical implication is to indicate when the implementation took place, when observations were carried out and, if this information is available, the costs of the implementation. This would help to describe the literature as a comparison between delayed benefits and persistent trade-offs, instead of as a comparison of different national context.

5.5 Verification and feedback

Verification is the process of linking information and data used to select suppliers with information and data used to assess them in the future. The design of this connection is considered as an independent management task, as done in Gualandis et al. (2015). The organisation should be aware of what claims are being ticked, who is doing the ticking and what the result of the ticking means for a purchasing decision. A requirement that does not affect an award, an improvement plan or a renewal decision may be retained administratively, but has unclear operational implications. On the other hand, if the same questionnaire is repeated as monitoring, a material change of the supplier's behavior might be not noticed.

Feedback should enable an organization to understand if its criteria had the desired effect. When a supplier with a high initial score fails over and over again to achieve an expected result, then the response should involve investigation of the scorecard and the purchasing relationship. The Evaluation should reflect the improvement if the initial provider(s) make significant improvement(s) after development. This does not mean it is ok to reduce standards when performance is less than

what is desired. It involves the difference between legitimate requirement and the ineffectual indicator, and the difference between supplier responsibility and limitations imposed on the supplier by buyer's decision. The resulting learning process is not an empirical result of this review, but rather a proposed feature of the framework.

6 Integrative framework and research propositions

6.1 Connecting criteria with outcomes

This article's framework links four activities: Declaring sustainability criteria, choosing and allocating business, carrying out purchasing and supplier-development practices and checking results. Feedback from verification is given back to the former activities. This is a conditional connection as each transition may fail independently. It is possible to set up suitable criteria but measure inappropriate data, to select a competent supplier but propose unsuitable commercial terms, or to implement a beneficial intervention but measure too far removed an outcome to determine its contribution. Favourable results at one stage are not a guarantee of success throughout the process, as shown in these possibilities.

The proposed connections are shown on Figure 1. The figure is an analytical synthesis, rather than an estimated structural model. The availability of resources to buyers and suppliers and the nature of the purchased category and the distribution of the decision authority impact each stage. Quality of evidence also exists throughout the process. It has a bearing on the inferences that can be drawn at selection, the observability of implementation and the attributability of an outcome to the intervention. These conditions are made explicit as there is no assumption made that the adoption of a named sustainability practice brings an automatic performance gain.

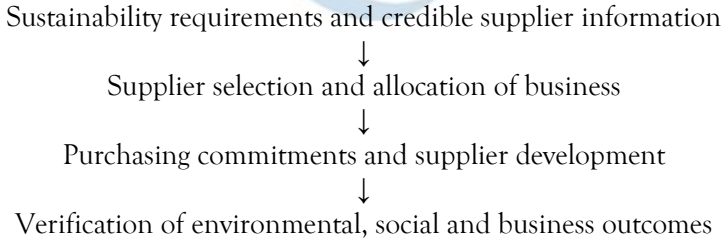


Figure 1: Proposed connections between purchasing decisions and outcomes

The framework builds on past supplier-management integration by focusing on transitions relating to evidence. It asks: "If you have one of these claims that a procedure identifies a desirable supplier, what further observation is necessary to claim that purchasing improved sustainability? Implementation is based on a selection model, which in turn needs information on actual adoption. Outcome observations are necessary before implementing practice can result in improvement. For an observed improvement to lead to a causal conclusion there must be a credible account of other explanations. The framework is not set on any specific type of research design for each question, but each question must align with the evidence.

6.2 Propositions for empirical investigation

The following propositions delineate testable relationships that are posited by the synthesis. They are intentionally conditional and should not be viewed as results already shown in the publications selected. The capabilities and relational perspectives discussed above offer a conceptual underpinning, and the mixed research designs provide a guide as to which relations would require some further direct investigation.

Proposition 1. Supplier-selection procedures that separate minimum sustainability from comparative preferences will result in fewer selections that fail to meet performance criteria for a designated critical condition, than will supplier-selection procedures which allow unrestricted compensation across all criteria. A

test would have to have the supplier information as well as the decision context fixed and would have to specify the critical condition up-front and compare selections. It should also consider the commercial implications and not presume that enhanced protection is free. The proposition is about decision properties, if an additional follow-up is desired to account for subsequent supplier behaviour, it should be done separately.

Proposition 2. When supplier commitments and development resources are available to support the changes needed, there will be greater correlation between sustainability based supplier selection and subsequent supplier improvement. This proposition focuses on the process of implementation as an intervening process. Can be explored through commitment, resources, and change over time as opposed to surveying all constructs at one point in time. Other reasons might be that suppliers were already capable and buyers were picking suppliers that they thought were likely to improve anyway. Taking up these explanations would make an evaluation of the contribution made by the purchasing relationship more robust.

Proposition 3. Information from multiple and relevant sources will minimize inconsistencies in reported supplier sustainability compared to independent observations due to a single organizational respondent. The statement is about information quality NOT any extra information is useful. Sources need to be pertinent to the condition being investigated and disagreements need to be discussed. When it comes to social results, supplier management and affected workers might offer different insights. Different measurement boundaries may be identified through an internal assessment and through independently recorded operational data for environmental outcomes.

Proposition 4. The type of supplier evaluation and development that can be sustained by SMEs will be dependent on the resources available and a reliance on a few commercial suppliers. Performance is assumed to be linked to the effort of implementation. Where significant specialist involvement is required, an intervention can function differently if it is shared across various business functions. These relationships could be explored in comparative research to see if there is a difference between the two when intermediaries or shared services or buyer support are involved. The easiest way to do this would be to look for new capability, but not the dependency on the organization that provides it.

These propositions are to establish a linkage between the research questions which are frequently explored using different methodologies. Decision experiments can explore the results of alternative evaluation rules. Relationships can be determined in process studies and the changes in relationships with implementation. Under appropriate conditions outcomes may be examined by longitudinal or quasi-experimental designs. Also, the propositions allow for a null or adverse determination. Benefits to a buyer from a practice in documentation, but not to a supplier in costs and at no change in other relevant conditions, should not be "silenced" in an overall sustainability score.

7. Implications for Pakistan and future research

7.1 Implications for purchasing managers

The framework allows Pakistani purchasing managers to make a series of decisions to be adjusted based on the category and supplier relationship. The first step is to identify the outcome of interest and the range of possible variation that will be considered as progress. Questions of sustainability might vary for a textile purchasing team, a manufacturer buying equipment and an SME buying workwear. The review does not provide a set of criterion weights for these settings on the national level. It rather offers a justification for making the connection between the requirements and purchasing decisions and the monitored results transparent. What comes next is the ability to differentiate between necessary factors and factors that can be compared. Buyers should be able to tell them why they have put it in the requirement, what proof they will take for that requirement, and what the implications are if they are uncertain. This method also aids determine information that is being collected repeatedly, but not affecting buying. Lessening these kinds of admin workloads can free up time to focus on material issues (the impact needs to be assessed, not taken

for granted). If specialized evidence is required, the team should see if the supplier is able to get it and if a shared verification is possible.

Development arrangements should outline what changes should be anticipated from both parties. A supplier improvement plan can involve equipment, training, or changes in work organization, but can also rely on more reliable order information, and reasonable implementation times. These can be recorded by purchasing managers when agreement is reached upon the plan. They are then able to compare progress with the actual commitments. The recommendation is based on the framework and not necessarily that any one contracting model is found to be the best in Pakistan. Depending on the purchasing category, risk and commercial relationship, the appropriate arrangement should be made.

Outcome monitoring needs to differentiate if the buyer or supplier is the one undergoing change. For example, if the buyer's audit effort is reduced, this should be represented as better working conditions with supporting evidence. In the same way, moving an expensive activity from a buyer to a supplier does not mean a total reduction in the supply chain cost. The presence of benefits and when they occur should be documented so that trade-offs are readily apparent and managers have a better foundation for their discussions about corrective action. It can also help avoid the appearance of positive buying data when actually there are problems in other aspects of the relationship.

A pragmatic approach is better for SMEs than to presume that all the advanced techniques employed in research need to be an in-house system. Proportionality here refers to commensurateness with the material risk and the needs of a decision, with a meaningful requirement. Shared technical expertise, common information requests and coordinated training are possible areas of support, but need to be investigated locally. Industry associations and universities could be involved in such arrangements by running pilots with industry partners. The number of participating organizations is not the only factor that should be included in any assessment – the administrative cost, reliability of information and outcomes for suppliers should also be included.

7.2 A research agenda grounded in the framework

Researchers should first of all learn how to take the step from a selection recommendation to a real purchase. They could see whether consumers do follow a model's rank, what they say about the rules they break and if exceptions are instances of missing operating conditions. This work would help elucidate the implications of models improvements for practical applications. It would also identify criteria that seem convincing during expert elicitation that do not matter when capacity and/or timing and/or contractual obligations become hard constraints. Documenting the exit can be more beneficial to the depiction of the choice and not simply a result of the managers not utilizing the approach.

The second priority is to link selection with implementation in the long term. In a prospective study, suppliers could be assessed prior to being awarded the contract, development and purchasing commitments could be documented throughout the relationship and outcomes could be observed at multiple follow-up time points. The design would have to take into account baseline differences such as the rationale of offering development support to specific suppliers. Even in situations where it is not possible to allocate participants randomly, a clear comparison approach and due process documentation can enhance interpretation. The contribution would be in the form of a description of change and reasonable mechanisms instead of yet another simultaneous association of self-reported constructs.

Third, social outcomes require research designs which are able to compare accounts from different participants. Buyer surveys should be used in conjunction with supplier records, interviews and relevant worker level evidence. It is not to give the impression that one participant is invariably the one who gets the facts right, but to see what each can see and where each misses. Studies should also be made on how commercial changes go hand in hand with the social programs. Without that information, it is

difficult to differentiate between underperforming program design and improvement efforts hampered by the general purchasing dynamic.

Fourth, when comparing industries and firm sizes in Pakistan, context mechanisms being tested should be mentioned. Instead of viewing a country indicator as an explanation, studies may consider looking at issues of available information, resource limitations, purchasing reliance or the presence of intermediaries. Industry comparisons can be useful to identify the criteria and interventions that can be transferred and those that will need adaptation. Where there are multiple employees per SME, it is important that the method of determining the unit of analysis is reported in the SME research. Sample size and generalizability depend on identification of respondents, firms and buyer–supplier relationships.

Lastly, the framework can undergo more in-depth evidence review in the future. A systematic map of the proposed transitions could be realized by a preregistered search of bibliographic databases, duplicate removal, full text retrieval and independent screening. Where constructs, units and effect estimates are reasonably comparable, a meta-analysis should be attempted. A methodological rigor by itself would not dismantle the difference between a ranking study and an outcome study, the review would have to maintain the distinction. The present integrative account provides questions and categories for that latter work.

8. Limitations

Some journals, contexts and research traditions may be privileged by language of publication and/or by being available online. There were a variety of evidence in the reference set and the methodological detail was not always available. Those sources also make only limited descriptive claims, which can be accepted at the abstract level, but are not an exhaustive evaluation of internal validity. Independent screening and reproducible export of the bibliographic data have been absent, leading to the possibility that relevant publications have been missed. These limitations limit the conclusions of the review to an analytical synthesis of the sources identified.

The outcomes and units of analysis of the selected empirical studies also vary. The results from their research cannot be joined together to form an estimation of the impact of sustainable supplier selection in Pakistan. The individual organisation's expectation is not indicated by a positive association or by the lack of any statistical significance. The framework and propositions are therefore to be developed and tested. Practical recommendations should be put to use, bearing in mind the purchasing category and taking note of the outcomes. The article does not claim that the process proposed has been applied or tested in a Pakistani organization.

9. Conclusion

Sustainable purchasing includes considerations of purchasing criteria, supplier relationships, and evidence for performance evaluation. The first conclusion of the review is that a sustainability label has to be understood in the context of the real-life economic, environmental and social issues that are represented. The second is that the validation of a decision procedure and the evaluation of its effects afterwards are a response to different questions. Third, it is that assessment, development and commercial arrangements have relevance to the explanation of the relationship between selection of supplier and achievement of desired outcome. These connections are impacted by context, through capabilities, constraints and relationships that need to be explored.

In the case of Pakistan, there are studies to get evidence to start with and see a common causal statement. The integrative framework identifies potential sites of the additional observations that would enrich the argument: at actual purchasing, in implementation commitments, in independently credible measures and changes over time. It's usefulness is that it provides a way to consider the overall purchasing process while maintaining boundaries of responsibility and outcome. Progress can then be measured by the impact on the buyer, supplier and impacted stakeholders, not just by the sophistication of a selection method or the existence of a sustainability score.

REFERENCES

Bashir, A., Jawaad, M., Hasan, T., & Nasir, T. (2024). Linking environmental management systems to firm performance: Empirical evidence from an emerging economy. *The Lahore Journal of Business, 11*(2), 55–88. https://doi.org/10.35536/ljb.2024.v11.i2.a3

Carter, C. R., & Rogers, D. S. (2008). A framework of sustainable supply chain management: Moving toward new theory. *International Journal of Physical Distribution & Logistics Management, 38*(5), 360–387. https://doi.org/10.1108/09600030810882816

Dyer, J. H., & Singh, H. (1998). The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review, 23*(4), 660–679. https://doi.org/10.5465/amr.1998.1255632

Geng, R., Mansouri, S. A., & Aktas, E. (2017). The relationship between green supply chain management and performance: A meta-analysis of empirical evidences in Asian emerging economies. *International Journal of Production Economics, 183*, 245–258. https://doi.org/10.1016/j.ijpe.2016.10.008

Gidiagba, J. O., Tartibu, L. K., & Okwu, M. O. (2025). A systematic review of machine learning applications in sustainable supplier selection. *Decision Analytics Journal, 14*, Article 100547. https://doi.org/10.1016/j.dajour.2025.100547

Gimenez, C., & Tachizawa, E. M. (2012). Extending sustainability to suppliers: A systematic literature review. *Supply Chain Management: An International Journal, 17*(5), 531–543. https://doi.org/10.1108/13598541211258591

Govindan, K., Rajendran, S., Sarkis, J., & Murugesan, P. (2015). Multi criteria decision making approaches for green supplier evaluation and selection: A literature review. *Journal of Cleaner Production, 98*, 66–83. https://doi.org/10.1016/j.jclepro.2013.06.046

Gualandris, J., Klassen, R. D., Vachon, S., & Kalchschmidt, M. (2015). Sustainable evaluation and verification in supply chains: Aligning and leveraging accountability to stakeholders. *Journal of Operations Management, 38*(1), 1–13. https://doi.org/10.1016/j.jom.2015.06.002

Hart, S. L. (1995). A natural-resource-based view of the firm. *Academy of Management Review, 20*(4), 986–1014. https://doi.org/10.5465/amr.1995.9512280033

Junejo, I., Sohu, J. M., Alwadi, B. M., Ejaz, F., Nasir, A., & Hossain, M. B. (2025). Green supply chain management and SMEs sustainable performance in developing country: Role of green knowledge sharing, green innovation and big data-driven supply chain. *Discover Sustainability, 6*, Article 342. https://doi.org/10.1007/s43621-025-01055-6

Karakoç, Ö., Memiş, S., & Sennaroglu, B. (2024). A review of sustainable supplier selection with decision-making methods from 2018 to 2022. *Sustainability, 16*(1), Article 125. https://doi.org/10.3390/su16010125

Khan, S. A. R., & Qianli, D. (2017). Impact of green supply chain management practices on firms' performance: An empirical study from the perspective of Pakistan. *Environmental Science and Pollution Research, 24*(20), 16829–16844. https://doi.org/10.1007/s11356-017-9172-5

Klassen, R. D., & Vereecke, A. (2012). Social issues in supply chains: Capabilities link responsibility, risk (opportunity), and performance. *International Journal of Production Economics, 140*(1), 103–115. https://doi.org/10.1016/j.ijpe.2012.01.021

Lund-Thomsen, P., & Lindgreen, A. (2014). Corporate social responsibility in global value chains: Where are we now and where are we going?. *Journal of Business Ethics, 123*(1), 11–22. https://doi.org/10.1007/s10551-013-1796-x

Luthra, S., Govindan, K., Kannan, D., Mangla, S. K., & Garg, C. P. (2017). An integrated framework for sustainable supplier selection and evaluation in supply chains. *Journal of Cleaner Production, 140*, 1686–1698. https://doi.org/10.1016/j.jclepro.2016.09.078

Miemczyk, J., Johnsen, T. E., & Macquet, M. (2012). Sustainable purchasing and supply management: A structured literature review of definitions and measures at the dyad, chain and network levels. *Supply Chain Management: An International Journal, 17*(5), 478–496. https://doi.org/10.1108/13598541211258564

Pagell, M., & Wu, Z. (2009). Building a more complete theory of sustainable supply chain management using case studies of 10 exemplars. *Journal of Supply Chain Management, 45*(2), 37–56. https://doi.org/10.1111/j.1745-493X.2009.03162.x

Reuter, C., Foerstl, K., Hartmann, E., & Blome, C. (2010). Sustainable global supplier management: The role of dynamic capabilities in achieving competitive advantage. *Journal of Supply Chain Management, 46*(2), 45–63. https://doi.org/10.1111/j.1745-493X.2010.03189.x

Rubab, U., Khan, M. R., & Tufail, M. M. B. (2024). Fostering environmental performance through green supply chain practices: Empirical evidence from manufacturing firms in Pakistan. *The Lahore Journal of Business, 11*(2), 89–114. https://doi.org/10.35536/ljb.2024.v11.i2.a4

Sancha, C., Gimenez, C., & Sierra, V. (2016). Achieving a socially responsible supply chain through assessment and collaboration. *Journal of Cleaner Production, 112*, 1934–1947. https://doi.org/10.1016/j.jclepro.2015.04.137

Sánchez-Flores, R. B., Cruz-Sotelo, S. E., Ojeda-Benitez, S., & Ramírez-Barreto, M. E. (2020). Sustainable supply chain management—A literature review on emerging economies. *Sustainability, 12*(17), Article 6972. https://doi.org/10.3390/su12176972

Schramm, V. B., Cabral, L. P. B., & Schramm, F. (2020). Approaches for supporting sustainable supplier selection—A literature review. *Journal of Cleaner Production, 273*, Article 123089. https://doi.org/10.1016/j.jclepro.2020.123089

Seuring, S., & Müller, M. (2008). From a literature review to a conceptual framework for sustainable supply chain management. *Journal of Cleaner Production, 16*(15), 1699–1710. https://doi.org/10.1016/j.jclepro.2008.04.020

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Tong, L. Z., Wang, J., & Pu, Z. (2022). Sustainable supplier selection for SMEs based on an extended PROMETHEE II approach. *Journal of Cleaner Production, 330*, Article 129830. https://doi.org/10.1016/j.jclepro.2021.129830

Vellesalu, A., Salomonson, N., & Hjelmgren, D. (2026). Shaping SME purchasing for sustainability: A case study of supplier-buyer relationships and extended product lifetimes. *Journal of Business & Industrial Marketing, 41*(13), 282–298. https://doi.org/10.1108/JBIM-04-2025-0335

Zimmer, K., Fröhling, M., & Schultmann, F. (2016). Sustainable supplier management—A review of models supporting sustainable supplier selection, monitoring and development. *International Journal of Production Research, 54*(5), 1412–1442. https://doi.org/10.1080/00207543.2015.1079340

Zorzini, M., Hendry, L. C., Huq, F. A., & Stevenson, M. (2015). Socially responsible sourcing: Reviewing the literature and its use of theory. *International Journal of Operations & Production Management, 35*(1), 60–109. https://doi.org/10.1108/IJOPM-07-2013-0355