2022

Scale development and validation of Transaction Cost Economics typology for contracts: A systems thinking approach

Sedat Cevikparmak

Hasan Celik

Saban Adana

Hasan Uvet

Brian Sauser

See next page for additional authors

Authors

Sedat Cevikparmak, Hasan Celik, Saban Adana, Hasan Uvet, Brian Sauser, and David Nowicki

Scale development and validation of Transaction Cost Economics typology for contracts: A systems thinking approach

Sedat Cevikparmak [a,*], Hasan Celik [b], Saban Adana [c], Hasan Uvet [d], Brian Sauser [e], David Nowicki [e]

[a] Division of Business, DeSales University, 2755 Station Avenue, Center Valley, PA, 18034, USA
[b] School of Business, Robert Morris University, 6001 University Blvd, Moon Twp, PA, 15108, USA
[c] Boler College of Business, John Carroll University, 1 John Carroll Blvd, University Heights, OH, 44118, USA
[d] School of Business, Georgia Gwinnett College, 1000 University Center Ln, Lawrenceville, GA, 30043, USA
[e] G. Brint Ryan College of Business, University of North Texas, 1307 W Highland St, Denton, TX, 76201, USA

ARTICLE INFO

Keywords:
Transaction cost economics
TCE contract Typology
Systemigram
Classical contracts
Neo-classical contracts
Relational contracts

ABSTRACT

A variety of contract typologies that exist in the literature are helpful in the exploration of different approaches in contractual relations, but only when measured with the right instruments. Although Transaction Cost Economics (TCE) has a distinct, high-level, and abstract typology for contracts, it still lacks a measurement scale. In this paper, a measurement scale for the TCE contract typology (classical, neo-classical, and relational contracts) was developed and validated, using systems thinking approach and experimental design to contribute to the empirical tests of TCE within the contracting realm. First, the antecedents of contract selection within the TCE literature were analyzed using the systemigram technique to visualize and parse out complex relationships that lead to contract selection. The analysis of the TCE Systemigram helped the development of the scale and revealed the need to revisit the risk neutrality assumption embedded in TCE. Second, a measurement scale for the TCE contract typology (classical, neo-classical, and relational contracts) was developed adhering to the original texts of seminal papers and reviews from the TCE literature. Third, the 14-item measurement scale was validated using a series of three vignette-based experimental studies.

Key message: This research explores the antecedents of the TCE contract typology and develops a measurement scale for essential characteristics of classical, neo-classical, and relational contracts, as defined in TCE, using systems thinking approach and a novel vignette-based experimental design.

1. Introduction

Even though Transaction Cost Economics (TCE) theory has been applied to purchasing and supply management (PSM) in many contexts, its contract selection and management insights have largely been overlooked, partly due to the lack of an established measurement scale of TCE contract types. The main goal of this research is to address this gap in the literature by developing and validating a measurement scale for TCE contract types, which will enable further research on contract selection in the purchasing and supply management field through the lenses of TCE theory.

Contract selection is a key yet understudied phase of contracting within the purchasing and supply management field (Wynstra et al., 2019). The magnitude of annual revenues represented in contracts with suppliers for firms in utilities, aerospace, defense, and food manufacturing can be as high as 90% yet on average only 1% of annual revenue is spent across industries for contract development and vendor management (Belotserkovskiy et al., 2018). Suboptimal contract terms and lack of effective contract management might have an impact of as high as 9% of annual revenues (Belotserkovskiy et al., 2018), implying that there is an additional value that could be unlocked with a better understanding of contracting.

Make or buy decision is a central and recurring issue that confronts the managers in purchasing and supply management circles. Outsourcing is a powerful managerial choice to leverage a firm's internal and external resources, capabilities, and competencies (Mantel et al., 2006). Outsourcing would allow firms to free up resources that are much needed for innovative, more profitable products or their core

competencies. In this regard, make or buy decision takes on a strategic prominence that could help business thrive (Leenders et al., 2006). From a contracting perspective, both decisions (insourcing and outsourcing) entail some form of contract, which implies critical importance. Concentrating on transactions between and within firms, the TCE research produced valuable knowledge in the contracting field to both academia and practitioners.

TCE has drawn much attention in the past decades among a diverse group of scholars ranging from economics to organizational theory to even political science. Marketing scholars were the pioneers, especially in the empirical exploration of TCE, mainly since TCE takes economic exchange as the focal unit of analysis between firms. Purchasing and supply management discipline also extensively used TCE. In their assessment of the use of external grand theories in purchasing and supply management, Spina et al. (2016) reported that TCE was by far the most adopted framework (57 out of 102 occurrences of 12 grand theories they searched for). Wynstra et al. (2019), echoed this finding and established that TCE was the most frequently applied theory in PSM over the past two decades. This is no surprise since TCE endeavors to explain the make-or-buy dilemma and the ensuing buyer-supplier relationship, which is one of the topics covered most by PSM literature (Spina et al., 2016). Of all the PSM topics studied, contract management was found to be the most grand-theory-based topic (Spina et al., 2016). It is also interesting to note that although Giunipero et al. (2018), in their effort to identify the most applied theories in sourcing, have also confirmed TCE's prominence, they also found that the use of TCE in sourcing was in decline in recent years. This paper aims to open a new avenue of research in contract selection by offering a vignette-based measurement scale of TCE contract types. Treating a firm as the epicenter of governance of economic activity and focusing on transactions between and within firms, TCE has much to offer for contracting research in purchasing and supply management (Hoffmann et al., 2013). Many of the constructs and assumptions in TCE, such as governance structures, asset specificity, bounded rationality and opportunism have been heavily studied in prior research (Rindfleische, 1996; Schermann et al., 2016). However, contract types, transaction frequency, and risk neutrality received less attention among scholars (Saussier, 2000; Wynstra et al., 2019).

As a theoretical and methodological approach, systems thinking has also been rarely applied to TCE. The complex nature of relationships between TCE constructs calls for a systems understanding that enables scholars to visualize and parse out effects of various constructs on each other and the system as a whole, as firms strive to select the best governance structure/contract type to minimize transaction costs. Utilizing systemigram - a systemic diagramming technique - was deemed appropriate to reveal the essence of the original conceptual thinking using the structured seminal texts from the TCE literature.

TCE literature provides testable insights on contract selection. To mitigate the risks arising from the inherent nature of contracting, Williamson (1979) argues that some situations require protective safeguards ranging from simple realignment of incentives upfront (classical contracts) to creating a specialized governance structure for resolving disputes (neo-classical contracts), to introducing trading regularities which signal and insinuate continuity (relational contracts). Thus, drawing from the contract law literature, Williamson (1979) proposes a distinct, abstract, and high-level contract typology: classical, neo-classical, and relational contracts. The complex mix of conjectured relationships is built on this well-defined contract typology.

Extant literature uses a mix of typologies of contracts construed for a variety of disciplines or research programs. Various typologies could broaden the number and width of angles to be examined in contractual relations, but only if measured with the right instruments. Establishing the nomological validity of a typology and its measurement scale against others is of utmost importance for proper scholarly interpretation. The main purpose of this paper is to take a step forward in this direction with regards to TCE contract typology. Even though TCE is reported as one of the most utilized theories in contract management research (Giunipero et al., 2018; Spina et al., 2016) within the purchasing and supply management realm as well as within other related fields, it still lacks a measurement scale for TCE contract typology. In the few studies where TCE's approach to contracting was examined, scholars either used a different typology (Adler et al., 1998) or used contract completeness (Nyrhinen and Dahlberg, 2007) or interviews (Lacity and Willcocks, 1995) to operationalize the TCE contract typology. Adler et al. (1998) explored how TCE dimensions of asset specificity, uncertainty, and contract incompleteness could explain various contract types, which differ along with transactions in terms of content, performance incentives, and division of gains. However, they used a different contract typology developed by Brittelli et al. (1983): cost-plus fixed fee, firm-fixed-price, and incentive contracts. How well this typology could be approximated to the TCE contract typology is yet to be debated and left untouched in this research. The other two studies mentioned above (Nyrhinen and Dahlberg, 2007; Lacity and Willcocks, 1995) did use the TCE contract typology but did not develop a measurement scale to operationalize contract types of TCE. TCE contract typology is embedded in four different purchase scenarios as discussed in Williamson's (1979) paper and any attempt that falls short of addressing the purchase scenarios and TCE contract types together, differentiating contract types from the governance structures may not fill the void within the contracting research using TCE framework. On the other hand, a vignette-based measurement scale of contract types that address each of the purchasing scenarios, solely developed and validated for the TCE framework will enable scholars to put TCE propositions into more robust testing, eliminating possible confounding effects that might surface with the use of a mixed contract type measurement scale or insufficient operationalization.

This paper aims to contribute to the study of purchasing and supply management in several ways. First, a comprehensive TCE Systemigram was created to visually illustrate the intricate interrelationships of the TCE framework. Using the systemigram technique, a systems thinking approach was employed to the main premises of TCE to envisage and deconstruct complex relationships described in TCE with a focus on improving the understanding of contracting. Second, a measurement scale was developed and validated for classical, neo-classical, and relational contracts as defined in TCE. A measurement scale that captures the key characteristics of TCE contract typology might become a valuable instrument. The use of this scale for TCE research might improve the validity of future studies in the contracting context. In all, this research (1) produces a Systemigram of TCE constructs as they relate to contract selection, thereby enhances and visualizes the understanding of TCE theory; (2) develops, validates, and verifies a measurement scale for TCE contract typology; enabling future research, in which risk neutrality and other behavioral assumptions of TCE for contract selection will be put to more robust empirical testing.

This paper is organized into five sections. Section 2 discusses relevant literature and provides a theoretical background to the scale development and validation in Section 3. Section 4 discusses the results of the analysis of the TCE Systemigram and scale development process. The paper ends with concluding remarks, implications, and contributions in Section 5.

2. Literature review on contracts in Transaction Cost Economics

An in-depth grasp of the complex interrelationships of concepts and constructs in the TCE framework is necessary to understand the theoretical background underneath the TCE contract typology. TCE's analytic framework relies on three main assumptions about human behavior (i.e., bounded rationality, opportunism, and risk neutrality) and three key dimensions or characteristics of transactions, asset specificity, uncertainty, and transaction frequency. (Geyskens et al., 2006; Rindfleisch and Heide, 1997; Williamson, 1979). Transaction cost economics suggests that firms' choice of hierarchies (in-house production/vertical

integration), hybrid governance, or markets as an economic governance structure depends on the costs and difficulties associated with market transactions (Borys and Jemison, 1989).

The governance approach in TCE suggests that structure decisions are influenced mainly by the urge to economize on transaction costs (Williamson, 2002). The basic questions revolve around make or buy dilemma. Each generic mode of governance is supported by a distinctive form of contract law. As investments become more specific to the buyer/seller relationship, anticipation is that cost-minimizing institutional choice will respond by moving from simple anonymous market contracting (classical contracts) to more complex long-term contractual arrangements with protective provisions (relational or neo-classical contracts) and ultimately to vertical integration (Williamson, 1985; Joskow, 1988).

The contract typology in TCE has its differences and resemblances to other extant contractual conceptualizations. At the grand theory level, there is a deeper philosophical distinction (in terms of contracting) between TCE and other grand theories such as the Agency Theory and the Resource-Based View (RBV) (Gulbrandsen et al., 2009). The primary motivation for contracting is incentive alignment in the Agency Theory, whereas the RBV emphasizes resource use and deployment (Gulati, 1995). In TCE, transactional relationships are at the center of contracting literature. Thus, the discussion of the characteristics of contracts varies depending on what theory is being used.

Williamson's (1979) contract typology laid out the groundwork for the development of a variety of contractual typologies in the literature (Lee et al., 2003). Later studies expanded from TCE typology and adapted a diverse set of means to derive alternative frameworks. Heide (1994), on the basis of TCE, resource dependence perspective, and relational contracting theory suggested a framework of market, unilateral/hierarchical non-market, and bilateral governance, where market governance is described as a discrete exchange and non-market governance as relational (Dwyer et al., 1987). Lusch and Brown (1996) based their work on Macneil (1980) social contract concept and proposed explicit/hard contracts and normative/soft contracts as the two major forms of contracts to govern the relationships. Another angle taken by some scholars was based solely on control structures (Dahlstrom and Nygaard, 1999), which was coined as formalization and interfirm cooperation. Houston and Johnson (2000), on the other hand, proposed the dichotomy of buyer-supplier contracts vs. joint ventures in their effort to examine interfirm relationship structures. All these frameworks or typologies have their differences from the TCE contract typology of classical, neo-classical, and relational contracts.

A classical contract is when a firm goes out to the market and proposes a fixed price to pay for some good or service, and receives bids from potential suppliers (Macneil, 1977; Williamson, 1979; Eckerd and Girth, 2017). This is an advantage for the buyer since the risk of cost overrun is fully transferred to the supplier. Supplier bears all risks and manages all activities so, the buyer has less involvement, leading to minimal transaction costs (Williamson, 1979). The downside with the classical contracts is that as uncertainty and specificity of transaction rise, it becomes harder to draft a complete contract, which may lead to unsatisfactory delivery at the end of the contract duration (Macneil, 1977; Williamson, 1979).

A neo-classical contract is dependent upon trilateral governance, where "third-party assistance" is used for resolving disputes or evaluating performance. (Macneil, 1977; Williamson, 1979). This contract type assumes the terms and conditions are more uncertain. Most risks are taken by the buyer. It is often perceived by the buyer that the supplier might behave opportunistically, which requires costly monitoring over the progress and the quality of the supplier's performance (Williamson, 1979). The upside of a neo-classical contract is that the chances of satisfactory delivery are much higher, frictions between buyer and supplier are much easily resolved due to third-party assistance, albeit for a higher cost.

A relational contract is based upon a relationship of trust between the parties. The explicit terms of the contract (if there are any) are just a framework (Macneil, 1977). But there are implicit terms and understandings which determine the behavior of the parties. It requires extensive buyer involvement, thus increasing overall transaction costs along with overall costs due to incentivized cost mechanism (Williamson, 1979). However, the chances of satisfactory delivery are much higher, and in the long run, the total cost might come down due to continuity (Macneil, 1977; Williamson, 1979).

In TCE, parties of a contract share two central behavioral assumptions. First is that their ability to receive, store, retrieve, and process information is strictly limited due to bounded rationality (Williamson, 1979). Second, they are both motivated for "self-interest seeking with guile", i.e., opportunism which levels the rules of the game for both parties, in which calculated efforts to misinform, mask, and confuse are thus self-proclaimed. (Williamson, 1979, 1985). Williamson argues that problems of a contract are immensely complicated by bounded rationality and opportunism and asserts that these two behavioral assumptions are givens in any contracting context, whereas asset specificity and transaction frequency are the two variables that might be present at various levels depending on the type of transaction at hand (Williamson, 1979). Overall, Williamson argues that the ramifications of differences in bounded rationality, opportunism, asset specificity, and transaction frequency coupled with uncertainties are the key issues to be tackled in contracting (Williamson, 1985).

Williamson explored the foregoing relations among governance structures and contract types in his 1979 article as shown in Table 1. Treating governance structures and contract types as a function of transaction frequency and asset specificity, Williamson (1979) made distinct predictions on how firms would behave when faced with different combinations of these two variables. For various purchasing scenarios, the type of contracting predicted was different due to variation in levels of transaction frequency and asset specificity of that purchase. Standard component/material, standard equipment, customized component/material, and customized equipment-plant are the main four categories of purchases that correspond to a certain type of contract within TCE contract typology; classical for the first two purchases, relational for the third, and neo-classical for the fourth.

As discussed before, risk neutrality is also one of the main assumptions for these predictions. An in-depth review and analysis of the antecedents to the predictions in Table 1 is necessary to reveal the underlying mechanisms and create a deeper understanding of Williamson's (1979) line of thought, before delving into the effort to develop a scale for TCE contract typology. The systemigram technique might be helpful in this effort.

The use of diagrams to epitomize a system's form, fit, function,

Table 1
Governance Structures and Contract Types (Merged from two tables in Williamson, 1979).

		Asset Specificity		
		Nonspecific	Mixed	Transaction-specific
Transaction Frequency	Occasional	Market Governance	Hybrid Governance	
		Classical contracting	Neo-classical contracting	
		Purchasing standard equipment	Purchasing customized equipment-constructing a plant	
	Recurrent	Market Governance	Bilateral governance	Unified governance
		Classical contracting	Relational contracting	Relational contracting
		Purchasing standard material/ component	Purchasing customized material/ component	Site-specific transaction across stages

component, or environment is prevalent in systems engineering to help create a shared vision about the system (Cloutier et al., 2015). Due to the inherent limitations of these diagrams in their ability to capture multiple and divergent perspectives, systemigram, as a novel systemic diagramming technique has been proposed to alleviate shortcomings (Cloutier et al., 2015). Systemigrams are deployed for studying the remarkable elements/factors within a system of interest (Cloutier et al., 2015). Indeed, a systemigram is a network. It has nodes and links, a beginning, a flow, and an end. Noun phrases (nodes) specify stakeholders (people, organizations, conditions, constructs), whereas verb phrases (links) describe the complex relationships between these nodes (Cloutier et al., 2015). The first node on the top left is the focal stakeholder and the bottom right node is the objective of the system. Sub-systems and their vertical and horizontal inter-relationships are also revealed by a systemigram.

To this end, in this paper, the systemigram technique was used to parse out complex relationships within TCE constructs and assumptions towards contracting. The systemigram will also help visualize the inherent risks to the parties and their salience in contract selection. In this regard, a TCE Systemigram was created from the structured text of seminal papers in the TCE literature by adhering to Williamson's seminal work (1975; 1979; 1985; 1991; 2002; 2008), and to the reviews and conceptual papers in TCE (David and Han, 2004; Foss and Jensen, 2019; Rindfleisch and Heide, 1997; Schermann et al., 2016). By characterizing the interrelationships and structuring them into a graphical presentation, the TCE Systemigram captures the essence of the original conceptual thinking.

The first step in creating the TCE systemigram was to capture the main objective of TCE in one sentence, which is called the mainstay in the systemigram technique (the mainstay of the TCE Systemigram is represented in bold black arrows in Fig. 1.): *TCE posits that firms faced by the threat of opportunism operate in a risk-neutral business environment and are restrained with bounded rationality as they strive to decide on the most suitable governance structures and the ensuing contract types to minimize transaction costs* (Williamson, 1975).

The next steps are to identify how the constructs and assumptions in the mainstay are described in the structured text and create "scenes" for each one of them also depicting any inter-relationships. Fig. 2 depicts the opportunism construct within the TCE Systemigram. The "mixed" and "idiosyncratic (transaction-specific)" categories in Table 1 for asset specificity was merged in the Systemigram because the focus of this paper is on contracts not on governance, and the predictions of Williamson (1979) concerning contracts do not vary for "mixed" and "idiosyncratic" realms. Therefore, asset specificity is treated in two levels: nonspecific (low) and transaction-specific (high). In this regard, as shown in Fig. 2: *Opportunism surfaces when asset specificity is high and creates a safeguarding problem, which in the end increases total transaction costs. At the same time, high transaction frequency increases opportunism.*

Fig. 3 shows how the bounded rationality construct works within the TCE framework: *Bounded rationality coupled with environmental and behavioral uncertainty creates adaptation and performance evaluation problems respectively, causing the firm to incur renegotiation and monitoring costs, which as a result increases the total cost of a transaction.*

Fig. 4 describes contract selection within the TCE framework. Transaction frequency is defined as recurrent (high) or occasional (low), describing the rate of repetition of a certain transaction. In TCE, governance structures and contract types are dependent on various levels of transaction frequency and asset specificity (Rindfleisch and Heide, 1997; Williamson, 1979).

Various combinations of transaction frequency and asset specificity and the ensuing contract types, depicted in Table 1, can be followed in Fig. 4 from right to left: *(1) when transaction frequency is occasional and asset specificity is high, hybrid governance and the neo-classical contract might be chosen to minimize transaction costs (1*st *line from the right); (2) when transaction frequency is recurrent and asset specificity is high, hybrid governance and the relational contract might be chosen to minimize transaction costs (2*nd *line from the right); (3) when transaction frequency is occasional and asset specificity is low, market governance and the classical contract might be chosen to minimize transaction costs (3*rd *line from the right); (4) when transaction frequency is recurrent and asset specificity is low, market governance and the classical contract might be chosen to minimize transaction costs (4*th *line from the right). (5) When transaction frequency is recurrent and asset specificity is very high, hierarchical governance (vertical integration) might be chosen to minimize transaction costs (5*th *line from the right)*;

The final step of creating the TCE Systemigram (Fig. 5) is combining

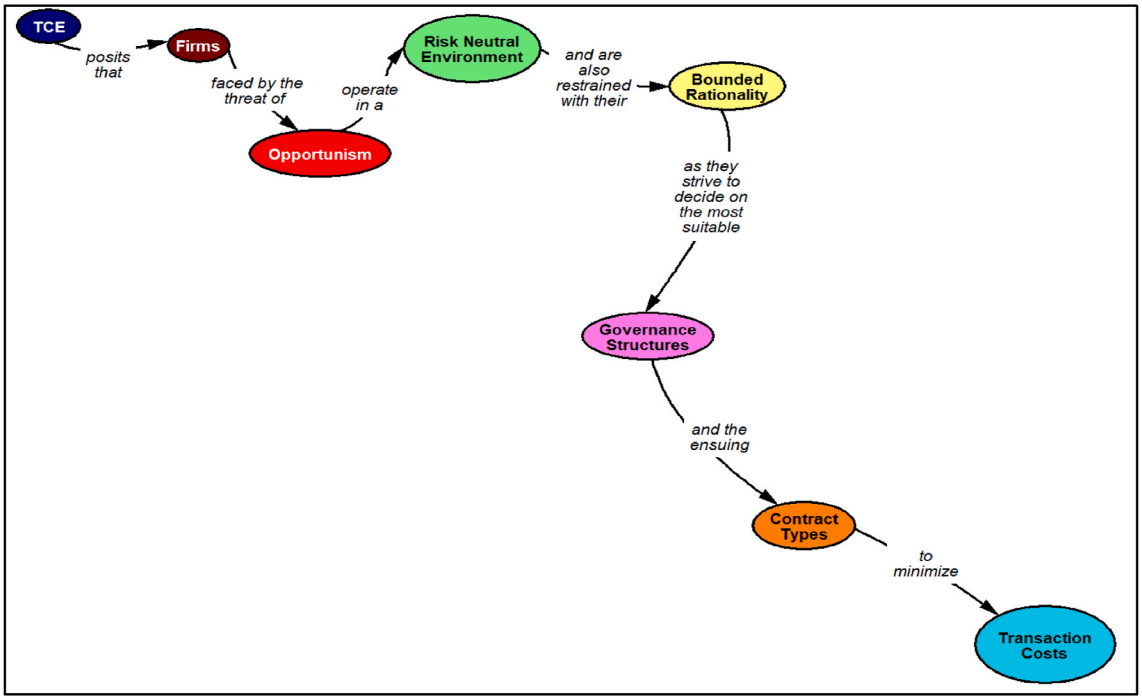

Fig. 1. TCE systemigram - mainstay.

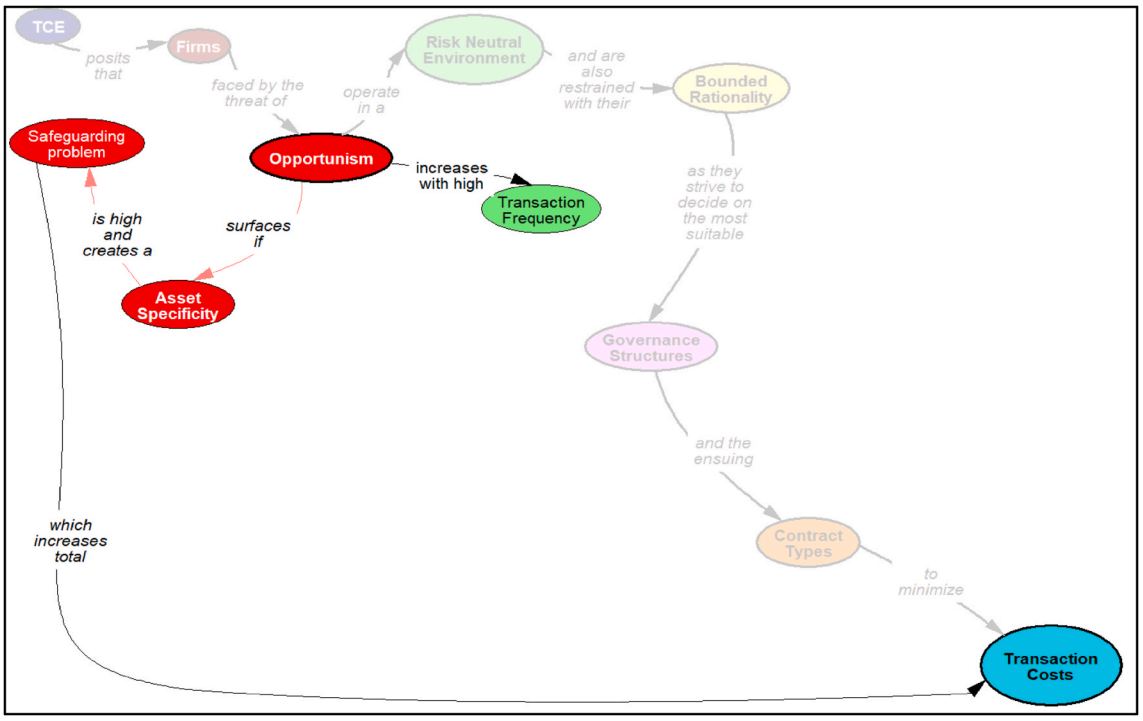

Fig. 2. TCE systemigram - opportunism.

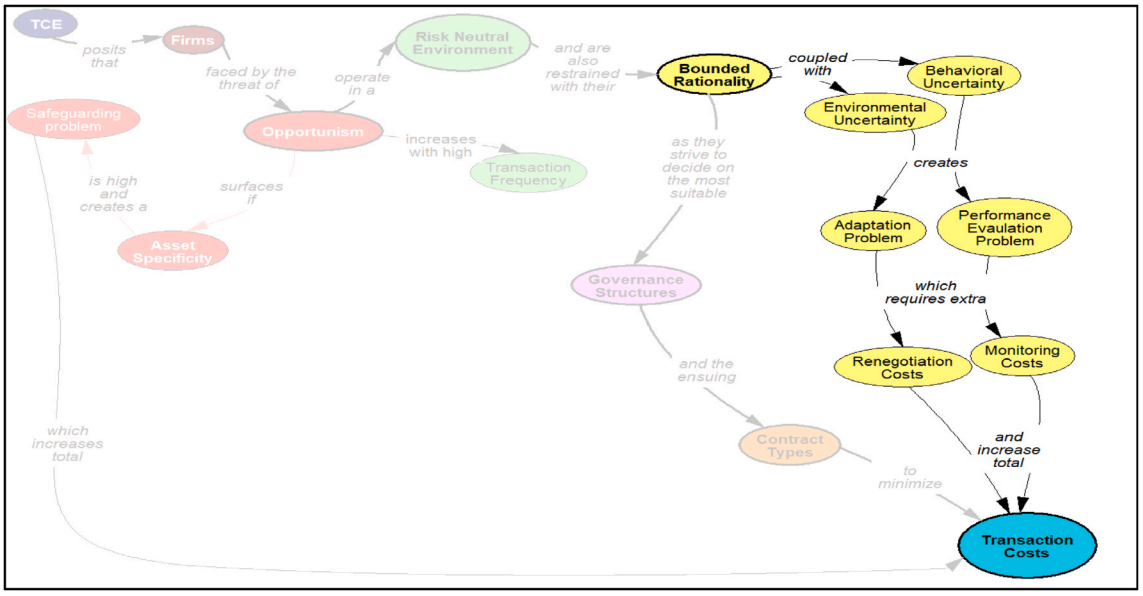

Fig. 3. TCE systemigram – bounded rationality.

all the previously created "scenes" that show all the complex interrelationships among TCE constructs as they pertain to the minimization of transaction costs. As observed in the Systemigram, bounded rationality, uncertainties (environmental and behavioral), and opportunism are ever-present, external factors in the overall context of any transaction. They all need to be considered and have an indirect effect on contract selection according to TCE theoretical framework. However, transaction frequency and asset specificity are the two variables that are inherent in any transaction and play a more crucial role in contract selection. In other words, the interaction between various levels of transaction frequency and asset specificity together is more of an internal factor and serves as the direct antecedents to contracting decisions. Applying systems thinking approach to TCE contracting framework using the systemigram technique was instrumental to visualize and illustrate these subtle nuances in the over-convoluted TCE theoretical framework.

Another outcome of the analysis of the TCE Systemigram was the highlighting of the presence of contract types as substantive variables distinct from the governance structures. For decades, exploration of contractual relations in TCE literature has paused at the governance structure level and either assumed that certain governance structures automatically lead to certain contractual paradigms as envisaged by Williamson (1979) or overlooked the contractual ramifications of governance structures totally (Foss and Jensen, 2019; Rindfleisch and

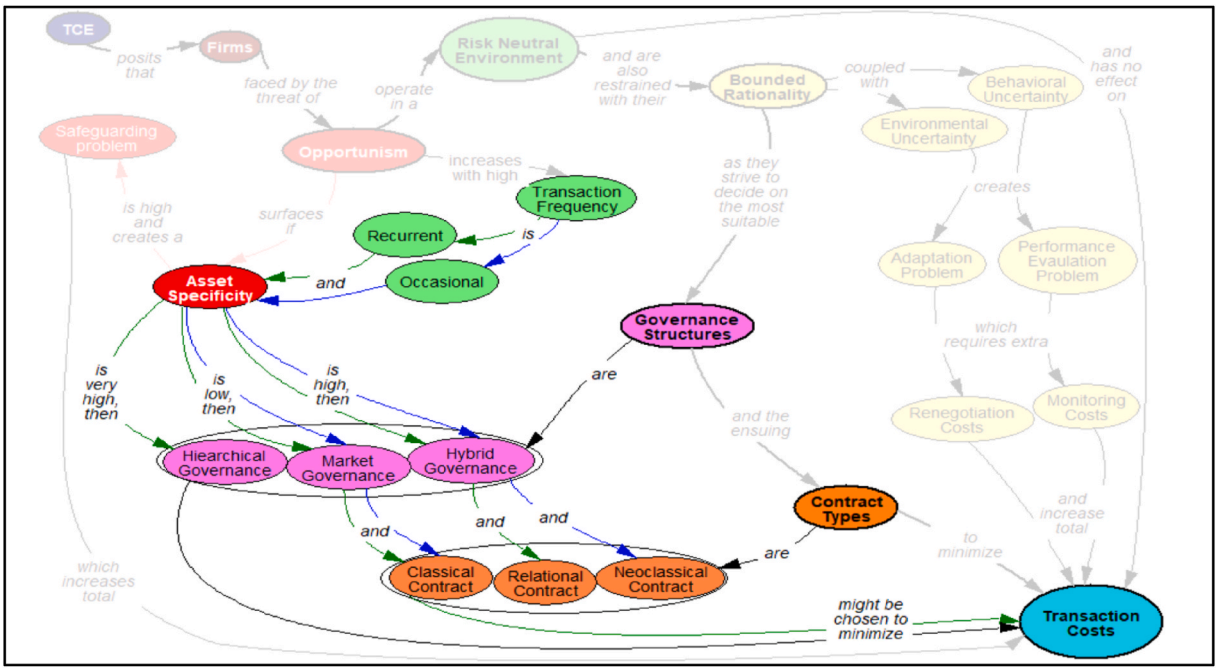

Fig. 4. TCE systemigram – contract selection.

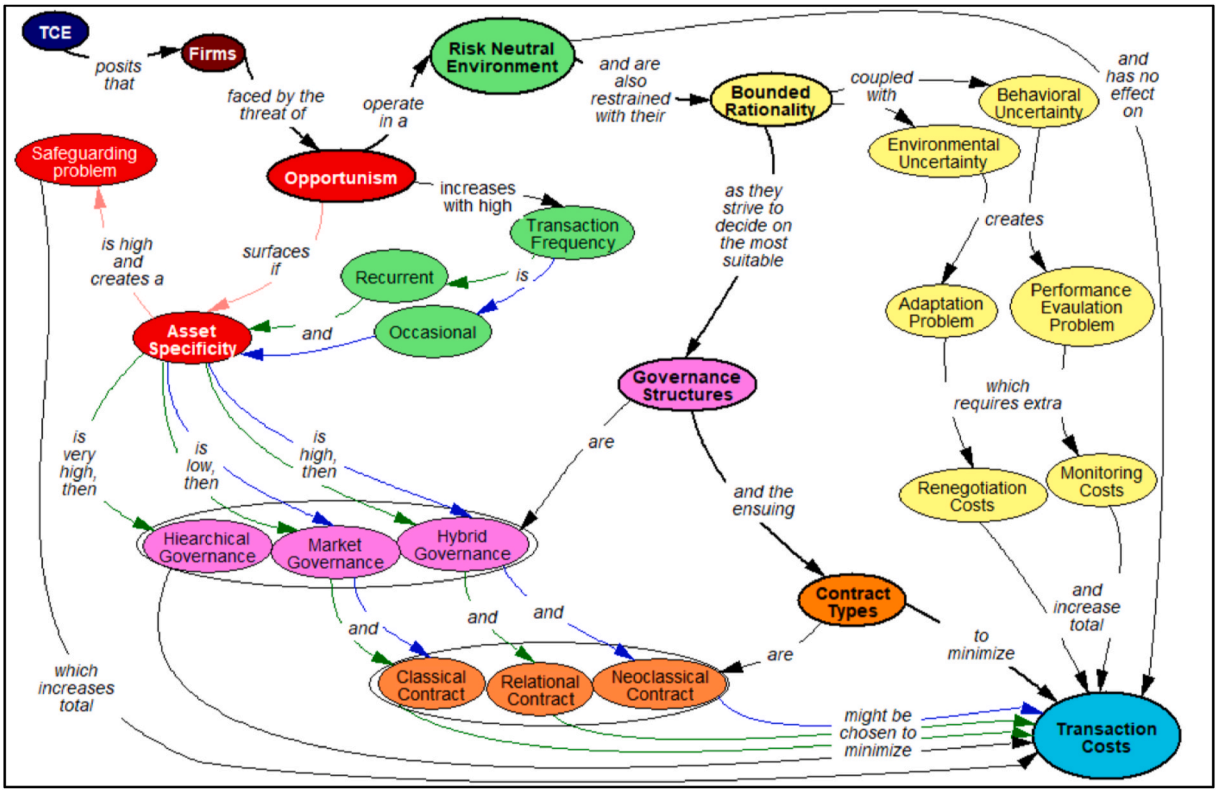

Fig. 5. TCE systemigram.

Heide, 1997; Schermann et al., 2016). On the contrary, one can argue that contractual arrangements are at the forefront of inter-firm relations and should carry no less weight than governance structures in academic research. Yet, as Adler et al. (1998) aptly argued, without the empirical verification of the predicted relationship between governance structures and contract types, the claims that TCE predictions for governance would also hold for contracts are merely anecdotal than factual. Furthermore, lack of empirical evidence to substantiate the existence of contract types as a continuum of governance structures was one of the issues that sparked criticism of TCE in the late 1990s (Moran and Ghoshal, 1996; Ghoshal and Moran 1996).

In the few studies that explored contracting within the TCE framework, the well-defined TCE contract typology received only insubstantial attention. In some cases, scholars abstained from using TCE

terminology, and used a different typology, as was the case with Adler et al. (1998) who preferred to use firm-fixed-price, cost-plus-fixed-fee, and incentive contracts as contract types in their research to test the validity of TCE dimensions on contract selection. In other cases, where scholars used TCE contract typology, TCE contract types were either operationalized by only as levels of contract completeness (shortening, lengthening the description of contract type in the questionnaire) (Nyrhinen, and Dahlberg, 2007) or by categorization of the contract types as a result of the coding of interviews conducted with managers (Lacity, and Willcocks, 1995). This drawback can be attributed to the lack of a distinct measurement scale for TCE contract typology. The robustness of future research on the TCE contracting paradigm may be immensely enhanced by the development of a measurement scale for the essential characteristics of TCE contract types.

3. Methodology of scale development and validation

A vignette-based experimental methodology was selected for scale validation for several reasons. First, the purpose of this research is to explore the distinct characteristics of classical, neo-classical, and relational contracts that separate these contract types from each other from the perspective of TCE. A vignette-based experimental design is well-matched to examine the subtle differences in human decision-making because it allows researchers to observe the direct impact of a change in a factor on a dependent variable (Eckerd, 2016). Second, the problem of retrospective biases, memory lapses, and post-rationalization when asked to recall past experience are almost non-existent in vignette-based experiments (Finch, 1987; Grewal et al., 2008; Lu et al., 2019). Third, vignette-based experiments have been demonstrated as a viable methodology in various purchasing and supply chain management contexts ((Davis-Sramek et al., 2018; Kaufmann et al., 2018; Lu et al., 2019; Polyviou et al., 2018). Fourth, as reported by Spina et al. (2016), experimental design is a rarely applied methodology in PSM research, inhibiting the study of behavioral influences in much detail. Therefore, a vignette-based experimental design for scale validation was deemed appropriate.

Established scale building recommendations and past scale-development studies and reviews were followed to develop a scale for measuring contract types within the TCE contract typology (e.g., Ambulkar et al., 2015; Bottger et al., 2017; Churchill, 1979; Gerbing and Anderson, 1988; Hensley, 1999; Shah and Ward, 2007; Baxter, R., 2009; Menor and Roth, 2007; Roth et al., 2008). The initial item generation was followed by three vignette-based experimental studies (Eckerd, 2016), including scale purification and initial validation, an examination of nomological validity, an exploration of the convergent and discriminant validity, and tests for the experimental and predictive validity. A detailed roadmap of scale development methods and data collection is shown in Fig. 6 (Froehle and Roth, 2004; Roth et al., 2008).

In line with Baxter (2009), the first step of the scale development was devoted to considerations on using formative or reflective specifications. The use of reflective scales means that scale items are indicators of the measured construct, and thus they correlate moderately strongly. On the other hand, in the formative scales, the indicators are independent "causes" of the measured construct with little correlation, and that their collective presence is necessary to adequately specify the measured construct (Baxter, 2009). As debated in the literature, conceptualization is the key to choose from formative or reflective scales, and that both could be appropriate in line with the research questions. In this paper, the use of reflective scale development was deemed appropriate for several reasons. First, various characteristics that are attributed to classical, neo-classical, and relational contracts within the TCE literature are not independent causes that form the different contract types. These characteristics are described as indicators of the TCE contract typology (Williamson, 1979). Second, contract types are discussed within specific purchase scenarios, the result of which is those specific contract characteristics such as the ability to fix the price, the need for arbitration, or the need for a high level of relationship with the supplier are not caused by the contract selection but they are more so reflected better within certain contract schemes.

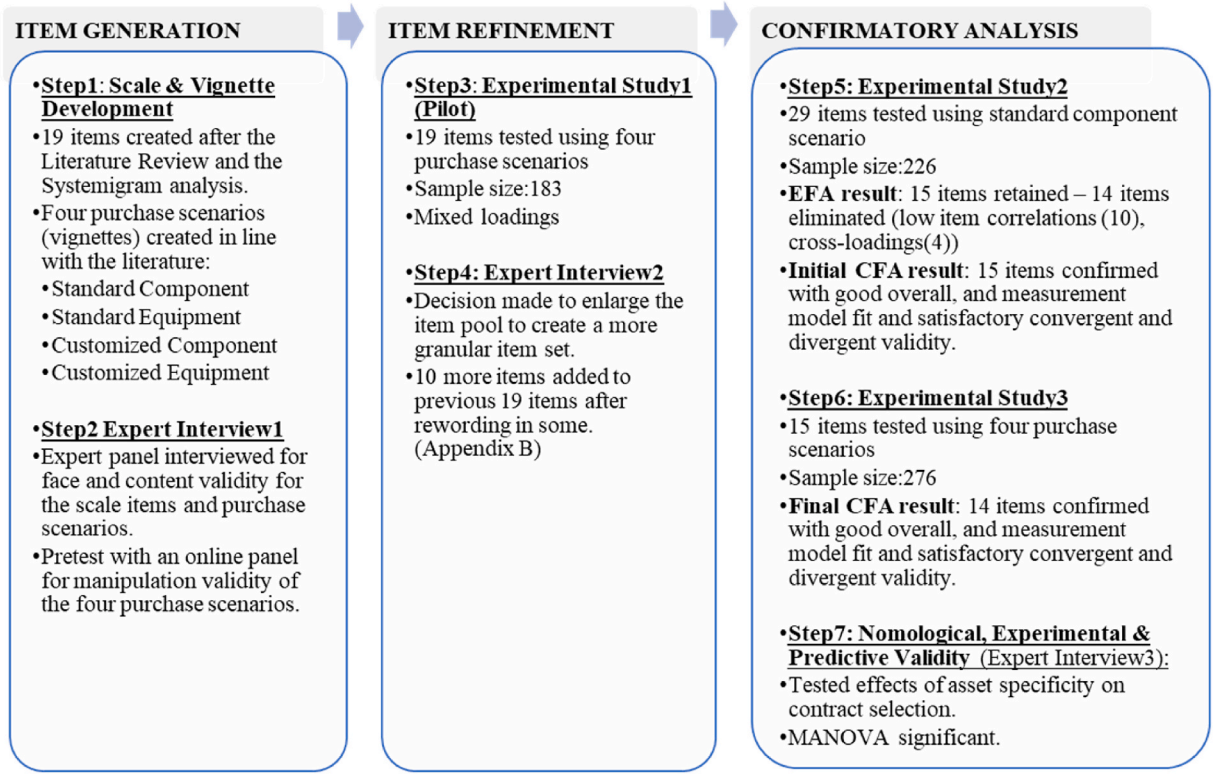

Fig. 6. Scale & vignette development and validation process.

3.1. Item generation

As a first step, a set of items encompassing general aspects of TCE contract typology was developed. Initially, 19 items were generated for classical, neo-classical, and relational contracts from a thorough examination of the TCE literature (e.g. Joskow, 1988; Macneil, 1974; 1977; Rindfleisch and Heide, 1997; Williamson, 1979, 1985, 2002, 2003, 2010). The seminal work that defines TCE contract typology is Williamson's 1979 article named "*Transaction-Cost Economics: The Governance of Contractual Relations*". The scale items were mainly generated adhering to the above article. Williamson's later articles and books and other seminal work from the TCE literature cited above were called upon to increase the face validity and capture the progression of the TCE contract typology over time. The items generated (a detailed explanation of TCE contract types and references are discussed in detail within the Literature Review section of this paper) captured general characteristics of classical, neo-classical, and relational contracts. A panel of eight experts (professors and Ph.D. students with past purchasing and procurement experience from an R1 public university in the southwest USA) evaluated the initial pool of generated items for face and content validity to assess how well the measurement items are reflective of the constructs of interest i.e TCE contract types: classical, neo-classical and relational contracts (Anderson and Gerbing, 1991; Froehle and Roth, 2004). Items were retained following the expert review.

3.2. Vignette design and validation/study 1 (pilot

The literature review has not revealed any studies that used a vignette, or a scenario to measure contract type preferences. Thus, the vignette-based experimental instrument was created after reviewing various scenario-based studies on topics ranging from make or buy decision (Mantel et al., 2006), resilience (Ambulkar et al., 2015), and supply chain sustainability (Davis-Sramek et al., 2018). The novel vignettes in this research were created by adhering to Williamson's (1979) purchase scenarios: standard component/material, standard equipment, customized component/material, and customized equipment. To ensure that vignettes were realistic, believable, and effective, best practices were followed as delineated in Bachrach and Bendoly (2011) and Rungtusanatham et al. (2011). The four vignettes differed from each other in terms of the frequency and specificity of the transaction. The participants were requested to choose the characteristics of the contract that they see fit for each purchase scenario. Following recommendations by Rungtusanatham et al. (2011), in the predesign stage, the language and description were aligned with prior research in behavioral decision-making literature and TCE literature. Additionally, in all vignettes, the text was common and invariant except manipulations and avoided cues as to the expectations of the experiment (Martin, 2004).

In line with Williamson (1979), four purchase scenarios were created (see Appendix A) such that asset specificity and transaction frequency were manipulated using a complete block design, where all other factors were held fixed. The participants were instructed to select the best fitting contract characteristics as the acting purchase/procurement manager of a global electronics company after they read the scenario randomly assigned to them. The two independent variables had two levels; low or high. This was manipulated by the choice of words that described the purchase scenario (*standard vs. customized for asset specificity, and component vs. equipment for transaction frequency*). Second, the same manipulation was repeated by stating that the item to be purchased was *specific vs. not specific* and the *frequency was low vs. high*. Further, four separate component/equipment names were used to help the participants visualize the purchase scenarios more realistically. The two independent variables were manipulated in the description of each scenario as follows: "… which; (1) *is/is not* easily available in the market at an equilibrium price, (2) *has a standard/requires a customized* design that fits your needs, (3) will be purchased *frequently/only occasionally*." (See Appendix A).

Following the pretest with an online panel, the experimental procedure was refined to ensure manipulation validity (Rungtusanatham et al., 2011).

3.3. Item refinement/study 1 (pilot)

A pilot study using an online panel (MTurk) of practitioners was conducted (N = 183) using four different purchase scenarios. As a first step, participants were instructed to respond to two identifier questions about their experience in contracting and job expertise. Only those with experience and relevant expertise were allowed to continue. A preliminary exploratory factor analysis was conducted for responses from each scenario in SPSS, using principal component analysis with varimax rotation. The limited number of measure items remained (2–3) for each contract type due to mixed factor loadings lead to an assessment with the expert panel and resulted in the decision to broaden the initial item pool. The new item pool included not only the most-pronounced characteristics of TCE contract types but also the second-order constructs of contract types such as price mechanisms, governance types (levels of involvement, need for arbitration and terms and conditions), and relational attitudes (trust, partnership, past experience, and goal alignment) (Williamson, 1979). The purpose of the broadening of the item pool was to increase the granularity of the measurement scale consistent with the results of the pilot study. The initial pool of 19 items was retained after some changes in wording, and a total of ten more items were generated in line with the literature. A detailed table of the item generation process and final scale is in Appendix B, in line with Shah and Ward (2007).

3.4. Reliability and validity analysis

3.4.1. Study 2

After the results of the pilot Study 1, a second study (Study 2) with 29 scale items was conducted with an online panel (MTurk) of practitioners using a standard component purchase scenario. First, participants were asked to respond to two identifier questions measuring their experience in contracting and job expertise. Only 341 respondents with contracting experience and suitable job expertise (logistics, supply chain, transportation, production, purchasing, operations management) were allowed to continue with the experiment. Out of the 341 responses submitted, 115 failed in both attention check questions (Kung et al., 2018). The remaining 226 responses were used for analysis. On average, the respondents reported 5.56 years of experience in contracting from a wide spectrum of industries ranging from automotive to health. Scale items for contract types were measured after each participant read the introduction and the standard component purchase scenario (See Appendix A and B). Participants were asked the following question: *"Given the above information, please indicate the essential characteristics of contract type you would recommend when you are purchasing a standard component"*. Responses to all scale items were measured on a 5 point Likert scale anchored between strongly disagree (1) and strongly agree (5).

Measure verification followed a multi-step process. The scale items measuring contract types were first subjected to exploratory factor analysis (EFA). The results of the EFA obtained from a principal component analysis with varimax rotation in SPSS are presented in Table 2.

First, the factorability of the measure items was examined. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .82 (greater than the recommended value of 0.60), and Bartlett's test of sphericity was significant ($\chi^2 = 1090.17$, $p < .001$) (Black and Porter, 1996).

Further, the commonalities were all above 0.3 indicating that each item shared some common variance with other items. Principal components analysis with varimax rotation was used because the main purpose was to identify and calculate composite scores for the factors underlying contract types. Initial Eigenvalues indicated that the first three factors explained 29%, 19%, and 8% of the variance respectively.

Table 2
EFA (study 2)-Rotated factor structure.

Scale Item	1	2	3
Contract Characteristics			
Classical Contract *(1)*			
Question: "When purchasing a standard component …."			
- Giving out price incentives to the supplier are unnecessary. (7C)	0.76		
- The final cost to us can be determined upfront. (2C)	0.71		
- Cost overrun is unacceptable. (3C)	0.69		
- We do not anticipate additional costs to the supplier other than those defined in the contract. (4C)	0.60		
- Setting a fixed price upfront would not jeopardize the deliverables. (5C)	0.60		
Neo-classical contract *(2)*			
Question: "When purchasing a standard component …."			
- The complexity of the purchase justifies third-party assistance and the extra cost that comes with it. (12N)		0.81	
- Third-party assistance (arbitration or legal help) is needed albeit the cost. (11N)		0.78	
- We anticipate disagreement with the supplier which dictates costly outside arbitration or legal help. (13N)		0.76	
- Uncertainties prevent setting a fixed price upfront. (14N)		0.75	
- Maximal friction between us and the supplier is likely. (19N)		0.72	
Relational Contract *(3)*			
Question: "When purchasing a standard component …."			
- We would like to think of the supplier as our partner. (21R)			0.76
- Goal alignment between us and the supplier is necessary. (24R)			0.72
- Implicit terms and understandings between us and the supplier would serve us well. (27R)			0.66
- Positive past experience with the supplier is a must. (20R)	0.47		0.60
- Cost issues should not supersede our relationship with the supplier. (8R)			0.47
Cronbach's Alpha	**0.71**	**0.85**	**0.75**

As cross-loading measure items (above 0.32) are removed, the initial six-factor solution was consolidated into a three-factor solution, which explained 56% of the variance (Tabachnick and Fidell, 2001). The final factor structure had all items (except 8R) loading over 0.60 or better, indicating solid factors (Osborne and Costello, 2009). One item (20R) had a cross-loading of 0.47 where primary loading was 0.60 but was retained since the item was theoretically justifiable. In all, the three-factor solution with a total of 15 items (5 each for three constructs) was deemed appropriate due to theoretical support and the 'leveling off' of the Eigenvalues on the scree plot after three factors.

The scale items for the classical, neo-classical, and relational contracts were next examined for internal consistency (using Alpha scores), and convergent and discriminant validity using inter-item correlation scores (see Tables 2 and 3). All Alpha scores were above 0.70 indicating acceptable levels of internal consistency (Nunnally, 1978). All inter-factor correlations within factors were higher than the correlations across factors, which satisfies the essential criteria for discriminant validity (Churchill, 1979).

The data were next subjected to the initial confirmatory factor analysis using a structural equation modeling procedure (AMOS, SPSS) using a variance-covariance matrix (Table 3) (Froehle and Roth, 2004). The fit indices for the three-factor structures (χ^2 = 151.71, df = 84, p-value = .001; NFI = 0.86, IFI = 0.94, CFI = 0.93, TLI = 0.92, and CMIN/DF = 1.80, RMSEA = 0.06) were within acceptable ranges (Anderson and Gerbing, 1991; Bagozzi and Yi, 1988; Menor and Roth, 2007). The sample size of 226 used in this study was sufficient for CFA. Construct validity for the scale items measuring constructs were assessed using AVE (average variance extracted), and construct reliability (CR). All the AVE estimates were close to 0.50, while all the CRs were above 0.78, which satisfies convergent validity criteria (Fornell and Larcker, 1981; Hair et al., 2006; Lam, 2012). Further, all multi-item measures' AVEs' square roots are higher than their correlation estimates with other factors and thus indicating discriminant validity. Overall, these findings indicate acceptable internal consistency, convergent and discriminant validity, and construct validity for the scale items used in this study (see Tables 2 and 3).

Common method bias is a prospective concern in cross-sectional studies where single informants are used (Podsakoff et al., 2003). Common latent factor method was used (Eichhorn, 2014) to address this issue. The common latent factor method indicated 32% variance for a single factor, which is less than the generally accepted benchmark of 50% (Eichhorn, 2014; Podsakoff and Organ, 1986).

Study 2 was conducted with one purchasing scenario to provide an initial test of the measurement items towards contract types. The standard component purchase scenario was deemed appropriate for Study 2 since it provides a striking contrast for contract feature selection by its simplicity and straightforwardness as a scenario, eliminating potential confounding effects. As seen in the initial CFA results, it is noteworthy that the three contract types were differentiated substantially. Yet, to increase the robustness of the initial CFA, a third experimental study (Study 3) was conducted using the full set of four purchasing scenarios.

3.4.2. Study 3

Following Study 2, Study 3 was conducted on MTurk using the four purchasing scenarios. To increase the validity and reliability of the results, the participants were first asked to answer an identifier question measuring their level of experience in purchasing (Kung et al., 2018). Out of 1535 participants, only 435 indicated some experience in purchasing and passed the screening question and thus were allowed to continue with Study 3. Of the 435 participants, 151 failed to respond correctly to the attention check questions (Kung et al., 2018), and therefore were removed from the data. 8 of the responses were identified as outliers using Mahalanobis Distance, Cook's distance, and boxplot analysis. The remaining 276 responses were used for analysis. The participants reported 6.87 years of experience in purchasing on average.

The data analysis followed a stepwise approach. First, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was checked and was found as 0.86 (greater than the recommended value of 0.60), and Bartlett's test of sphericity was significant ($\chi 2$ = 2276.90, p < .001) (Black and Porter, 1996). Further, the commonalities were all above 0.3 (except 8R = 0.241) indicating shared some common variance.

Next, the final confirmatory factor analysis (CFA) using a structural equation modeling procedure in AMOS-SPSS was conducted (Table 4) to check construct reliability and construct validity. The final measurement model was obtained after several runs of checks for lower loading factors, modification indices fixes, and standardized residual covariances. The final model contained 14 scale items as depicted in Table 4.

The overall model fit was analyzed for absolute model fit, incremental fit, and parsimonious fit, and results indicated that the measurement model was satisfactory (Anderson and Gerbing, 1991; Bagozzi and Yi, 1988; Paswan et al., 2015). For the absolute model fit, the final model χ^2 is 140.667, df = 73, (p = .000), indicating mediocre fit between the model implied and sample covariance matrices, which is expected due to the large sample size (276). On the other hand, RMSEA

Table 3
CFA 1 (study 2) - assessment of construct validity.

	Classical	*Neo-classical*	*Relational*	CR[b]	AVE
Classical	0.76[a]			0.87	.58
Neo-classical	−0.36	**0.67**[a]		0.81	.45
Relational	0.23	0.53	**0.65**[a]	0.78	.42

[a] Square root of average variances extracted are reported along a diagonal line. Correlations of factors' composite scores are below the diagonal line.
[b] CR = Construct Reliability.

Table 4
CFA 2 (study 3) - structural equation model estimates.

Parameters	Standardized Estimates	P Value
Contract Characteristics		
Classical Contract *(1)*		
Question: "When purchasing a …."		
- The final cost to us can be determined upfront. (2C)	0.80	***
- Setting a fixed price upfront would not jeopardize the deliverables. (5C)	0.71	***
- We do not anticipate additional costs to the supplier other than those defined in the contract. (4C)	0.53	***
- Giving out price incentives to the supplier are unnecessary. (7C)	0.50	***
Neo-classical Contract *(2)*		
Question: "When purchasing a …."		
- Third-party assistance (arbitration or legal help) is needed albeit the cost. (11N)	0.80	***
- The complexity of the purchase justifies third-party assistance and the extra cost that comes with it. (12N)	0.76	***
- We anticipate disagreement with the supplier which dictates costly outside arbitration or legal help. (13N)	0.72	***
- Maximal friction between us and the supplier is likely. (19N)	0.66	***
- Uncertainties prevent setting a fixed price upfront. (14N)	0.60	***
Relational Contract *(3)*		
Question: "When purchasing a …."		
- We would like to think of the supplier as our partner. (21R)	0.73	***
- Positive past experience with the supplier is a must. (20R)	0.71	***
- Goal alignment between us and the supplier is necessary. (24R)	0.66	***
- Implicit terms and understandings between us and the supplier would serve us well. (27R)	0.64	***
- Cost issues should not supersede our relationship with the supplier. (8R)	0.57	***

(0.058) and GFI (0.93) suggest a perfect model fit. Furthermore, for incremental fit, IFI (0.80) and as for the parsimonious fit, both CMIN/DF (1.927) and AGFI (0.90) indicate a satisfactory model fit.

To check for convergent and discriminant validity and construct reliability, calculations of AVE and CR values are given in Table 5 (Fornell and Larcker, 1981; Hair et al., 2006). All the standardized λs were close or above 0.6 and significant (See Table 4), all of the CRs were above 0.74, and all of the AVEs were above 0.42 (See Table 5). All of the inter construct correlations (φ) were less than the square roots of AVEs for the corresponding constructs. These results indicate a good model fit and satisfactory levels of internal consistency, convergent, discriminant, and construct validity (Hair et al., 2006; Paswan et al., 2015).

The experimental, predictive and nomological validity of the scale was established using the scale in the vignette-based experimental study (Study 3) to test the effects of asset specificity on contract selection

Table 5
CFA 2 (study 3) - assessment of construct validity.

	Classical	*Neo-classical*	*Relational*	CR[b]	AVE
Classical	0.65[a]			0.74	.42
Neo-classical	0.01	**0.71**[a]		0.84	.51
Relational	0.54	0.54	**0.66**[a]	0.80	.44

[a] Square root of average variances extracted are reported along a diagonal line. Correlations of factors' composite scores are below the diagonal line.
[b] CR = Construct Reliability.

(Narasimhan et al., 2001). A multiple analysis of variance (MANOVA) was performed to test for the overall significance of the difference in the means for dependent variables (contract types) at levels of asset specificity. The overall MANOVA result was statistically significant (Pillai's Trace = 0.11, F (3, 272) = 11.21, p = .0001), and the multivariate effect size was estimated as 0.11. The results were shared with members of the expert panel and received positive feedback, indicating experimental, predictive, and nomological validity. As a final step, data were analyzed for common method variance using the common latent factor method (χ^2 difference). The results indicated that there is 30% variance (Common Latent Factor method), which is below the generally accepted benchmark of 50% (Podsakoff and Organ, 1986; Eichhorn, 2014). Also, the χ^2 difference between constrained and unconstrained models when a common latent factor added was not statistically significant (χ^2 unconstrained = 94.2 df = 68, χ^2 constrained = 94.4 df = 69, p-value for the χ^2 difference = 0.65, df = 1).

4. Discussion of results

The purpose of this study was twofold. First, a systemigram was created of TCE theory to examine complex relationships between TCE constructs as to how firms make contract type decisions. The analysis of the TCE Systemigram revealed the fundamental relationships between various constructs and contract types, which was instrumental in the scale development process. The TCE Systemigram analysis underlined the presence of contract types as substantive factors in the TCE framework. Throughout the prior research on TCE, contract types were either treated as an automatic continuum of governance structures as predicted by Williamson (1979) or neglected (Foss and Jensen, 2019; Rindfleisch and Heide, 1997; Schermann et al., 2016). The possibility of contractual arrangements playing roles in more than one governance structure is a viable future research topic as an outcome of this research. A related topic is to identify the cognitive and behavioral mechanisms through which variance between governance structures and contract types can occur.

The TCE Systemigram was also contributory in identifying the nature of the relationship among some of the TCE constructs, by visually depicting that bounded rationality, uncertainties and opportunism are influential factors to be considered to minimize transaction costs, but only play an external role in contract selection decision process. In other words, these factors are ever-present in the milieu but get pronounced when they interact with levels of the other two variables: asset specificity and transaction frequency, which, in comparison, are the internal factors that play a more direct role in contract selection.

Another result of the analysis of the TCE Systemigram is the visual confirmation of the underlying but downplayed and underrepresented role of risk neutrality in this mix of relationships that lead to contract selection. The risk neutrality assumption of TCE is embraced as a simplifying assumption to devote more attention to transactions as the unit of analysis (Williamson, 1985). However, the findings of prospect theory and risk propensity research depict a different picture of risk neutrality (Martynov and Schepker, 2017; Sitkin and Pablo, 1992; Tversky and Kahneman, 1992). Thus, as a viable future research area, this paper implies that contract selection might be moderated by the risk propensities of managers as their perception of the risks that are inherent in an economic exchange will vary depending on their propensity.

Second, adhering to TCE literature and the Williamson (1979) contract typology a 14-item measurement scale was developed and validated. To this date, much of the TCE research has stopped at the governance structure level and has not empirically tested the predicted relationship between governance structures and contract types (Adler et al., 1998). The prior research conducted on contracts within the TCE framework either refrained from using the TCE contract typology, and used a different typology – firm-fixed-price, cost-plus-fixed-fee, and incentive contracts (Adler et al., 1998) or measured the TCE contract

types by only levels of contract completeness (shortening, lengthening the description of contract type in the questionnaire) (Nyrhinen, and Dahlberg, 2007) or classified them by conducting interviews with managers and coding their responses (Lacity, and Willcocks, 1995). The lack of a distinct measurement scale for TCE contract typology could be argued to diminish the robustness of prior research on contracting in the realm of TCE theory. Therefore, the theoretical contribution of the scale to the study of transactions and contracting through the TCE framework will be paramount.

Managerial implications of this research are also noteworthy. Enhancing the understanding of factors affecting contract selection, this paper not only points out the differences of each factor alone but also sheds light on how interactions among them are related to contract selection. Good contract management can save businesses much-needed resources, and it all starts with the selection of contract type that suits best the transaction at hand. Another outcome of this research is the highlighting of TCE contract typology as an abstract and high-level taxonomy, which suits better for strategic decision-making in contracting. The scale developed for TCE contract typology could be used as a starting point in the discussion of which contract would serve the firm better in their effort to minimize transaction costs and remain competitive.

Some limitations for this research should be noted. The experimental design maximizes the internal validity of the findings but comes with some shortcomings. Often, researchers tend to create extreme variances in scenarios to ensure that the respondents receive intended treatment manipulations (Davis-Sramek et al., 2018). This is because experimental manipulations tend to be notoriously weak, especially in social sciences research (Davis-Sramek et al., 2018). Despite realism checks, one cannot guarantee maximal realism. Using a multi-method approach could remedy this weakness in future studies.

As noted earlier, TCE contract typology is abstract and the measurement scale items devised to adhere to the theoretical underpinnings of the TCE framework may have been observed by the participating practitioners as somewhat hard to grasp, which is reflected in the satisfactory yet border-line AVE scores. This also very well be the case because relational contracts in comparison to classical or neo-classical contracts pose a difficulty in scale development. Relational attributes such as trust, partnership, goal alignment/commitment are powerful constructs that people crave in their lives and might have pulled respondents unequivocally towards them. This might have possibly created some bias in the results towards relational measure items. Similar effects surfaced in prior research. Suprapto et al. (2016) found that better relational attitudes, teamwork quality, and incentives fully mediates the relationship between various contract types (traditional vs relational) and project performance. However, even though classical contracts like fixed-price contracts have been shown to have a higher risk of project failure as specification increases (Jorgensen et al., 2017), nevertheless, they are still being used even for infrastructure projects (Suprapto et al., 2016). How do we account for what we observe in the real world and the research findings? Is there a bias towards relational attributes that is causing this discrepancy? What other factors could be at play? These are valid research questions that could be answered by conducting multiple studies with different samples, which might help assess the existence and/or the strength of the potential bias towards relational attributes in contract selection by the respondents.

This research was designed by adhering to the well-documented best practices in the literature for use of the Amazon MTurk (Matherly, 2019). Yet, limitations for data collection innate with the Amazon MTurk platform could have played some role in the results of this paper (Stritch et al., 2017). Hence, further studies using various samples might enhance the validity of the measurement scale developed in this paper.

5. Conclusion

This paper takes systems thinking approach to contracting through the lens of Transaction Cost Economics (TCE) theory to explore interrelationships among main antecedents of contracting in TCE: opportunism, bounded rationality, uncertainty, risk neutrality, transaction frequency, and asset specificity. A TCE Systemigram was created to visualize and examine the complex relationships between contract types and various levels of transaction frequency and asset specificity as well as the underlying assumptions of opportunism, bounded rationality, uncertainty, and risk neutrality. One of the outcomes of the literature review and the TCE Systemigram analysis was the categorization of the antecedents to contract selection as external and internal. The TCE Systemigram analysis revealed that asset specificity and transaction frequency play a central role as the main antecedents to contract selection in the TCE framework. Bounded rationality, uncertainty, opportunism, and risk neutrality are also influential. These factors are always present in any transaction context yet have a more external role than the first two antecedents (asset specificity and transaction frequency), which are more specific to transactions and therefore predicted to exert a more pronounced role on contract selection. The analysis of the TCE Systemigram has also shown that the fit between risk neutrality and other constructs needed further exploration. The literature review also revealed that risk neutrality along with risk perceptions/propensities of managers are understudied topics in contracting. To pave the way for future research in this area, a measurement scale for TCE contract typology was developed and validated to empirically test the Williamson (1979) predictions of how transaction frequency and asset specificity interact to shape decisions of contract types under risk neutrality assumption, using a vignette-based experimental design.

The main contribution of this paper to purchasing and supply management field as well as to other related fields is the introduction of a measurement scale for TCE contract typology of classical, neo-classical, and relational contracts, which was created and validated using a robust experimental vignette-based design adhering to the purchasing scenarios of Williamson's seminal 1979 article on contractual relationships. Considering the influence of TCE in many fields of research and the decades-long history of this grand theory's evolvement, the limited number of past research on the contracting premises of TCE separated from the governance structures is astonishing. TCE contract types are unique and abstract and are interwoven with carefully constructed purchasing scenarios that may have implications for modern contracting practices. The abstractness of the TCE contract typology makes a suitable tool for rigorous academic research towards the theoretical underpinnings of contracting and sourcing. With this novel measurement scale based on the four purchasing scenarios, researchers will have the opportunity to delve into new research topics.

Declaration of competing interest

Authors have no competing conflict of interest in the publication of the manuscript.

Acknowledgements

We would like to thank Dr. Timothy Hawkins (University of North Texas) for his comments and insights.

Appendix A. Scale Development – Purchase Scenarios (Vignettes)

Introduction:

Most of the spending in many industries is locked in contracts. There is an additional value that could be unlocked with wise contracting. Contract types substantially differ from each other. Thus, contract type selection during the pre-contract stage is a crucial factor for a business to thrive. The specificity and frequency of a transaction are two important factors in contract type selection. Specificity is the value of a certain asset outside the focal relationship between buyer and supplier. Transaction frequency is the number of occurrences of a certain exchange between the buyer and the supplier(s), ranging from low to high. Below, you see a table of possible combinations of two attributes and corresponding purchases.

		Transaction Frequency	
		Recurrent (High)	Occasional (Low)
Asset Specificity	**High**	Customized Component	Customized Equipment
	Low	Standard Component	Standard Equipment

Standard Material/Component Purchase:

- The item is standard, so we can expect plenty of qualified suppliers.
- The item is standard, so we can expect less involvement on our part and lower transaction costs (search, bargaining, renegotiation, performance evaluation costs, etc.)
- We have a good idea about market prices.
- We can expect to know upfront how much we will pay.
- We can expect no unforeseen costs to the supplier.
- The supplier could be any qualified firm that agrees to the price we offer.
- The level of relationship (trust, goal alignment, etc.) with the supplier is less relevant despite the frequency of purchase is high.
- Terms and conditions can be easily drafted.
- We can expect a smooth transaction, less need for arbitration or legal help.

Standard Equipment/Plant Purchase:

The above characteristics for a standard material/component purchase would mostly apply to the purchase of a standard equipment/plant; except:

- Most firms would need equipment purchase or plant/building construction much less frequently.
- Therefore, a long-term relationship with the supplier might not be necessary.

Customized Material/Component Purchase:

- The item is customized, so we can expect limited qualified suppliers.
- The item is customized, so we can expect more involvement on our part and higher transaction costs (search, bargaining, renegotiation, performance evaluation costs, etc.)
- We might not have a good idea about market prices.
- It might be hard to know upfront how much we will pay.
- We can expect some unforeseen costs to the supplier that would require reimbursement negotiations.
- Setup costs, design costs might be hard to quantify.
- Not all suppliers would be able to deliver. The supplier selection is more important.
- The level of relationship (trust, goal alignment, etc.) with the supplier might be more important because the frequency of purchase will be high.
- Terms and conditions might be hard to draft.
- We can expect a bumpy transaction, which requires more need for trust between parties.
- Giving incentives to the supplier might increase our costs but also might help the transaction go more smoothly.

Customized Equipment/Plant Purchase:

The above characteristics for a customized material/component purchase would mostly apply to the purchase of a customized equipment/plant; except:

- Most firms would need equipment purchase or plant/building construction much less frequently.
- Therefore, a long-term relationship with the supplier might not be necessary.
- We can expect a bumpy transaction, which requires more need for arbitration or legal help from a third party.
- Giving incentives to the supplier will increase our costs and might be unnecessary since the frequency of transactions is low.

Please carefully read the scenario on the next page and answer the questions accordingly.

Scenario 1- Standard Component Purchase:

You are the procurement/purchasing manager of one of the world's premier electronics manufacturing firm. So, it is your responsibility to evaluate and make recommendations on contract type selection for all kinds of procurement. You work at the corporate headquarters in New York and the production takes place in the manufacturing facility in Ohio. Your contract type selection should depend on the specificity and frequency of the transaction. The top management of the firm has a risk-neutral attitude and has put no restrictions on what contract type to choose, so it is your decision to make. The firm has decided to outsource **one standard component, i.e. a standard capacitor** which;

- **is easily available** in the market at an equilibrium price,
- has a **standard** design that fits your needs,
- will be purchased **frequently.**

Therefore, the standard component is **not specific** and the frequency of purchase is **high**.

Scenario 2 - Standard Equipment Purchase:

You are the procurement/purchasing manager of … ….
The firm has decided to outsource **one standard equipment, i.e. a standard voltmeter** which;

- **is easily available** in the market at an equilibrium price,
- has a **standard** design that fits your needs,
- will be purchased only **occasionally.**

Therefore, the standard equipment **is not specific** and the frequency of purchase is **low**.

Scenario 3 - Customized Component Purchase:

You are the procurement/purchasing manager of … ….
The firm has decided to outsource **one customized component, i.e. a customized switch** which;

- **is not easily available** in the market at an equilibrium price,
- requires a **customized** design that fits your needs,
- will be purchased **frequently.**

Therefore, the customized component **is specific** and the frequency of purchase is **high**.

Scenario 4 - Customized Equipment Purchase:

You are the procurement/purchasing manager of … ….
The firm has decided to outsource **one customized equipment, i.e. a customized semiconductor manufacturing machine** which;

- **is not easily available** in the market at an equilibrium price,
- requires a **customized** design that fits your needs,
- will be purchased **occasionally.**

Therefore, the customized equipment **is specific** and the frequency of purchase is **low.**

APPENDIX B. Item Generation and Final Scale

Given the above information, please indicate the essential characteristics of contract type you would recommend when you are purchasing a standard component/standard equipment/customized component/customized equipment ….” (Likert Scale: 1–5, from strongly disagree to strongly agree) Notes: X: Included X (r): Included after rewording.	Expert Review1	Study1 (Pilot) & Expert Review 2	Study2 (EFA & CFA 1	Study3 (CFA 2)
We can set a fixed price, so we know exactly our total cost.	X	X (r)	X	
The final cost to us can be determined upfront.	X	X	X	X
Cost overrun is unacceptable.	X	X (r)	X	
We do not anticipate additional costs to the supplier other than those defined in the contract.	X	X (r)	X	X
Setting a fixed price upfront would not jeopardize the deliverables.	X	X	X	X
Cost reimbursement to the supplier to make up for the extra costs is unacceptable.	X	X (r)	X	
Giving out price incentives to the supplier are unnecessary.	X	X (r)	X	X
Cost issues should not supersede our relationship with the supplier.		X	X	X
I would prefer a supplier I have worked before despite higher costs.		X	X	
Incentivizing the supplier monetarily to strengthen our relationship is appropriate although total cost might increase.		X	X	
Third-party assistance (arbitration or legal help) is needed albeit the cost.	X	X	X	X
Complexity of the purchase justifies third-party assistance and the extra cost that comes with it.		X	X	X
We anticipate disagreement with the supplier which dictates costly outside arbitration or legal help.		X	X	X
Uncertainties prevent setting a fixed price upfront.		X	X	X
Constant performance evaluation of the supplier is necessary, which will require extra money and resources.	X	X (r)	X	
We do not need much involvement after the contract has been signed.	X	X (r)	X	
Extensive monitoring of the supplier performance is necessary, which requires constant involvement on our part.	X	X (r)	X	

(*continued on next page*)

(*continued*)

Given the above information, please indicate the essential characteristics of contract type you would recommend when you are purchasing a standard component/standard equipment/customized component/customized equipment ….” (Likert Scale: 1–5, from strongly disagree to strongly agree) Notes: X: Included X (r): Included after rewording.	Expert Review1	Study1 (Pilot) & Expert Review 2	Study2 (EFA & CFA 1	Study3 (CFA 2)
Any qualified supplier would get the job done.	X	X (r)	X	
Maximal friction between us and the supplier is likely.	X	X	X	X
Positive past experience with the supplier is a must.		X	X	X
We would like to think of the supplier as our partner.		X	X	X
Mutual trust based on past business experience is an absolute necessity.	X	X (r)	X	
A long-term relationship with the supplier is a must.	X	X (r)	X	
Goal alignment between us and the supplier is necessary.	X	X	X	X
We can draft a detailed and formal terms and conditions.	X	X	X	
Written terms and conditions monitored by a third party is a must.		X	X	
Implicit terms and understandings between us and the supplier would serve us well.	X	X	X	X
Mutual understanding is more important than written terms and conditions.		X	X	
Adaptation of the initial agreement will be necessary as we go along.	X	X	X	

References

Adler, T.R., Scherer, R.F., Barton, S.L., Katerberg, R., 1998. An empirical test of transaction cost theory: validating contract typology. J. Appl. Manag. Stud. 7 (2), 185–200.

Ambulkar, S., Blackhurst, J., Grawe, S., 2015. Firm’s resilience to supply chain disruptions: scale development and empirical examination. J. Oper. Manag. 33, 111–122.

Anderson, J.C., Gerbing, D.W., 1991. Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. J. Appl. Psychol. 76 (5), 732.

Bachrach, D.G., Bendoly, E., 2011. Rigor in behavioral experiments: a basic primer for supply chain management researchers. J. Supply Chain Manag. 47 (3), 5–8.

Bagozzi, R.P., Yi, Y., 1988. On the evaluation of structural equation models. J. Acad. Market. Sci. 16 (1), 74–94.

Baxter, R., 2009. Reflective and formative metrics of relationship value: a commentary essay. J. Bus. Res. 62 (12), 1370–1377.

Belotserkovskiy, R., Sewak, J., Teodorian, A., Lietke, B., 2018. Contracting for Performance: Unlocking Additional Value. *Operations (by McKinsey and Company*. March 2018.

Black, S.A., Porter, L.J., 1996. Identification of the critical factors of TQM. Decis. Sci. J. 27 (1), 1–21.

Borys, B., Jemison, D.B., 1989. Hybrid arrangements as strategic alliances: theoretical issues in organizational combinations. Acad. Manag. Rev. 14 (2), 234–249.

Bottger, T., Rudolph, T., Evanschitzky, H., Pfrang, T., 2017. Customer inspiration: conceptualization, scale development, and validation. J. Market. 81 (6), 116–131.

Brittelli, J., Lynch, P., Emmelhainz, P., 1983. Principles of Contract Pricing. Wright-Patterson Air Force Base. Extension Course Institute, Ohio.

Churchill Jr., G.A., 1979. A paradigm for developing better measures of marketing constructs. J. Market. Res. 16 (1), 64–73.

Cloutier, R., Sauser, B., Bone, M., Taylor, A., 2015. Transitioning systems thinking to model-based systems engineering: systemigrams to SysML models. IEEE Transactions on Systems, Man, and Cybernetics: Systems 45 (4), 662–674.

Dahlstrom, R., Nygaard, A., 1999. An empirical investigation of ex post transaction costs in franchised distribution channels. J. Market. Res. 36 (May), 160–170.

David, R.J., Han, S.K., 2004. A systematic assessment of the empirical support for transaction cost economics. Strat. Manag. J. 25 (1), 39–58.

Davis-Sramek, B., Thomas, R.W., Fugate, B.S., 2018. Integrating behavioral decision theory and sustainable supply chain management: prioritizing economic, environmental, and social dimensions in carrier selection. J. Bus. Logist. 39 (2), 87–100.

Dwyer, F. Robert, Schurr, Paul H., Oh, Sejo, 1987. Developing buyer-seller relationships. J. Market. 51 (April), 11–27.

Eckerd, S., 2016. Experiments in purchasing and supply management research. J. Purch. Supply Manag. 22 (4), 258–261.

Eckerd, A., Girth, A.M., 2017. Designing the buyer–Supplier contract for risk management: assessing complexity and mission criticality. J. Supply Chain Manag. 53 (3), 60–75.

Eichhorn, B.R., 2014. Common Method Variance Techniques. Cleveland State University, Department of Operations and Supply Chain Management, Cleveland, OH. SAS Institute Inc.

Finch, J., 1987. The vignette technique in survey research. Sociology 21 (1), 105–114.

Fornell, C., Larcker, D.F., 1981. Structural Equation Models with Unobservable Variables and Measurement Error: *Algebra and Statistics*.

Foss, N.J., Jensen, H., 2019. Managerial meta-knowledge and adaptation: governance choice when firms don’t know their capabilities. Strat. Organ. 17 (2), 153–176.

Froehle, C., Roth, A., 2004. New measurement scales for evaluating perceptions of the technology-mediated customer service experience. J. Oper. Manag. 22 (1), 1–21.

Gerbing, D.W., Anderson, J.C., 1988. An updated paradigm for scale development incorporating unidimensionality and its assessment. J. Market. Res. 25 (2), 186–192.

Geyskens, I., Steenkamp, J.B.E., Kumar, N., 2006. Make, buy, or ally: a transaction cost theory meta-analysis. Acad. Manag. J. 49 (3), 519–543.

Ghoshal, S., Moran, P., 1996. Bad for practice: a critique of the transaction cost theory. Acad. Manag. Rev. 21 (1), 13–47.

Giunipero, L.C., Bittner, S., Shanks, I., Cho, M.H., 2018. Analyzing the sourcing literature: over two decades of research. J. Purch. Supply Manag., 100521

Grewal, D., Roggeveen, A.L., Tsiros, M., 2008. The effect of compensation on repurchase intentions in service recovery. J. Retailing 84 (4), 424–434.

Gulati, Ranjay, 1995. Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances. Acad. Manag. J. 38 (1), 85–112.

Gulbrandsen, B., Sandvik, K., Haugland, S.A., 2009. Antecedents of vertical integration: transaction cost economics and resource-based explanations. J. Purch. Supply Manag. 15 (2), 89–102.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L., 2006. Multivariate Data Analysis 6.

Heide, Jan B., 1994. Interorganizational governance in marketing channels. J. Market. 58 (January), 71–85.

Hensley, R.L., 1999. A review of operations management studies using scale development techniques. J. Oper. Manag. 17 (3), 343–358.

Hoffmann, P., Schiele, H., Krabbendam, K., 2013. Uncertainty, supply risk management and their impact on performance. J. Purch. Supply Manag. 19 (3), 199–211.

Houston, M.B., Johnson, S.A., 2000. Buyer–supplier contracts versus joint ventures: Determinants and consequences of transaction structure. J. Market. Res. 37 (1), 1–15.

Jorgensen, M., Mohagheghi, P., Grimstad, S., 2017. Direct and indirect connections between type of contract and software project outcome. Int. J. Proj. Manag. 35 (8), 1573–1586.

Joskow, P.L., 1988. Asset specificity and the structure of vertical relationships: empirical evidence. J. Law Econ. Organ. 4 (1), 95–117.

Kaufmann, L., Esslinger, J., Carter, C.R., 2018. Toward relationship resilience: managing buyer-induced breaches of psychological contracts during joint buyer–supplier projects. J. Supply Chain Manag. 54 (4), 62–85.

Kung, F.Y., Kwok, N., Brown, D.J., 2018. Are attention check questions a threat to scale validity? Appl. Psychol. 67 (2), 264–283.

Lacity, M.C., Willcocks, L.P., 1995. Interpreting information technology sourcing decisions from a transaction cost perspective: findings and critique. Account. Manag. Inf. Technol. 5 (3–4), 203–244.

Lam, L.W., 2012. Impact of competitiveness on salespeople’s commitment and performance. J. Bus. Res. 65 (9), 1328–1334.

Lee, K.S., Lim, G.H., Lim, W.S., 2003. Family business succession: Appropriation risk and choice of successor. Acad. Manag. Rev. 28 (4), 657–666.

Leenders, M.R., Johnson, P.F., Flynn, A.E., Fearon, H.E., 2006. Purchasing and Supply Chain Management, thirteenth ed. McGraw-Hill Irwin, Boston.

Lu, J., Kaufmann, L., Carter, C.R., 2019. Small talk, big impact–The influence of casual collegial advice on purchasing negotiations. J. Purch. Supply Manag. 25 (5), 100576.

Lusch, Robert F., Brown, James R., 1996. Interdependency, contracting, and relational behavior in marketing channels. J. Market. 60 (October), 19–38.

Macneil, I.R., 1974. A primer of contract planning. S. Cal. L. Rev. 48, 627.

Macneil, I.R., 1977. Contracts: adjustment of long-term economic relations under classical, neoclassical, and relational contract law. Nw. UL Rev. 72, 854.

Macneil, I.R., 1980. The New Social Contract. Yale University Press, New Haven.

Mantel, S.P., Tatikonda, M.V., Liao, Y., 2006. A behavioral study of supply manager decision-making: factors influencing make versus buy evaluation. J. Oper. Manag. 24 (6), 822–838.

Martin, D.W., 2004. Doing Psychology Experiments, 6th ed. Wadsworth, Belmont, CA.

Martynov, A., Schepker, D.J., 2017. Risk preferences and asset ownership: integrating prospect theory and transaction cost economics. Manag. Decis. Econ. 38 (2), 125–143.

Matherly, T., 2019. A panel for lemons? Positivity bias, reputation systems and data quality on MTurk. Eur. J. Market. 53 (2), 195–223.
Menor, L., Roth, A., 2007. New service development competence in retail banking: construct development and measurement validation. J. Oper. Manag. 25 (4), 825–846.
Moran, P., Ghoshal, S., 1996. Theories of economic organization: the case for realism and balance. Acad. Manag. Rev. 21 (1), 58–72.
Narasimhan, R., Das, A., 2001. The impact of purchasing integration and practices on manufacturing performance. J. Oper. Manag. 19 (5), 593–609.
Nunnaly, J.C., 1978. Psychometric Theory. McGraw-Hill.
Nyrhinen, M., Dahlberg, T., 2007, January. Is transaction cost economics theory able to explain contracts used for and success of firm-wide it-infrastructure outsourcing?. In: 2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07). IEEE, 240a-240a.
Osborne, J.W., Costello, A.B., 2009. Best practices in exploratory factor Analysis: four recommendations for getting the most from your analysis. Pan-Pacific Management Review 12 (2), 131–146, 2009.
Podsakoff, P.M., Organ, D.W., 1986. Self-reports in organizational research: problems and prospects. J. Manag. 12 (4), 531–544.
Paswan, A.K., Gai, L., Jeon, S., 2015. Alcohol and college students: Reasons, realization and intention to quit. J. Bus. Res. 68 (10), 2075–2083.
Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., Podsakoff, N.P., 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J. Appl. Psychol. 88 (5), 879.
Polyviou, M., Rungtusanatham, M.J., Reczek, R.W., Knemeyer, A.M., 2018. Supplier non-retention post disruption: what role does anger play? J. Oper. Manag. 61, 1–14.
Rindfleisch, A., Heide, J.B., 1997. Transaction cost analysis: past, present, and future applications. J. Market. 30–54.
Roth, A., Schroeder, R., Huang, X., Kristal, M., 2008. Handbook of Metrics for Research in Operations Management - Multi-Item Measurement Scales and Objective Items. Sage Publications, Thousand Oaks, p. 764.
Rungtusanatham, M., Wallin, C., Eckerd, S., 2011. The Vignette in a scenario-based role-playing experiment. J. Supply Chain Manag. 47 (3), 9–16.
Saussier, S., 2000. Transaction costs and contractual incompleteness: the case of Électricité de France. J. Econ. Behav. Organ. 42 (2), 189–206.
Schermann, M., Dongus, K., Yetton, P., Krcmar, H., 2016. The role of transaction cost economics in information technology outsourcing research: a meta-analysis of the choice of contract type. J. Strat. Inf. Syst. 25 (1), 32–48.
Shah, R., Ward, P.T., 2007. Defining and developing measures of lean production. J. Oper. Manag. 25 (4), 785–805.
Sitkin, S.B., Pablo, A.L., 1992. Reconceptualizing the determinants of risk behavior. Acad. Manag. Rev. 17 (1), 9–38.
Spina, G., Caniato, F., Luzzini, D., Ronchi, S., 2016. Assessing the use of external grand theories in purchasing and supply management research. J. Purch. Supply Manag. 22 (1), 18–30.
Stritch, J.M., Pedersen, M.J., Taggart, G., 2017. The opportunities and limitations of using mechanical Turk (MTurk) in public administration and management scholarship. Int. Publ. Manag. J. 20 (3), 489–511.
Suprapto, M., Bakker, H.L., Mooi, H.G., Hertogh, M.J., 2016. How do contract types and incentives matter to project performance? Int. J. Proj. Manag. 34 (6), 1071–1087.
Tabachnick, B.G., Fidell, L.S., 2001. Principal components and factor analysis. Using Multivariate Statistics 4, 582–633.
Tversky, A., Kahneman, D., 1992. Advances in prospect theory: cumulative representation of uncertainty. J. Risk Uncertain. 5 (4), 297–323.
Williamson, O.E., 1975. Markets and Hierarchies. New York, 2630.
Williamson, O.E., 1979. Transaction cost economics: the governance of contractual relations. J. Law Econ. 22 (2), 233–261.
Williamson, O.E., 1985. The Economic Institutions of Capitalism. Simon and Schuster.
Williamson, O., 1991. Comparative economic organization: the analysis of discrete structural alternatives. Adm. Sci. Q. 36 (2), 269–296.
Williamson, O.E., 2002. The theory of the firm as governance structure: from choice to contract. J. Econ. Perspect. 16 (3), 171–195.
Williamson, O.E., 2008. Outsourcing: transaction cost economics and supply chain management. J. Supply Chain Manag. 44 (2), 5–16.
Wynstra, F., Suurmond, R., Nullmeier, F., 2019. Purchasing and supply management as a multidisciplinary research field: unity in diversity? J. Purch. Supply Manag., 100578