ELSEVIER

Journal of Operations Management 16 (1998) 361–385

JOURNAL OF OPERATIONS MANAGEMENT

A definition of theory: research guidelines for different theory-building research methods in operations management

John G. Wacker *

Department of Management, Iowa State University, Ames, IA 50011-2065, USA

Abstract

This study examines the definition of theory and the implications it has for the theory-building research. By definition, theory must have four basic criteria: conceptual definitions, domain limitations, relationship-building, and predictions. Theory-building is important because it provides a framework for analysis, facilitates the efficient development of the field, and is needed for the applicability to practical real world problems. To be good theory, a theory must follow the virtues (criteria) for 'good' theory, including uniqueness, parsimony, conservation, generalizability, fecundity, internal consistency, empirical riskiness, and abstraction, which apply to all research methods. Theory-building research seeks to find similarities across many different domains to increase its abstraction level and its importance. The procedure for good theory-building research follows the definition of theory: it defines the variables, specifies the domain, builds internally consistent relationships, and makes specific predictions. If operations management theory is to become integrative, the procedure for good theory-building research should have similar research procedures, regardless of the research methodology used. The empirical results from a study of operations management over the last 5 years (1991–1995) indicate imbalances in research methodologies for theory-building. The analytical mathematical research methodology is by far the most popular methodology and appears to be over-researched. On the other hand, the integrative research areas of analytical statistical and the establishment of causal relationships are under-researched. This leads to the conclusion that theory-building in operations management is not developing evenly across all methodologies. Last, this study offers specific guidelines for theory-builders to increase the theory's level of abstraction and the theory's significance for operations managers. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Theory; Research guidelines; Operations management

1. Introduction

Although many business professionals, social scientists, and other academics have very similar beliefs on the definition of theory, there are still some differences of opinion in a theory's exact nature. For example, some practitioners and academics believe that theory and its application are very limited and, therefore, not very useful in the real world of business. Others feel that very little theory exists in the academic world. For example, consider the following discussions of theory and scientific investigation:

Theory, for theory's sake, can easily degenerate into an uninteresting art form. Yet, practice without theory can quickly become a dull and dangerous occupation. Unfortunately, the world is a complicated place and complicated solutions and processes are

* Tel.: +1-515-294-8111; e-mail: wacker@iastate.edu

PII S0272-6963(98)00019-9

often required to make complex organizations run. The ability to live with uncertainty and the insight into both one's professional powers and limitations is the sign of a mature management science (Shubik, 1987).

... Of all our valid knowledge of the social world, most of it seems to have been the product of lay rather than professional inquiry... A typical situation in social science is that scientific inquiry only modestly raises the validity of a lay proposition by qualifying it (Lindblom, 1987, p. 517).

... (N)o theory of inventory exists unless the theory dictates the manner in which the measurements of the parameters of the theory are to be made. This amounts to saying that today we have no theory of inventory. We will only have such a theory when we can specify the information necessary to test the theory and justify this specification (Churchman, 1961, 132).

The implication of the first statement is that theory does not necessarily require application. This statement means that theory is abstract and does not have to be applied or tested to be a 'good' theory. If this is the case, then theory can remain totally abstract and non-applied. The crux of whether theory can be totally abstract and non-applied depends on the definition of theory.

The second statement implies that scientific investigation and theory-building are not necessarily useful for the development of social science or managerial decision-making fields of academic study. This statement implies that 'good' theory is 'found' through trial and error rather than through scientific investigation in a systematic matter. Further, this statement implies that 'good' theory-building in social sciences and managerial decision-making is derived from lay investigation rather than by scientific investigation. The validity of this statement depends upon how theory is defined.

The third statement implies that unless theory explicitly indicates how it is measured, it is not a theory. Churchman (1961) states that since inventory theories do not indicate how they are to be measured, they are not theories. The statement infers that a theory should offer how it is to be measured for empirical testing, and without this testing, a theory cannot be 'good' theory. Churchman's conclusion that no inventory theory exists in management also depends upon the definition of theory.

All three of the above statements cause concerns for academics since they infer three common criticisms of theory: (1) theory does not have to be applied, (2) it does not make significant improvements in the external world, and (3) theory does not exist due to lack of measurement of definitions. Each concern has some measure of truth. Unless operations management research addresses relevant practical problems to explain complex phenomena, it cannot develop into a theory-building discipline. Each of the above concerns depends upon the definition of theory and, more importantly, what criteria is used to develop 'good' theory. It does not seem logical to call any theory a 'good' theory without defining theory or the criteria for 'good' theory.

Besides these concerns, there are additional reasons why theory is important for researches and practitioners: (1) it provides a framework for analysis; (2) it provides an efficient method for field development; and (3) it provides clear explanations for the pragmatic world. The first reason is that good theory provides a framework of analysis for operations management since it provides structure for where differences of opinion exist. For example, a 'theory of internationally competitive manufacturing' would provide a structure to evaluate factories' manufacturing international competitiveness. This theory enables academics to enumerate the exact conditions to classify firms into degrees of international competitiveness. Since it is unlikely that all academics would agree on which factors are most important in a hierarchy of factors, a framework of 'good' theory procedures provides an understanding of where these differences of opinion lie.

The second reason for developing theory is efficiency. Theory development reduces errors in problem-solving by building upon current theory. Building upon current theory is equivalent to incorporating all that is known from the current literature (theoretical, mathematical, empirical, and practitioner research) into a single, integrated consistent body of knowledge. For researchers, using a single integrated body of knowledge for analytical and empirical testing gives the results a deeper theoretical meaning by differentiating between the competing theories. An integrated body of knowledge can only be pursued efficiently if integrated theory is developed through consistent theory-building methodologies.

A third reason why theory is important is its applicability. Consider this example: a group of managers hires a consultant to improve their firm's competitiveness. The consultant provides a set of suggestions. If the consultant does not provide anything else, there is little reason to be surprised if the managers disregard these suggestions, since managers need more rationale than a set of rules. More specifically, they want to know why this set of rules hold for their specific manufacturing facility (what, who, where, when) and how and why these specific relationships could, would, or should improve (prediction claim) their manufacturing facility. Although managers seek advice, it is essential for them to be skeptical of any suggestions since they are directly responsible for both good and bad results. Consequently, they must logically be convinced that nothing important is overlooked. They want the analytical reassurance that these suggestions are logically and practically compatible with each other. Additionally, they want to know specific instances where these rules have been successful (empirical support). Therefore, even in the most practical of instances, use of the formal definition of theory is important for all managers wishing to achieve measurable results. For this reason, many academics believe that: ''There is nothing as practical as a good theory'' (Hunt, 1991; Van de Ven, 1989). ''Good theory is practical precisely because it advances knowledge in a scientific discipline, guides research toward crucial questions, and enlightens the profession of management'' (Van de Ven, 1989).

Because of the need for 'good' theory, the purpose of this study is to set guidelines for integrating theory-building research. In order to achieve this lofty goal, this study first determines what a 'good' theory is. Next, it closely ties 'good' theory to specific research procedures. After the good-theory-building procedures are outlined, this study categorizes different research methodologies by their theory-building purpose and depicts how all theory-building procedures are similar. Then this study empirically investigates the prevalence of different research methodologies in operations management to determine where operations management is weak in theory-building research. Last, the article concludes with the guidelines for theory-building research. As are all articles suggesting theory-building and research investigation, this article is an opinion article since ''the value laden nature of theory can never be eliminated'' (Bacharach, 1989). Therefore, the opinions expressed here should be evaluated on their logical appeal and internal consistency.

2. A definition of theory

Before beginning any discussion on theory, this study must differentiate between the common notion of 'theory' and a formal definition of theory. In this article, the term theory is interpreted as following the formal definition and operationalization of theory. This operationalization of the definition of theory should directly be tied to the necessary components of theory. Generally, academics point to a theory as being made up of four components, (1) definitions [1] of terms or variables, (2) a domain [2] where the theory applies, (3) a set of relationships of variables, and (4) specific predictions (factual claims) (Hunt, 1991; Bunge, 1967; Reynolds, 1971). Theories carefully outline the precise definitions in a specific domain to explain why and how the relationships are

[1] *Theoretical definitions* are not observable as such. Rather, their existence and properties are asserted in order to account for what is observable. Theoretical definitions are conceptual in nature. Even a relatively simple term such as manufacturing lead time has conceptual foundations that transcend its measurement. By definition, manufacturing lead time is ''The total time required to manufacture an item, exclusive of lower level purchasing time'' (Crawford, 1987). Conceptually, manufacturing lead time represents the internal time a manufacturing system takes to manufacture an item. This concept is not observable because the precise moments from when all materials become available to when the order is completed cannot be exactly specified.

[2] *Domain of a theory*: the domain of the theory is the exact setting or circumstances where the theory can be applied. For example, a just-in-time theory domain may be manufacturing facilities that focus on few products and compete on cost with acceptable quality. The just-in-time theory domain would give extensive factors that limit the instances of when and where it is applied.

logically tied so that the theory gives specific predictions. Therefore, the precision of good theory causes a theory to be very exacting for all the key components of a theory, or as Poole (1989) and Van de Ven (1989) state: ''A good theory is, by definition, a limited and fairly precise picture.'' A theory's precision and limitations are founded in the definitions of terms, the domain of the theory, the explanation of relationships, and the specific predictions.

Authors usually agree that the goal of 'good' theory is a clear explanation of how and why specific relationships lead to specific events. Consequently, these explanations of relationships are critical for 'good' theory-building. Other authors' statements on theory indicate the importance of relationship-building:

(Theory is)... an ordered set of assertions about a generic behavior or structure assumed to hold throughout a significantly broad range of specific instances (Sutherland, 1976: 9).

Researchers can define theory as a statement of relationships between units observed or approximated in the empirical world. *Approximated* units mean constructs, which by their very nature cannot be observed directly. ... A theory may be viewed as a system of constructs and variables in which the constructs are related to each other by propositions and the variables are related to each other by hypotheses (Bacharach, 1989).

These statements indicate the importance of relationship-building in explaining how and why specific phenomena will occur. Sometimes how and why and specific predictions are condensed into the expression 'adequate explanation', which implies that unless an explanation can predict, it is not considered adequate (Hunt, 1991).

A very important aspect of a theory definition is phrased in the common questions that researchers require to exactly specify a theory. Consider this statement: ''The primary goal of a theory is to answer the questions of how, when (or where), and why... unlike the goal of description, which is to answer the question of what (or who)'' (Bacharach, 1989). In short, any definition of theory should answer common questions that researchers face. First, theory defines all variables by answering the common questions of who and what. The domain specifies the conditions where the theory is expected to hold by using the common questions of when and where. The relationship-building stage specifies the reasoning by explaining how and why variables are related. And last, the predictive claims specify the whether ''Could a specific event occur?'', ''Should a specific event occur?'', and ''Would a specific event occur?'' In short, the definition of theory provides guidelines to answer the common questions that occur in natural language. From the pragmatic perspective of operations managers, the predictive claims from theory answer the could, should, and would questions which are quite critical for managers' future success. Consequently, the should, could and would questions are very important for theory to be considered useful to managers.

In summary, the definition of theory suggested by this study has these four components: definitions, domain, relationships, and predictive claims to answer the natural language questions of who, what, when, where, how, why, should, could and would.

3. Virtues of 'good' theory

The purpose of this section is to limit the domain of theory-building to the domain of what Popper (1957) calls 'good' theory. Theories in the domain of 'good' theory are superior to other theories because they possess what philosophers of science call the theory's virtues (Popper, 1957; Kuhn, 1980; Quine and Ullian, 1980). 'Good' theory's superiority is important because researchers must evaluate the relative significance of opposing theories. Although there is no general agreement between philosophers of science concerning the relative importance of each virtue of 'good' theory, there seems to be a fairly widespread agreement as to what they are. A common set of virtues of a theory is: uniqueness, parsimony, conservatism, generalizability, fecundity, internal consistency, empirical riskiness, and abstraction (Quine and Ullian, 1980).

Table 1 provides an overview of the virtues of 'good' theory. The first five virtues are fairly straightforward: each theory must be differentiated from other theories (uniqueness): new theories cannot replace existing theories unless they are better

Table 1
The virtues of 'good' theory: key features and why they are important to 'good' theory development

Virtue	Key feature	Why important for 'good' theory and for the development of the field
Uniqueness	The uniqueness virtue means that one theory must be differentiated from another.	If two theories are identical, they should be considered a single theory. Although it applies to all criteria for theory, this virtue directly applies to definitions since definitions are the most elemental of building blocks for theory.
Conservatism	A current theory cannot be replaced unless the new theory is superior in its virtues.	Therefore, current theory is not rejected for the sake of change. This criteria is needed so that when a new theory is proposed, there is a good reason to believe all other theories are lacking in some virtue (Quine and Ullian, 1980; Kuhn, 1980; Popper, 1957).
Generalizability	The more areas that a theory can be applied to makes the theory a better theory.	If one theory can be applied to one type of environment and another theory can be applied to many environments, then the second theory is a more virtuous theory since it can be more widely applied. Some authors call this virtue the utility of the theory since those theories that have wider application have more importance.
Fecundity	A theory which is more fertile in generating new models and hypotheses is better than a theory that has fewer hypotheses	Theories which expand the area of investigation into new conceptual areas are considered superior to theories which investigate established research areas. This means
Theory parsimony theory simplicity, theory efficiency Occum's razor	The parsimony virtue states, other things being equal, the fewer the assumptions the better.	If two theories are equal in all other aspects, the one with fewer assumptions and the fewer definitions is more virtuous. This virtue also includes the notion that the simpler the explanation, the better the theory. This virtue keeps theories from becoming too complex and incomprehensible.
Internal consistency	Internal consistency means the theory has identified all relationships and gives adequate explanation.	Internal consistency refutation means that the theory *logically* explains the relationships between variables. The more *logically* the theory explains the variables and predicts the subsequent event, the better the theory is. This *internal consistency* virtue means that the theory's entities and relationships must be *internally compatible* using symbolic logic or mathematics. This *internal consistency* means that the concepts and relationships are logically *compatible with* each other.
Empirical riskiness	Any empirical test of a theory should be risky. Refutation must be very possible if theory is to be considered a 'good' theory.	If there are two competing theories, the theory that predicts the most unlikely event is considered the superior theory. In the opposite case, if the theory predicts a very likely event then it is not seen as being a very valuable theory. This criteria is sometimes put in a different way: ''Every good theory has at least one prohibition; it prohibits certain things from happening'' (Popper, 1957).
Abstraction	The abstraction level of theory means it is independent of time and space. It achieves this independence by including more relationships.	The abstraction level means it is better to integrate many relationships and variables into a larger theory. If one of two competing theories integrates more internally consistent concepts, it is more virtuous than a theory that integrates fewer internally consistent relationships.

theories evaluated in the light of their virtues (conservation): new theories can be introduced that have rich new areas of investigation (fecundity virtue, some academics would call this area grounds for a paradigm shift); and last, any theory should not be overly complex (theory parsimony).

However, the remaining virtues are worthy of additional discussion since they are the focus of most of the academic discussion. There are two theory refutation virtues: internal consistency and empirical riskiness. The internal consistency virtue implies that the theory is internally consistent using either mathematics or symbolic logic (Bacharach, 1989). Without internal consistency, each relationship is independent from every other relationships causing the theory to not be integrated within itself. This virtue is different from face validity because each relationship in a theory may be explained individually, and therefore the relationship may have face validity. However, each relationship may not be consistent with other relationships in the theory, thus, causing the entire theory to be internally inconsistent. Since theories tie together many concepts, it usually is not easy to identify internal inconsistency. Therefore, it is necessary to develop the theory using mathematics or symbolic logic.

The criterion of empirical riskiness has been the focus of most of critical evaluators of 'good' theory. Most academics believe that empirical tests of theory should be risky so that there is a good chance of the theory being refuted. On the other hand, if an empirical test of theory supports the occurrence of a likely event, then the theory is deemed to be a 'weak' theory. One empirical riskiness criteria is frequently called the 'single event refutation' criteria which states that it takes only one sample to question a theory's legitimacy. Put another way, every legitimate empirical test is designed to disprove the theory and should be risky (Popper, 1957).

Sometimes the two refutation criteria are not compatible. The incompatibility between the two refutation criteria typically occurs when empirical evidence does not support an existing theory. What are the deciding criteria to determine which evidence is more important, empirical or logical? Generally, the logical evidence is considered superior to the empirical evidence since there is less chance of error. In the philosophy of science the phrase used is the 'power of deduction rules'. The 'power of deduction rules' provides the underlying structure for theory-building research (Popper, 1957). A brief example may prove useful to clarify why the 'power of deduction rules'. Suppose a researcher gathered a large set of data on a bank and statistically found that when the arrival rate is greater than the service rate, the line is shorter. To most operations management academics, there is no amount of empirical evidence that could convince them that having greater arrival rates than service rates will lead to shorter lines. This conclusion is contrary to queuing theory's deductive conclusion. In short, the ''power of deduction rules''.

Yet, most operations management academics would investigate the specific circumstances to determine which queuing model should be applied. Or alternatively, academics would develop a new theory to explain the empirical evidence. Therefore, based upon the empirical evidence, operations management academics may develop a 'new' internally consistent theory (model) to replace the existing theory (model).

The last virtue is the theory's abstraction level, which is usually classified into three levels: high, middle, and low. High abstraction level theories (general or grand theories) have an almost unlimited scope, middle abstraction level theories explain limited sets of phenomena (which serve as the raw materials for the construction of more general theories), and lower level theories (called empirical generalizations of very limited scope) serve as simple relationship identifications (Bluedorn and Evered, 1960). These three levels of abstraction are points on an abstraction continuum. A low abstraction level theory can be directly applied to a specific instance since a low abstraction level theory completely defines the variables, domain, relationship(s), and then states the factual claim for this specific instance. However, the usefulness of low abstraction level theories to academics and practitioners is limited because in the majority of instances, the theory cannot be applied since the definitions and domain are so narrow that the theory only applies to a few instances. Yet, low abstraction level theories are used to build middle abstraction level theories which, in turn, lead to high abstraction level theories. At the highest abstraction level, the application of a theory is the broadest since the domains are broad. High level abstraction theory, therefore, has wide application. Consequently, an important goal of theory-building is to advance lower abstraction level theories to middle level theories, and then to high abstraction level theories. This leads to the generalization that as a field progresses, it tends to have theory in higher abstraction levels. This progression from low abstraction to high abstraction is called ladder-climbing (Osigweh, 1989).

In operations management, ladder climbing is becoming evident by several of the empirical generalizations in just-in-time (JIT) theories. For example, the following are three low-level abstraction theories: (1) shorter set-up times facilitate smaller lot sizes, (2) smaller lot sizes cause lower work-in-process, and (3) lower work-in-process causes increased quality. These three different smaller theories are being integrated into a large theory through a more careful explanation as to how they are related to each other. When addressed individually, each of these theories is a low abstraction level theory. Yet, when combined into single integrated theory, it becomes a mid-range theory (Miller and Roth, 1995; Demeyer and Arnoud, 1989; Wacker, 1987, 1996). Still, the integration into high abstraction level theories seems to be somewhat in the future for operations management.

Though all of these theory virtues are highly significant for theory-building, the relative weight given to each virtue becomes important for comparing competing theories, since virtues trade off with each other. For example, a new theory may explain phenomena in many domains (generalizability), but be relatively complex (hampering the parsimony criteria). Or a theory may be rich in hypotheses development (fecundity), but can only be applied to a very limited set of conditions (hindering the generalizability criteria). Or theories that are generalizable, may not be internally consistent (hindering the internal consistency virtue). To some degree, each ‘good’ theory weighs each virtue against other virtues. This evaluation requires value judgments to decide which theory is superior to other theories based on each theory’s virtues. Although there are always trade-offs among virtues, the internal consistency and empirical riskiness virtues are generally considered the most important virtues since without refutation being possible or likely, the theory becomes tautological (Popper, 1957).

4. When does a theory become a ‘good’ theory?

One of the most difficult questions to answer about theory is: when does a theory become a ‘good’ theory? The answer to that question is inherently controversial since it involves the degree to which individuals believe in adhering to the formal definition of theory and follow the virtues of ‘good’ theory. For the purpose of this article, a ‘good’ theory should meet the definitional criteria of theory as well as follow the virtues of ‘good’ theory. Therefore, ‘good’ theory must first be a theory (i.e., have definitions, have a domain, have relationships, and make predictions) and must meet each ‘good’ theory virtue to some degree.

Any theory which adheres to both the definition of theory and the virtues of a ‘good’ theory is a ‘good’ theory. This adherence, however, does not mean that the ‘good’ theory is valid since ‘good’ theories can be ‘just plain wrong’. Yet, ‘good’ theories which are wrong are more quickly identified as being wrong since they are more easily refuted (internally inconsistent or empirically invalid). Still, ‘good’ incorrect theories serve a very important purpose for field development, since ‘good’ theory is a beginning point to determine why the theory is wrong. Therefore, because ‘good’ theory is easily refuted and is a beginning point for future investigation, using ‘good’ theory for empirical tests seems to be a laudable objective for the development of an academic field.

On the other hand, a theory that violates one of the virtues of ‘good’ theory is difficult to refute. Since the virtues of theory apply to the definitional criteria of theory, there are innumerable ways a ‘bad’ theory can be developed. Consequently, it is possible to give myriad examples of how a ‘bad’ theory hinders the development of an academic field by detracting research efforts away from ‘good’ theories.

5. The general procedure for ‘good’ theory-building research

The purpose of this section is to demonstrate a general research procedure that satisfies the following specifications: theory’s definition, ‘good’ theory’s virtues, and answers to the natural language questions of who, what, when, where, why, how, should, could, and would. From the logical perspective, these guidelines are all necessary conditions for theory-building (Whetten, 1989; Van de Ven, 1989). The procedure must state the purpose of each step.

The academic literature suggests many different procedures for theory-building (Eisenhardt, 1989; Swamidass, 1986; Bacharach, 1989). Generally speaking, these procedures suggest how to operationalize specific types of research projects. The research procedures suggested here are similar to those operationalizations of research. However, this study identifies a minimal procedure for ensuring that all guidelines for 'good' theory-building are met regardless of the type of research. Each step reflects a necessary condition to fulfill 'good' theory-building guidelines, since all theory-building requires definitions, domain, relationships, and predictions. Consequently, the same procedure is required regardless of the methodology used for any theory-building research. Therefore, from the viewpoint of theory's virtues, this study suggests that the simplest, most generalizable procedure be adopted. The requirement that theory-building research follows the definition of theory leads to similarities in theory-building research for both analytical and statistical procedures. More about these two basic research methods are discussed below.

Although the stages suggested here are listed sequentially, they are not necessarily sequential since they interact with each other. For example, when a theory is visualized, a new definition may be needed to identify a new concept. Or when a prediction or factual claim is made, it may require a different domain or a different relationship. In addition, the operationalization of variables for empirical tests may require some modifications of selected definitions. Consequently, although this procedure is listed in stages, a research project may not necessarily proceed inexorably sequentially through each stage to a 'good' theory.

For all stages of theory-building, the role of the literature search in the research procedure is extremely important since it provides the accepted definitions, domains of where a theory applies, previously identified relationships (along with empirical tests), and specific predictions of other theories. Therefore, to assure that all theory-building conditions are fulfilled, an extensive literature search of the academic as well as the practitioner articles is required.

Table 2 outlines the general procedures for developing theory. The precise definitions of variables are

Table 2
A general procedure for theory-building and the empirical support for theory

	Purpose of this step	Common question	'Good' theory virtues emphasized
Definitions of variables	Defines who and what are included and what is specifically excluded in the definition.	Who? What?	Uniqueness, conservation
Limiting the domain	Observes and limits the conditions by when (antecedent event) and where the subsequent event are expected to occur.	When? Where?	Generalizability
Relationship (model) building	Logically assembles the reasoning for each relationship for internal consistency.	Why? How?	Parsimony, fecundity, internal consistency, abstractness
Theory predictions and empirical support	Gives specific predictions. Important for setting conditions where a theory predicts. Tests model by criteria to give empirical verification for the theory. The riskiness of the test is an important consideration.	Could the event occur? Should the event occur? Would the event occur?	Empirical tests refutability

needed to limit the area of investigation by defining who and what. Generally, the literature provides a base for defining the variables. New definitions must demonstrate why current definitions are not adequate, otherwise the conservation virtue is violated. Ordinarily, currently defined concepts should be used to avoid violating the conservation virtue. Although there are many difficulties in precisely defining and measuring variables, and the complete discussion of the implications are beyond the scope this paper, there are several concerns which bear mentioning. One important difficulty with definitions is what has come to be known as 'concept stretching' which occurs:

> ... (W)hen concepts are broadened in order to extend their range of application, they may be so broadly defined (or stretched) that they verge on being too all-embracing to be meaningful in the realm of empirical observation and professional practice. ... Because of this broadness, it tends to confuse more than help the development of the field: ... Because concept stretching in organizational science results in amorphous unclear conceptualizations, what appears to have been a gain in extensional coverage (breadth) often has been matched and even surpassed by losses in connotative precision (depth) (Osigweh, 1989).

One common source of 'concept stretching' is natural language which is too broad for precise measurement, causing a research project to build an ill-defined theory, which in turn, gives misinformation (Osigweh, 1989). Because of the difficulties with natural language, artificial languages and definitions (sometimes called formal language) must be developed to avoid the confusions caused by natural language. Some feel that theory-building research cannot be effectively achieved unless the academic field has a precise artificial language (Teas and Palan, 1997). To develop an artificial language and avoid 'concept stretching', some theorists suggest that the definitions be examined by the negation principle. That is, definitions must be examined by what they specifically exclude. Although it is not true in the extreme, generally, the more a conceptual definition excludes, the more precise it is and the more likely it will be unique (Osigweh, 1989).

Once the precise definitions of variables are established, the domain is established to limit when and where the theory hold. The domain of the theory directly limits the its generalizability since the more specific the domain of the theory is, the lower the generalizability. Theory-building research extends domains to new broader areas by testing the theory in a new environment or a different time period. In short, theory-building research extends the domain of the theory.

Typically, after both the definitions are clarified and the domain is specified, relationship (model) building begins. This step is necessary to establish which variables have logical connections to other variables. Technically, every variable used should explicitly state how and why it is related or unrelated to each and every variable in the model. The summation of the related plus the unrelated variables should be equal to the total number of variables in the model. Hence, in any given theory, the relationship between any two variables in the theory must be explicitly stated, or else the theory cannot be shown to be internally consistent. The third stage is quite complex since the theory (and the model by which the theory is to be tested) should be fully developed and internally consistent. For this step, the academic literature suggests which relationships are important for the development of the theory. The more carefully a researcher builds the relationships from other research, the more theoretically important the research is, since it is integrating theory to raise its abstraction level.

In the relationship-building step, there are four types of theory-building relationships: those relationships that are assumed to be true (fundamental laws or axioms); those laws that are derived from the fundamental laws (derivative laws or theorems); those laws that span the gap from the theoretical to the empirical world (bridge laws or guiding hypotheses), and the relationships that are being investigated (research or theoretical hypotheses) (Hunt, 1991). For example, research on cost and small lot sizes analytically assumes that the number of orders that are placed in a year is equal to demand divided by the fixed order size (a fundamental law). Additionally, the average inventory is equal to maximum plus minimum inventory divided by two (a second fundamental law). The total cost of inventory is equal to the order cost times the number of orders plus the holding cost times the average inventory (a derived law or theorem). The cost optimization solution leads to the economic order quantity (EQQ) (a second derived law). A bridge law is necessary to tie the analytical results to the empirical world. For example, given the annual demand, annual holding cost, and ordering cost, firms should have a predicted order quantity (the research or theoretical hypothesis). One bridge law that would facilitate this empirical test is: rational firms act to minimize their total inventory costs. Note that without a bridge law, the EQQ derived law does not span into empirical world. The area of bridge laws is important for integration of theory in operations management. It is the lack of bridge laws that Churchman (1961) referred to the beginning of the study. Consequently, Churchman's criticism cannot be summarily dismissed, since many analytical research studies do not give any guidance as to how they are to be tested in the external empirical world.

In the above example, fundamental and derived laws are not tested, but are assumed to be true analytically. If these relationships were tested along with other relationships, the statistical complexity may be too complex for meaningful investigation. Because of this complexity, it is important that researchers limit their investigation using the literature, since the parsimony virtue may be violated. The literature provides the best guidelines as to which relationships are theoretically important for investigation and which relationships may be considered fundamental or derived laws that do not need consideration in the investigation. Consequently, theory-building research uses the literature as a guideline to decide which relationships are important for investigation.

The last stage is theory prediction. In order for a theory to meet the minimal requirements for the definition of theory, it is technically not necessary that a theory offer evidence to support predictive claims (only predictions are necessary). Yet, from a practical perspective, it is possible that a huge number of theories could be proposed. Therefore, many academics prefer that empirical evidence be presented to verify that a proposed theory has some merit in the empirical world. Different methodologies use different empirical evidence to verify their predictive validity. More will be discussed about this predictive stage below when discussing procedures for analytical and empirical research.

Though each stage of theory-building needs to follow all the virtues of 'good' theory, each stage is more closely related to certain specific virtues. For example, in the definition step, uniqueness and conservation are key virtues since they limit theory-building to developing current concepts before developing new unique concepts. Some researchers argue that many definitions such as total quality management, continuous improvement, and just-in-time overlap and are neither unique, nor new (Westbrook, 1987; Robinson, 1990). They argue that these concepts do not limit the definition and are not necessarily unique from each other which makes them malformed thus causing data-gathering difficulties as well as misinformation (Osigweh, 1989). In the domain specification stage, the generalizability virtue is important because the more domains in which a theory can be applied, the more important the theory is. In the relationship-building step, parsimony, fecundity, and abstraction virtues enhance the theory by using only necessary relationships, offering new areas for investigation, and integrating relationships for a higher abstraction level. Also in this stage, internal consistency is important to verify which relationships are logically compatible with each other. Generally, as more internally consistent relationships are integrated into a theory, the theory can explain more, therefore raising the theory's abstraction level. In the theory prediction step, the importance of internal consistency and empirical riskiness are both needed for the theory to make predictions.

Table 2 summarizes the procedure for 'good' theory-building research. Each of the stages are required for a theory to be considered a 'good' theory. Unless a proposed theory has all these stages, it does not meet the criteria of the formal definition of theory. The first column shows the components of theory. The second column states why these components are necessary. The third column gives the common question that each stage addresses. While the first three columns use the definitional criteria for theory, the last column gives the most relevant virtues for that stage to ensure that any theory developed is a 'good' theory.

'Good' theory-building research requires the fulfillment of the formal definition of theory. Additionally, a 'good' theory should not violate any 'good' theory virtue. The stages of a 'good' theory-building mirrors the four criteria for theory and meet the virtues of 'good' theory. However, since it is relatively easy to propose many 'good' theories that meet the theory criteria, usually some empirical evidence is needed for 'good' theory to be 'good' theory-building. Therefore, empirical evidence usually is included in the prediction stage to support the theory (it could be argued that this is an additional stage of theory-building). Consequently, 'good' theory-building should have empirical evidence to suggest why the theory has some legitimacy in the empirical world.

As a side note, the statements by Shubik (that theory can be totally impractical) cannot be summarily dismissed since a 'good' theory does not have to be tested to be 'good' theory. In short, it is possible that a new 'good' theory may not yet have external empirical evidence to support it. For this reason, many academics believe that empirical evidence should be offered before any theory is considered a 'good' theory (Whetten, 1989; Poole and Van De Ven, 1989).

6. The two objectives of research: some important observations

The two general objectives of research are theory-building and fact-finding. The difference between these two objectives is grounded in the purpose of the research. 'Good' theory-building research's purpose is to build an integrated body of knowledge to be applied to many instances by explaining who, what, when, where, how and why certain phenomena will occur. On the other hand, 'good' fact-finding research's purpose is to build a lexicon of facts that are gathered under specified conditions. The division line between the two types of research is quite fine since both types of research usually include data gathering and empirical estimation. However, the contrast between the two purposes comes from two areas: how and when the a priori predictions are made; and how the results are integrated with other studies.

First, theory-building research carefully defines concepts, states the domain, explains how and why relationships exist, and then predicts the occurrence of specific phenomena. After the prediction, it typically gathers evidence to see if the phenomena occurs. 'Good' fact-finding research also carefully defines concepts and states domains. Then, 'good' fact-finding research uses evidence to discover if relationships exist. Next, 'good' fact-finding research then explains how and why specific phenomena occurred. In the formal sense of the definition of theory, fact-finding research is not theory-building research since the evidence is gathered before the relationships are explained and before the relationships are predicted. Therefore, fact-finding research does not meet the conditions of 'good' theory-building, since it lacks a priori explanations and predictions before the data are gathered.

Even though fact-finding research is not theory-building research, it serves an important purpose by providing facts that can later be integrated into a theory. 'Good' fact-finding research serves to provide fertile ground for subsequent new theory-building. Fact-finding research is not constrained by existing theory-based relationships and, therefore, can more easily investigate new relationships that can later be integrated into a theory. However, to be of value for 'good' theory-building, 'good' fact-finding research must follow the first two conditions for theory, since without precise definitions and well-defined domains, facts are not meaningful for any future use. Because fact-finding research investigates relationships and then offers explanations as to why the results occurred, fact-finding research suggests new relationships for new theory development. Due to the freedom from the constraints placed on theory-building research, fact-finding research frequently provides the basis for new theory development. Consequently, there is a large measure of truth in the statement of Lindblom (1987) statement (made at the beginning of this study) that many new theories come from fact-finding investigation.

A second contrast between theory-building and fact-finding research is the degree of integration of the results with other studies. This degree of integration has two dimensions: the internal integration of all the empirical data and the integration of the results with other studies. Empirical evidence used in theory-building research stresses subtle systematic similarities between all data to raise the theory's abstraction level. These systematic similarities provide additional factors to explain all data in one integrated framework applied to diverse environments. By using one theory across many different environments, theory-building research raises the abstraction level to explain how and why the theory can be applied to predict events. When the empirical data do not support the theory, the reason for the lack of support is important for the further development of the theory. If the theory is logically internally consistent, then the reason for the lack of support must be a missing explanatory factor(s). This explanatory factor(s) is introduced into the theory and then the theory is examined for internal consistency. Once the 'new' theory is internally consistent, then the theory can be empirically tested. This sequence of empirical tests contradicting the theory, introducing new factor(s), building a 'new' internally consistent theory, and empirical testing of the 'new' theory is a common theory-building sequence for theory building. This sequence has the goal of building a single integrated theory to explain all data. In short, when empirical tests do not support a theory, theory-building research examines the theory to integrate new factors to raise the abstraction level. Consequently, 'good' theory-building research is always striving to find integrating factors to expand the theory to apply to diverse environments, thereby raising the theory's abstraction level.

In contrast, fact-finding research typically stresses the descriptive differences in data. These descriptive differences exist because all data, to some degree, are different. Fact-finding research attempts to discover differences in data and explains these differences. Unfortunately, without a theory to be tested in the empirical world, many times the explanations of descriptive differences are not specifically tested. Even more unfortunate, is that they may not be explanations at all. Consider a case where a researcher finds that Japanese manufacturers when compared to USA manufacturers, have statistically significantly higher manufacturing performance on some performance measures. The basic fact indicates that Japanese manufacturing managers are superior to USA manufacturing managers on these performance measures. However, this fact does not offer adequate (or any) explanation of how or why this result occurred since the source of the difference could be better planning (forecasting, and resource scheduling); control (forecast control, production activity control); advanced technology (better information technology, CAD, FMS, etc.); culture (work ethic, managerial behavior, etc.) or numerous other causes; all of which were not specifically tested. In short, fact-finding research frequently explains descriptive differences, but because the explanation is not specifically tested, any inferences and/or conclusions are deceptive. For this reason, these explanations are deceptive descriptive differences.

One type of integration used in theory-building research is the identification of subtle systematic similarities across different studies. For example, Ettlie (1995) noted the empirical similarities between several new product development research studies (Madique and Hayes, 1984; Crawford, 1987; Kekre and Srinivasan, 1990; Klemschmidt and Cooper, 1991). The similarities he deduced were between the two extremes on product development. At one extreme, firms which introduce variants of existing products reap significant benefits from increased market share and profits. At the other extreme, he noticed that firms which introduce totally new products are also reap significant benefits from increased market share and profits. Yet, firms that attempted an intermediate strategy were not successful. Hence, he concluded, ''(T)he relationship between product innovativeness and success in the market place is u-shaped.'' His integration of the three studies into one theory is accomplished by identifying the subtle similarities between different studies. This integration across several studies raises the abstraction level, therefore increasing the importance of his study's findings.

In short, researchers are faced with whether to find subtle systematic similarities or to explain deceptive descriptive differences between individuals, organizations, businesses, industries, and countries. Fact-finding research focuses on descriptive differences among data, while theory-building research concentrates on the underlying factors for similarities. Theory-building research raises the abstraction level by integrating subtle systematic similarities across the descriptive dimensions of individuals, organizations, businesses, industries, and countries.

Consequently, from the standpoint of 'good' theory-building, it seems that systematic similarities are more important than descriptive differences.

7. Types of theory-building research

The purpose of this section is to classify the theory-building types of operations management research by their theory-building purpose. There are two major classifications of research: analytical (formal) and empirical. These differences are best expressed by Sax (1968):

When an organized body is based primarily on deductive rules, it is called formal science, to distinguish it from those areas of knowledge that depend primarily upon empiricism and induction... called empirical science. Considered from this point of view, mathematics, logic, and library science are primarily formal sciences, whereas chemistry, psychology, and education are primarily empirical sciences.

This statement means that analytical (formal) sciences use deductive methods to arrive at theories while empirical sciences use induction methods to arrive at theories.

This section uses these traditional major classifications of analytical and empirical theory-building research and further divides them into three sub-categories for each major classification. The categorization of the research types serves two purposes: it illustrates the need for different research methodologies to raise abstraction levels of theory and, it illustrates that the 'good' theory-building procedures are applicable to different types of research (discussed in the Section 8). The first purpose for the classification is based on the goal of the 'good' theory-building research. This study argues that there are six different types of theory-building research, and each type serves to develop and verify theory using different research methodology. This theory verification usually is called triangulation and is extremely important for the final verification of theory (Meredith et al., 1989). Ideally, if a theory is tested by all six methods with supporting results, then these results would be compelling evidence to believe that the theory is confirmed.

7.1. Analytical research

The analytical research method uses deductive methods to arrive at conclusions (Swamidass, 1986). Analytical research methods primarily use logical, mathematical, and/or mathematical–statistical methods. The three different sub-categories of analytical research have subtle differences in how they use logic and mathematics for the development of the theory. Additionally, the analytical sub-categories have different theory-building purposes.

7.1.1. Analytical conceptual research

From a theory-building perspective, the purpose of analytical conceptual research is to add new insights into traditional problems through logical relationship-building. This research methodology comprises new insights through logically developing relationships between carefully defined concepts into an internally consistent theory. These studies usually use case study examples to illustrate these conceptualizations. There are several types of research in this sub-category. One example is called introspective research which uses the researcher's experience to formulate concepts. Consequently, it describes and explains relationships from past experience to develop theory. A second example in this sub-category is conceptual modeling. This modeling is where a mental model of deduced relationships is posited, which may then be evaluated using a framework that captures the essence of the systems under investigation. A third example in this sub-category is hermeneutics research which deduces facts from what is being observed (Meredith et al., 1989).

A representative article of analytical conceptual research is the Kim and Lee (1993) article where the researchers logically tied together manufacturing strategies and production systems. In their study, they logically integrated technical flexibility, technical complexity, and production system type with differentiation and cost efficiency strategies. This type of research falls into the analytical conceptual sub-category since the methodology used is logical and does not use empirical data for theory development.

7.1.2. Analytical mathematical research

The theory-building purpose of this research sub-category is to develop sophisticated relationships between narrowly defined concepts through developing new mathematical relationships to study how the models behave under different conditions. These articles mathematically develop the relationships and give numerical examples from their derivations or computations. Analytical mathematical research does not use external data to test the theory, but instead uses deterministic or simulated data to draw conclusions. The research in this sub-category is sometimes called 'operations research' and/or 'management science'.

There are many types of research in this sub-category: reason/logical deductive theorem proving; normative analytical modeling research; descriptive analytical modeling; proto-typing and physical modeling research methods; experimentation; and mathematical simulation. In all these methods, the models usually are built using formal logic and tested using artificially developed data. Since the data is derived artificially, all these methods are classified under the analytical mathematical research method (Meredith et al., 1989).

7.1.3. Analytical statistical research

This research sub-category integrates logical/mathematical models from analytical research and statistical models from empirical research into a single integrated theory (Moorthy, 1993). The analytical statistical research is different from the analytical mathematical method since its models are explicitly developed for future empirical statistical tests. This methodology uses large bodies of knowledge and integrates them into a single model for future empirical tests. Typically, the variables used for investigating relationships have measurement errors due to random variability caused by the external world. In sum, the purpose of analytical statistical research is to provide larger, more integrated models for empirical statistical testing.

7.2. Empirical research

In the empirical research major classification, the methodology must use data from external organizations or businesses to test if relationships hold in the external world. Empirical research methods could be classified more correctly as 'real world' empirical methodologies. However, since this is a fairly long phrase, it is abbreviated to empirical methods.

7.2.1. Empirical experimental research

The theory-building purpose of this sub-category is to investigate relationships by manipulating controlled treatments to determine the exact effect on specific dependent variables. The research design uses treatment variables that are manipulated to determine their effect on the dependent variable(s). Because direct manipulation of the treatment variables causes direct effects on the dependent variables, this research sub-category comes the closest to demonstrating causality between variables. This sub-category of empirical research is also called 'field experiments' (Meredith et al., 1989).

The empirical experimental research methodology is difficult to implement in operations management since the environment must be closed to contamination effects. In operations management, the system frequently is an open system and therefore can be subject to contamination effects. Yet, it seems possible that controls could be placed on some experiments to determine if one treatment causes a certain result. For example, a researcher may find an organization that is about to install a just-in-time (JIT) system. If the facility has two different lines or manufacturing areas that are separate, the researcher may experiment with the amount and type of training to determine the effect of training on selected performance measures (throughput time, defect rates, etc.) This type of experiment could offer convincing evidence of the necessity or non-necessity for specific amounts or types of training needed for JIT implementation.

This research type is not to be confused with the experimental design in mathematical simulation found in the analytical mathematical sub-category above (where the data are developed by the researcher in a closed simulated environment).

7.2.2. Empirical statistical research

This research sub-category's purpose is to empirically verify theoretical relationships in larger samples from actual businesses. Generally, the more complex the research issues are, the more likely the study will use this methodology. A typical topic would be: manufacturing strategy's effect on manufacturing performance.

There are many types of research that fall into this sub-category. Structured and unstructured (elite and intensive) interviewing processes which gather data for statistical analyses are examples of this sub-category of research. Other representative examples are surveying and historical/archival research, expert panels, and Delphi techniques. Each of these methods has the goal to statistically analyze data from relatively large external samples. Therefore, from a theory-building perspective, this methodology offers empirical support for theoretical relationships in larger samples in real world (Meredith et al., 1989).

7.2.3. Empirical case study

The purpose of this type of research is to develop insightful relationships within a limited set of companies. By limiting the number of companies investigated, this research method investigates small samples using a large number of variables to identify new empirical relationships. Frequently, this research method analyzes organizations across time and provides the dynamic dimension to theory to elevate the theory's abstraction level. This sub-category includes field studies and action research (Meredith et al., 1989).

A typical research article of this sub-category is the article of Marucheck et al. (1990) on manufacturing strategic practice, which examines the strategic practices in six firms. Using the practices of the six firms, they found that firms follow the general conceptual models developed in the academic literature. Additionally, they stated ''..(T)he real benefits of manufacturing strategy come from implementing the manufacturing strategy'' (p. 121). The research project of Marucheck et al. (1990) analyzed data from a limited number of firms to identify manufacturing strategy procedures. Since a limited number of firms are utilized to identify possible theories, this project is classified as empirical case study. Hence, the key difference between the empirical case study method and the analytical conceptual method is that the empirical case study method uses data to form the theory and the analytical conceptual method uses deduction to form theories.

7.3. An important conclusion

From the standpoint of good theory, one important conclusion drawn is that no single research category or sub-category is superior to any other research category or sub-category. Each method serves a very different, but important, purpose for theory development in operations management. The basic purpose of all three types of analytical methodologies is to develop logically internally consistent theories and models. First, the analytical conceptual research type serves as a forum for expressing new conceptual perspectives on theory to better explain and integrate underlying relationships. The sub-category of analytical mathematical research serves to logically evaluate the internal consistency of complex relationships. The analytical statistical research serves to integrate both the analytical mathematical results and the empirical statistical results into a larger theoretical body of knowledge for statistical estimation. In short, all these analytical research sub-categories serve to develop internally consistent theories through logical analyses.

The empirical methodologies provide empirical verification of models, while offering evidence for the development of new theory. The empirical experimental research uses experimental design to verify the causality of a specific theory while elevating relationships from a testable hypothesis to an empirically verified theory. A verified theory may not need to be tested in future research projects and may be assumed to be a fundamental law. The empirical statistical research methodologies verify models for their empirical validity in larger populations to reduce the number of relationships in future research. The empirical case studies provide new conceptual insights by empirically investigating individual cases for an in-depth understanding of the complex external world.

Table 3 provides a summary of the classification of research along with its relative importance to operations management. Basically, theories developed using the analytical research methodologies are refuted by empirical evidence, while theories developed using the empirical research methodologies are refuted by internal consistency. The last row of the table states the theory-building purpose of the research method. It would be difficult to disagree with the importance of any of these six sub-categories, since it would be paramount to stating that operations management does not need: new conceptual theory; internal mathematical consistent theory; integrative theory; causal verification of theory; large

Table 3
Specific research sub-category, refutation methods, and importance to operations management theory-building

	Analytical			Empirical		
	Conceptual	Mathematical	Statistical	Experimental design	Statistical sampling	Case studies
Types of research included	Futures research scenarios, introspective reflection, hermeneutics, conceptual modeling	Reason/logical theorem proving, normative analytical modeling, descriptive analytical modeling, proto-typing, physical modeling, laboratory experiments, mathematical simulation	Mathematical statistical modeling	Empirical experimental design, descriptive analytical modeling	Action research structured and unstructured research, surveying, historical analysis, expert panels	Field studies, case studies
Refutation methods	Empirical data from empirical methods	Empirical data from empirical methods	Empirical data from empirical methods	Analytical/logical inconsistency	Analytical/logical inconsistency	Analytical/logical inconsistency
Importance to operations management theory-building	Develops new logical relationships for conceptual models of theory	Explores the mathematical conditions underlying the relationships used in theory-building	Integrates the other five methods into a single theory for empirical investigation	Tests and verifies *causal* relationships between variables	Tests the theory by investigating statistical relationships to verify their existence in larger populations	Tests and develops complex relationships between variables to suggest new theory

population verification of theory; and new empirical relationship exploration for theory development. Consequently, all these methods are extremely important for the complete development of theory-building in operations management.

8. The need for similarity in theory-building research procedures

This section suggests that general research procedures should apply to all research methodologies if theory-building is to be important for operations management. If all research methodologies follow common definitions, common domains, common relationships, and common predictions, the integration across the six different sub-categories of theory-building methodologies would be more likely.

Table 4 provides an outline of the procedures for the different theory-building methodologies. All six methodologies have the same four stages in the procedure, since these stages map directly to the definition of theory. In the first stage, it is important that the conceptual definitions used in the analytical methodologies are the same definitions that are operationalized in the empirical methods. It does not seem logical for analytical methods to use conceptual definitions for their mathematical convenience if these definitions have no hope of ever being operationalized for measurement in empirical studies. Therefore, analytical researchers should give measurement guidelines to empirical researchers. Conversely, it is important for empirical methods not to redefine an existing conceptual definition for the sake of measurability. From the theory-building perspective, ideally, all six methodologies would use the same conceptual definitions for the same concept so that analytical and empirical evidence are building on the same theory. Therefore, for theory-building, the first step in theory-building research should be identical regardless of the sub-category of research pursued.

In the second stage of theory-building research, the domain of the theory is defined. In this stage, all three types of analytical methodologies should specify where their theories apply. Generally, these methodologies have broader domains than empirical methodologies. Yet a common challenge to the analytical methodologies is a cautious domain specification to identify exactly where the theory is to be applied. Without this specification, it is assumed that all domains are included, which causes these theories to be questioned (or refuted) by a single empirical observation from any domain. For empirical methodologies, the domain usually is carefully specified by the research design (empirical experimental design), sampling method (empirical statistical methodology), or by the specific case studied (empirical case study methodology). From the perspective of theory-building, ideally, the domains suggested by the analytical methods would be investigated by the empirical methods to verify relationships to raise the abstraction level.

In the third stage of theory-building research procedures, relationships are suggested for theory-building. In all three of the analytical methodologies, logic is used to determine the relationships. The analytical methodologies use fundamental laws to deduce derived laws to suggest which relationships are logically compatible (internally consistent) with each other. Empirical methodologies should test analytically developed relationships before offering new relationships since new relationships may not be internally consistent. Empirical methodologies should exhibit great care before offering new relationships since some relationships may have already been identified using analytical methods. If an empirical study fails to examine the analytically derived relationships and the study results are contrary to the analytical relationships, the empirical results are considered artifacts, since the ‘power of deduction rules’ (Hunt, 1991).

The last stage of theory-building research is the prediction and verification. ‘Good’ theory-building suggests that operations management theory should offer internally consistent predictions in the ‘real world’ empirical realm. This suggestion means that analytical methodologies should use bridge laws as part of their methodology to provide the means to test their theories’ empirical predictions in the ‘real’ world. This suggestion is not to be interpreted as mathematical models are not needed in operations management, but rather, these models should offer some hope of improving operations management as it is practiced. To develop practical models, it seems that offering ‘how’ these results will be tested in the

Table 4
The theory-building procedure for different theory-building types of research

Theory-building procedure	Analytical			Empirical		
	Conceptual	Mathematical	Statistical	Experimental design	Statistical sampling	Case studies
Definitions development	Conceptual definitions. However, many times, new definitions are offered.	Conceptual definitions from the literature.	Conceptual definitions from the literature.	Conceptual definitions. However, frequently the method may require new, more measurable concepts.	Conceptual definitions from the literature. However, many times, new constructs are developed to represent the theoretical concept.	Conceptual Definitions from the literature. However, many times, new relationships require new definitions.
Domain limitations	Logically developed	Mathematically developed	Mathematically and statistically developed	Experimental design with a narrow controlled domain.	From analytical statistical models or developed experimentally	Developed from cases studied.
Relationship (model) building	Usually, relationships are logically developed.	Usually, relationships are mathematically developed without stochastic error terms.	Usually, other studies are used to develop mathematical statistical models with error terms.	Usually developed with limited relationships between variables.	Usually using other statistical studies' suggested theories.	The combining of the relationships discovered from the case.
Theory predictions: evidence of prediction	Usually, predictions come from logical analyses. Empirical evidence comes from case studies.	Mathematically deduced predictions. Examples from mathematical calculations or simulated results	Mathematically and logically derived predictions. Uses empirical evidence from other studies.	Prediction from the experimental design. Statistical significance of the tests	Prediction from other studies. Results from sample's statistical tests for significance.	Supported by case studies

empirical realm is a valuable aid for ‘real world’ empirical testing.

In summary, for each stage of the theory-building procedure, there is a need for integration of the results from other research methods to raise the abstraction level of operations management theory. To be a theory-building study, each research project should not ignore results from other methodologies, since ignoring other studies lowers the abstraction level and reduces the significance of any findings.

9. The current state of theory-building methodologies in operations management

One question that should be answered is: how are the different research methods in operation management being utilized? Table 5 illustrates the current state of theory-building in operations management over 5 years (1991–1995 inclusively). The top eight academic and practitioner journals in operations management provide data for investigating the current state of operations management research (see Barman et al. (1993)). These journals are the International Journal of Operations and Production Management (IJOPM); Decision Sciences (DS); Journal of Operations Management (JOM); Management Science (MS); Production and Operations Management Journal (POM); International Journal of Production Research (IJPR); Harvard Business Review (HBR); and Production Inventory Management (PIM). One caveat is needed before any further discussion: this article did not investigate the relative degree of theory-building of each article but rather concentrated its efforts on classifying each article into the sub-categories of research methodologies.

There are two deliberative ‘judgment calls’ in these classifications. The ‘first judgment call’ is made when classifying each article in terms of whether its primary contribution is to the operations management literature, or to another academic field. This study uses a broad perspective to include a broad spectrum of topics from the traditional list of operations management topics (see Hahn et al. (1982)). Of the 2817 articles reviewed, only 2002 are classified as operations management. Almost all of the articles in the operations management journals (International Journal of Operations and Production Management (IJOPM), Journal of Operations Management (JOM), Production and Operations Management (POM), International Journal of Production Research (IJPR), and Production Inventory Management(PIM) are designated as contributions to the operations management literature (It is not surprising to operations management academics that Decision Sciences (43%), Management Science (35%), and the Harvard Business Review (17%) all have significantly less than 50% of the articles related to traditional POM topics. See Table A.1 in Appendix A.).

The ‘second judgment call’ is the sorting of articles into the six major categories. This sorting required considerable judgment since studies frequently use multiple research methodologies. In this study, the classifying procedure concentrated on the *predominant methodology used* in the study. For example, the productivity article (not included since pre-1991) of Hayes and Clark (1989) would be considered a cross between empirical and analytical research since it uses statistical samples along with in-depth case studies to logically derive a conceptual productivity theory. However, when considered more closely, the article’s primary methodology uses sev-

Table 5
Overall classification of the articles by research sub-category

Classification	1991	1992	1993	1994	1995	Total
Analytical conceptual	103 (27.39%)	110 (26.63%)	100 (23.58%)	114 (27.01%)	63 (17.17%)	**490 (24.48%)**
Analytical mathematical	221 (58.78%)	214 (51.82%)	246 (58.02%)	218 (51.66%)	206 (56.13%)	**1105 (55.19%)**
Analytical statistical	2 (0.53%)	5 (1.21%)	5 (1.18%)	3 (0.71%)	7 (1.91%)	**22 (1.10%)**
Empirical experimental	2 (0.53%)	4 (0.97%)	5 (1.18%)	2 (0.47%)	2 (0.54%)	**15 (0.75%)**
Empirical statistical	28 (7.45%)	41 (9.93%)	34 (8.02%)	45 (10.66%)	56 (15.26%)	**204 (10.19%)**
Empirical case study	20 (5.32%)	39 (9.44%)	34 (8.02%)	40 (9.48%)	33 (8.99%)	**166 (8.29%)**
Total						**2002**

eral case studies over time to develop the factory productivity theory. Consequently, it would have been classified in the empirical case study sub-category.

Classifying articles into the six research sub-categories addresses the question of: ''How are the articles distributed into research sub-categories over the last 5 years?'' Using Table 5, it is apparent that over the last 5 years over 55% of all operations management's published articles use the analytical mathematical method (see Table A.2 in Appendix A for the breakdown by journal). The next most popular methodology is analytical conceptual, followed by empirical statistical, and empirical case study. Both empirical experimental and analytical statistical are not popular methodologies in the operations management literature. For the academic journals, five of the journals (Management Science, International Journal of Production Research, Decision Sciences, Production Operations Management Journal and the Journal of Operations Management) had a majority of articles following the analytical mathematical methodology. Only the International Journal of Operations and Production Management has its largest percent of articles in any other methodology (analytical conceptual was the most popular). As one might expect, the broader based journal (Harvard Business Review) and the practitioner-oriented journal (Production and Inventory Management) had a higher percentage of their articles classified into analytical conceptual and empirical case studies. However, the fact remains, that in the respected academic journals of operations management, the predominant research methodology is analytical mathematical.

One important implication of these results is that empirical experimental design and analytical statistical methodologies are not being used to verify causality and integrate theory. The implications of this result leads to two important constraints for theory-building in operations management. First, because experimental design is not being used, there is a lack of studying causality in the empirical world which hinders the development of verified relationships that are assumed to be true. This lack of verified relationships means that all relationships must be tested through empirical testing, causing statistical models to be too complex for meaningful investigation (Hunt, 1991). Second, because analytical statistical methodology is under-researched, there is a lack of an integrated internally consistent theory across all methodologies used in operations management. Since analytical statistical methods' basic purpose is to develop integrated models across all research methods, it provides a critical linkage to elevate theory to higher-level abstraction levels. In short, the evidence presented here suggests that operations management is not fully utilizing all research methods to verify old relationships and to build integrated theory.

10. How this article follows the criteria for 'good' theory-building

This article developed a definition of theory from the academic literature. The definition refined the conditions for what constitutes a theory. Next, it limited the domain to 'good' theory using the traditional virtues of 'good' theory. Third, it suggested a procedure for all types of theory-building research which fulfills the definition of theory. Additionally, this study classified research into six categories and demonstrated that all sub-categories of theory-building research in operations management use the same stages. Finally, it presented evidence that the six predominant types of theory-building research in operations management are not all equally distributed. This unequal distribution led to the conclusion that research on causality and theory integration are hindering the elevation of the abstraction level of operations management theory.

11. Guidelines for theory-builders and conclusions

The purpose of this section is to offer theory-building guidelines for theory-builders. It is important for the academic field to follow the formal definition of theory for theory-building research and incorporate these elements: definitions, domain, relationships, and predictions. Consequently, the following is a set of questions to provide guidelines for theory-builders.

(1) For definitions: are the terms used in the study standard (artificial) terms agreed upon by academics? Are the terms unique? Are new terms care-

fully differentiated from the standard terms in existence (conservation criteria)? Do all the new and standard terms avoid 'concept stretching'? Do all the terms used carefully exclude other concepts? Are difficulties with measurement of definitions highlighted for future researchers?

(2) For domains: are the specific conditions of when and where the data were gathered carefully enumerated? Are conditions for when and where the results apply prudently articulated? Is the domain of the results wide enough to be of value for researchers using other research methodologies?

(3) For relationship and model building: are all the variables listed in the estimates necessary (parsimony criteria)? Do all the variables used specifically state how and why they are related (or unrelated) to each and every other variable in the study (internal consistency criteria)? Is the model as small as possible to explain all the results (parsimony criteria)? Do the relationships offer new areas for research (fecundity criteria)? Are all the relationships specified before the data is gathered and before the relationships are estimated (internal consistency and empirical riskiness criteria)? Does the model elevate the level of abstraction by adding variables and inter-relating theories (abstraction criteria)?

(4) For theory prediction: does the model used make specific predictions which could likely be falsified (empirical riskiness criteria)? Do the theory predictions prohibit some events from happening? Does the theory discuss specifically how the theory is to be used and tested in the external empirical world (i.e., specifying bridge laws which are needed for empirical riskiness)?

Although all these guidelines are necessary for all theory-building articles, there are four specific concerns for the full development of operations management theory. First, there is a need for cautious attention to the measurement of the defined terms, since without exacting definitions, all theory is tenuous at best. Second, the specific domain of when and where the theory is to be applied is needed for the theory to be empirically tested. Third, the relationship-building stage is most important for empirical researchers since all variables should specifically state whether they are related or unrelated to each and every other variable. Consequently, the 'how' and 'why' each variable is related other variables assures that the empirical models are internally consistent. On the other hand, the fourth criteria, prediction, is most important for analytical methodologies, since these methodologies should offer how the theory can be tested and refuted in the external empirical world (empirical riskiness virtue).

A final word about the above theory-building guidelines. These guidelines are to be used as suggestions rather than hard and fast rules which can never be violated. Each guideline offered is based upon 'good' theory's virtues. Since 'good' theory's virtues are weighed against each other, judgment is necessary to determine the relative importance of each virtue (and each guideline). Still, theory-builders can use the above guidelines to increase their research's significance by 'good' theory-building and raising their research's abstraction level. It was in this hope that this article was written.

Acknowledgements

The author wishes to thank Professor R. Kenneth Teas for his suggestions, comments, insights and lengthy discussions on this study. The author also wishes to thank numerous academics for their insights and discussion that inspired many of the thoughts presented in this study. In particular, the author wishes to thank Richard Van Iten (Iowa State University), Paul Swamidass (Auburn University), and Jack Meredith (Wake Forest University) for the helpful discussions. Next, the author wishes to thank the three anonymous referees and the editors of this issue. Last, the author wishes to thank Michelle Johnson, Matthew Hanson, and Vicki Bouillon for their help with the editing. Naturally, any errors, omissions, or oversights are entirely the responsibility of the author.

Appendix A

A.1. Table A1: The classification of articles in leading operations management journals for production / operations management topics

Year	IJOPM POM	Non-POM	DS POM	Non-POM	JOM POM	Non-POM	MS POM	Non-POM	IJPR POM	Non-POM	POM POM	Non-POM	HBR POM	Non-POM	PIMJ POM	Non-	Total
1991	62	6	29	45	19	3	35	41	161	1			7	57	63	0	
1992	64	6	33	52			40	72	175		28	0	10	55	63	0	
1993	65	5	21	21	28	0	48	71	172	1	16	0	11	50	63	0	
1994	78	5	19	19	22	1	34	81	172		19	0	15	47	63	0	
1995	50	5	15	20	14	0	42	99	153	1	24	0	9	51	60	0	
Total	**319**	**27**	**117**	**157**	**83**	**4**	**199**	**364**	**833**	**3**	**87**	**0**	**52**	**260**	**312**	**0**	**815***
Percent	**92.20%**	**7.80%**	**42.70%**	**57.30%**	**95.40%**	**4.60%**	**35.35%**	**64.65%**	**99.64%**	**0.36%**	**100.00%**	**0.00%**	**16.67%**	**83.33%**	**100.00%**	**0.00%**	**28.93%***

Non-POM articles overall and non-POM percentage of all those reviewed.

A.2. Table A.2: Overall classification of all articles

Journal	Research type	1991	1992	1993	1994	1995	Total	percent
International Journal of Operations and Production Management (IJOPM)	Analytical conceptual	20	20	23	32	15	110	34.48%
	Analytical mathematical	27	16	26	21	19	109	34.17%
	Analytical statistical	1	0	0	0	1	2	0.63%
	Empirical experimental	0	0	0	1	1	2	0.63%
	Empirical statistical	9	12	10	15	8	54	16.93%
	Empirical case study	5	16	6	9	6	42	13.17%
Decision Sciences (DS)	Analytical conceptual	2	7	2	2	1	14	11.97%
	Analytical mathematical	27	20	15	12	5	79	67.52%
	Analytical statistical	0	2	0	1	2	5	4.27%
	Empirical experimental	0	1	1	1	0	3	2.56%
	Empirical statistical	0	2	3	3	6	14	11.97%
	Empirical case study	0	1	0	0	1	2	1.71%
Journal of Operations Management (JOM)	Analytical conceptual	4		7	7	0	18	21.69%
	Analytical mathematical	9		17	11	8	45	54.22%
	Analytical statistical	0		0	0	1	1	1.20%
	Empirical experimental	0		0	0	0	0	0.00%
	Empirical statistical	4		3	4	5	16	19.28%
	Empirical case study	2		1	0	0	3	3.61%
Management Science (MS)	Analytical conceptual	2	5	9	4	4	24	12.06%
	Analytical mathematical	31	32	35	26	30	154	78.97%
	Analytical statistical	0	2	2	1	3	8	9.30%
	Empirical experimental	0	0	0	0	1	1	1.25%
	Empirical statistical	2	1	2	3	4	12	15.19%
	Empirical case study	0	0	0	0	0	0	0.00%

Production and Operations Management (POM)	Analytical conceptual		9	4	3	4	20	22.99%
	Analytical mathematical		12	11	10	12	45	51.72%
	Analytical statistical		0	1	1	0	2	2.30%
	Empirical experimental		0	0	0	0	0	0.00%
	Empirical statistical		6	0	3	7	16	18.39%
	Empirical case study		1	0	2	1	4	4.60%
International Journal of Production Research (IJPR)	Analytical conceptual	40	38	30	36	19	163	19.57%
	Analytical mathematical	113	123	131	127	119	613	73.59%
	Analytical statistical	1	1	1	0	0	3	0.36%
	Empirical experimental	2	3	3	0	0	8	0.96%
	Empirical statistical	4	8	2	5	11	30	3.60%
	Empirical case study	1	2	5	4	4	16	1.92%
Harvard Business Review (HBR)	Analytical conceptual	3	3	4	7	5	22	42.31%
	Analytical mathematical	0	0	0	0	0	0	0.00%
	Analytical statistical	0	0	0	0	0	0	0.00%
	Empirical experimental	0	0	0	0	0	0	0.00%
	Empirical statistical	2	0	1	0	1	4	7.69%
	Empirical case study	2	7	6	8	3	26	50.00%
Production and Inventory Management (PIM)	Analytical conceptual	32	28	21	23	15	119	38.4%
	Analytical mathematical	14	11	11	11	13	60	19.23%
	Analytical statistical	0	0	1	0	0	1	0.32%
	Empirical experimental	0	0	1	0	0	1	0.32%
	Empirical statistical	7	12	13	12	14	58	18.54%
	Empirical case study	10	12	16	17	18	73	23.40%
Total							2002	

References

Bacharach, S.B., 1989. Organizational theories: some criteria for evaluation. Acad. Manage. Rev. 14 (4), 496–515.

Barman, S., Tersine, R.J., Buckley, M.R., 1993. An empirical assessment of the perceived relevance and quality of POM journals by academicians. J. Operations Manage.

Bluedorn and Evered, 1960.

Bunge, M., 1967. Scientific Research 1: The Search For System. Springer-Verlag, New York.

Churchman, C.W. 1961. Prediction and Optimal Decisions. Prentice-Hall, Englewood Cliffs, NJ.

Crawford, C.M., 1987. New Products Management. Irwin, Homewood, IL.

Demeyer and Arnoud, 1989.

Eisenhardt, K.M., 1989. Building theories from case study research. Acad. Manage. Rev. 14 (4), 532–550.

Ettlie, J., 1995. Product-process development in manufacturing. Manage. Sci. 41 (7), 1224–1237.

Hahn, C., Pinto, P., Mabert, V., Biggs, J., 1982. A profile of undergraduate programs in production and inventory management. Prod. Inventory Manage. J., Third Quarter, 48–66.

Hayes and Clark, 1989.

Hunt, S.D., 1991. Modern Marketing Theory: Critical Issues in the Philosophy of Marketing Science. Southwestern Publishing, Cincinnati, OH.

Kekre, S., Srinivasan, K., 1990. Broader product line: a necessity to achieve success. Manage. Sci. 36 (10), 1216–1231.

Kim, Y., Lee, J., 1993. Manufacturing strategies and production systems: an integrated framework. J. Operations Manage. 11 (1), 3–15.

Klemschmidt, E.J., Cooper, R.G., 1991. The impact of product innovation on performance. J. Prod. Innovation Manage. 8, 240–251.

Kuhn, T.S., 1980. In: Klempe, E.D., Hollinger, R., Kline, A.D. (Eds.), Theory Choice, Introductory Readings in the Philosophy of Science. Prometheus Books, Buffalo, New York.

Lindblom, C.E., 1987. Alternative to validity: some thoughts suggested by Campbell's guidelines. Knowledge: Creation, Diffusion, Utilization 8, 509–520.

Madique, M., Hayes, R., 1984. The art of high-technology management. Sloan Manage. Rev. 25 (2), Winter.

Marucheck, A., Pannesi, R., Anderson, C., 1990. An exploratory study of the manufacturing strategy process in practice. J. Operations Manage. 9 (1), 101–123.

Meredith, J.R., Raturi, A., Amoako-Gyampah, K., Kaplan, B., 1989. Alternative research paradigms in operations. J. Operations Manage. 8 (4), 297–326.

Miller and Roth, 1995.

Moorthy, K.S., 1993. Theoretical modeling in marketing. J. Marketing 57, 92–106.

Osigweh, Chimizie, A.B., 1989. Concept fallibility in organizational science. Acad. Manage. Rev. 14 (4), 579–594.

Poole, 1989.

Poole, M.S., Van De Ven, A.H., 1989. Using paradox to build management and organizational theories. Acad. Manage. Rev. 14 (4), 562–578.

Popper, S.K., 1957. Philosophy of science: a personal report. In: Mace, C.A. (Ed.), British Philosophy in Mid-Century.

Quine, W.V., Ullian, J.S., 1980. Hypothesis. In: Klempe, E.D., Hollinger, R., David Kline, A. (Eds.), Introductory Readings in the Philosophy of Science. Prometheus Books, Buffalo, New York.

Reynolds, P.D., 1971. A Primer In Theory Construction, Bobbs-Merrill Educational Publishing, IN.

Robinson, A., 1990. The Shingo System: Modern Approaches to Manufacturing Improvement. Productivity Press, Cambridge.

Sax, 1968.

Shubik, M., 1987. What is an application and when is theory a waste of time?. Manage. Sci. 33 (12), 1511–1522.

Sutherland, 1976.

Swamidass, P.M., 1986. Manufacturing strategy: its assessment and practice. J. Operations Manage. 6 (4), 471–484.

Teas, R.K., Palan, K.M., 1997. The realms of scientific meaning framework for constructing theoretically meaningful nominal definitions of marketing concepts. J. Marketing, (forthcoming).

Van de Ven, A.H., 1989. Nothing is quite so practical as a good theory. Acad. Manage. Rev. 14 (4), 486–489.

Wacker, J.G., 1987. The complementary nature of manufacturing goals by their relationship to throughput time: a theory of internal variability of production systems. J. Operations Manage. 7 (2), 91–106.

Wacker, J.G., 1996. A theoretical model of manufacturing lead times and their relationship to a manufacturing goal hierarchy. Decision Sci. 27 (3), 483–617.

Westbrook, R., 1987. Time to forget 'Just-in-Time'?: observations on a visit to Japan. Int. J. Prod. Operations Manage. 8 (4), 5–21.

Whetten, D.A., 1989. What constitutes a theoretical contribution?. Acad. Manage. Rev. 14 (4), 490–495.