Diverse businesses, such as garbage collection, retail banking, and management consulting are often tied together under the heading of “services”, based on little more than a perception that they are intangible and do not manufacture anything. Such definitions inadequately identify managerial and operational implications common among, and unique to, services. We present a “Unified Services Theory” (UST) to clearly delineate service processes from non-service processes and to identify key commonalities across seemingly disparate service businesses. The UST defines a service production process as one that relies on customer inputs; customers act as suppliers for all service processes.

Conversely, non-services (such as make-to-stock manufacturing) rely on customer selection of outputs, payment for outputs, and occasional feedback, but production is not dependent upon inputs from individual customers. The UST reveals principles that are common to the wide range of services and provides a unifying foundation for various theories and models of service operations, such as the traditional “characteristics of services” and Customer Contact Theory. The UST has significant operational corollaries pertaining to capacity and demand management, service quality, services strategy, and so forth. The UST provides a common reference point to which services management researchers can anchor future theory-building and theory-testing research.

Key words: services management theory; service operations; business process models

Submissions and Acceptance: Received March 2005; revision received July 2005; accepted September 2005 by Michael Pinedo.

1. Introduction

One significant challenge faced by researchers interested in services management is that there is no single, comprehensive, and consistently used unifying structure that defines what services are and what they are not (Lovelock and Gummesson 2004; Nie and Kellogg 1999; Vargo and Lusch 2004). This deficiency makes it difficult to achieve coherence within the field and creates a barrier to discovering the managerial and operational implications (Cook, Goh, and Chung 1999; Roth and Menor 2003).

The problem of discriminating a “service” from other kind of process has been widespread and has affected many. Fortune magazine gave up trying to differentiate services firms from manufacturers when, after many years of publishing both the Fortune Industrial 500 and the Fortune Service 500, it collapsed both into a single list (Eiben and Davis 1995). Castells and Aoyoma (1994) laid out the problem this way (citations in original):

“The notion of ‘services’ is often considered at best ambiguous, at worst misleading (Gershuny and Miles 1983; Daniels 1993). In employment, it has been used as a residual notion embracing all that is not agriculture, mining, construction, utilities, or manufacturing. Thus, the category of services includes activities of all kinds, with roots in various social structures and productive systems. The only feature common to these service activities is what they are not (Castells 1976; Stanback 1979; Cohen and Zysman 1987; Katz 1988; Daniels 1993).”

This perspective of seeing services as a disjointed “residual”—leftovers when all other sectors are accounted for—is peculiar. In most developed nations,
that “residual” accounts for the majority of both employment and GDP and is larger than all other sectors combined. This “residual” perspective has been perpetuated by the way governments have classified economic activities (Schmenner 1995). In fact, these governmental classifications continue to shift with the political wind. In a 2004 U.S. corporate tax bill, the tax code definition of “manufacturer” was expanded to include businesses like plumbers, architects, civil engineers, NASCAR venues, and movie rental. Obviously, there is little scientific value in such an arbitrary and politically motivated definition.

Academics have also offered various and evolving definitions of service industries (Cook et al. 1999). For example, Ammer and Ammer (1984, p. 421) define them as “an industry that produces services rather than goods.” Gonçalves (1998, p. 1) asserts that “…a service business is one in which the perceived value of the offering to the buyer is determined more by the service rendered than the product offered.” Harvey simply defines a service as “a result that customers want” (1998, p. 583).

Some academic definitions have focused on the characteristics of services. Pearce (1981, p. 390) observed that services “…are sometimes referred to as intangible goods; one of their characteristics being that in general, they are ‘consumed’ at the point of production.” Bannock, Baxter, and Reese (1982, p. 372) said services are “…consumer or producer goods which are mainly intangible and often consumed at the same time they are produced … Service industries are usually labor-intensive.” Operations management textbooks commonly define services as intangible products. Karmarkar and Pitbladdo (1995, p. 397) stated that distinguishing characteristics of services include “…intangibility of service output, the lack of inventories, the difficulty of portability, and complexity in definition and measurement … and often involve joint production between the buyer and the supplier.” Harvey (1998, p. 596) states, “Customer contact and intangibility are the two most important distinguishing features of services.”

Pine and Gilmore (1998, p. 12) define a service as “a set of intangible activities carried out on [the customer’s] behalf.” In the bestselling introductory marketing textbook, Kotler (2006, p. 402) defines a service in this way: “A service is any act or performance that one party can offer to another that is essentially intangible and does not result in the ownership of anything.” (Kotler might thus not consider a restaurant to be a service.)

Such definitions can provide insights into important issues of services. However, we might question whether something is a service because it possesses such characteristics, or if it has those characteristics because it is a service. For example, software development results in a product that is intangible (computer code) and the production process is very labor intensive (computer programmers), but the output can indeed be inventoried and used or sold later. Therefore, the characteristics of intangibility and labor intensity do not inherently define something as a service. In other words, we could conclude that if something is a service, its outputs might tend to be intangible and labor intensive, but not the other way around.

Other definitions of services focus on the production process. Levitt (1972) describes a service as being “invariably and undeviatingly personal, as something performed by individuals for other individuals.” Violating this definition would be all automated services or services rendered on inanimate objects, such as an unattended car wash (Thomas 1978). Hill (1977) defines a service as “…a change in the condition of a person, or of a good belonging to some economic unit, which is brought about as the result of the activity of some other economic unit…” Some have defined a service as a product that is a process (Henkoff 1994; Shostack 1987). Chase (1978) sought to define services based on the amount of “customer contact,” which he defines as the physical presence of the customer in the system. Froehle and Roth (2004) pointed out the limitation of this definition in that it does not consider situations where the customer is not corporeally involved in the production environment, but is receiving service just the same (e.g., telephone support). Murdick et al. (1990) define services as “economic activities that produce time, place, form, or psychological utilities,” to which Riddle (1985, p. 12) adds “while bringing about a change in or for the recipient of the service.”

There is a lot of truth (and some confusion) in these statements from the literature, but they still do not sufficiently answer the fundamental question about why we are justified in studying disparate industries under the single heading of “services.” How can we rationalize simultaneously encompassing business processes from health care and garbage collection, consulting and ski resorts, airlines and pawn shops, pet grooming and architecture firms, universities and butcher shops? A single commonality that comprises all services, in effect defining them categorically in a managerially relevant way, is a principal contribution of this article.

The next section will discuss theory building and present why a unifying theory is so valuable. The subsequent section provides a formal statement of the Unified Services Theory and requisite definitions of the concepts upon which it is based. Then, a section reconciles the Unified Services Theory with other popular prescriptive and descriptive services concepts. A section shows how the Unified Services Theory clarifies supply chain issues for services. The penultimate section applies the Unified Services Theory to several
key operations concepts, such as capacity and quality, demonstrating its usefulness as an analytical mechanism. The final section offers concluding thoughts and presents some future research topics motivated by the Unified Services Theory.

2. Value of Theories and Paradigms
The basis for any scientific discipline is theory (McMullin 1993). Theories can be, and should be, much more than vague suppositions about the unknown, but provide order and meaning to things that are well-studied and known to be true (Stafleu 1987, p. 15). Theories are valuable in explaining observations and data (Dictionary.com 2005, “theory”). Even the process of direct observation relies upon theory for interpretation and characterization (Cartwright 1993).

Closely related to theories are paradigms, which the esteemed scientific philosopher Thomas Kuhn characterizes as assumptions shared by members of a given discipline (Kuhn 1970). Paradigms form the basis for advancement: “A paradigm shapes the formulation of theoretical generalizations, focuses data gathering, and influences the selection of research procedures and projects” (Lovelock and Gummesson 2004, p. 21).

The discipline of services management cannot exist and advance without theories and paradigms. Yet it is not enough to have theories; they must be “good” theories (McMullin 1993, p. 63). There are many perspectives about what constitutes a “good” theory. The motivational theorist Locke (2005) identified six characteristics of a good inductive theory:

1. It is based on observations and data,
2. It defines concepts in a way that differentiates from other concepts,
3. It integrates concepts and resolves apparent contradictions,
4. It identifies causal relationships,
5. It typically takes time to develop, and
6. It is open ended, allowing for extensions and re-applications.

We will revisit this list in the final section. For now, it is important to consider the third item, integration. Locke (2004) states that theories that integrate other theories are especially useful, which is the idea behind a unifying theory. This is consistent with Feigl’s observation that “The aim of scientific explanation throughout the ages has been unification, that is, the comprehending of a maximum of facts and regularities in terms of a minimum of theoretical concepts and assumptions” (Feigl 1970, p. 12; Kitcher 1988, p. 167).

To progress, the field of services management needs to be unified (Roth and Menor 2003). Service researchers need a shared foundation and a common language upon which various areas of study and practice can build (Cook et al. 1999). Service marketing perspectives need to be reconciled with service operations perspectives (Karmarkar 1996). We need to understand how professional services, menial-labor services, and capital-intensive services can all possibly fall under the same discipline. The distinctions between front-office and back-office services have been well documented (Chase 1978; Chase 1981), but the commonality also needs to be established.

Perhaps the most well-known example of a unifying theory is the Unified Field Theory, wherein Einstein proposed that electromagnetism and gravity were different manifestations of a single phenomenon (Tonne-lat 1966; Yurth 1998). Similarly, the goal of the Unified Services Theory is to provide a “good” theory that unifies the various branches of services management to a common trunk.

3. Building a Unified Services Theory
The foundational core of the Unified Services Theory is as follows: “With service processes, the customer provides significant inputs into the production process.” This statement simultaneously defines what services are and what makes them services. The presence of customer inputs is a necessary and sufficient condition to define a production process as a service process. Any production process that is a service process must have customer inputs (they are necessary), and the presence of customer inputs establishes a production process as a service process (their presence is sufficient). Why does it make sense to define a service process this way? Answering that question and demonstrating to the reader that the presence of customer inputs is indeed a necessary and sufficient condition to identify a service process in a managerially relevant way are the objectives of this section.

Some production processes do not involve customer inputs. Because of that, these processes behave differently, and should be managed differently, than service processes. Such processes are often referred to as manufacturing or extractive processes. In manufacturing, groups of customers might contribute ideas to the design of the product, but individual customers’ roles are limited to the selection and consumption of the outputs, not the contribution of inputs specific to production for that particular customer. And, just as many products are more accurately thought of as bundles of tangible and intangible benefits, production systems tend to be mixtures of service and non-service processes (Boyer and Metters 2004). To fully represent the Unified Services Theory, each of the components of the above maxim—inputs, customer, and the production process—must be defined.

3.1. Inputs
An input is “something put into a system or expended in its operation to achieve output or a result” (Dictio-
nary.com 2005). Inputs to a production process may come from various sources and suppliers, and may include “materials, personnel, capital, utilities, and information” (Gaither and Frazier 1999, p. 15). “Customer inputs” are defined as inputs that are provided by customers (Davis and Heineke 2005, p. 7). The literature has typically identified three general types of customer inputs (Wemmerlöv 1990): the customer’s self, his belongings or other tangible objects, and information.

Customer-self inputs are common, for example, in services involving co-production (i.e., the employment of customer labor in the process) and in services involving the physical presence of the customer. Common examples are health care offices, buffet restaurants, and taxi services. These service providers can prepare for production, but they cannot execute the actual service process until necessary customer-self inputs are present. Some services, such as movie theaters, require the customer’s physical presence, although the customer’s mind is primarily being acted upon (movie theaters can show movies even without customers present, but that would hardly be called “production.”) Lovelock separates customer-self inputs into those involving the body and those involving the mind (Lovelock 1992; Lovelock 1996).

Tangible belongings (or property) and physical objects make up another type of input a customer can provide to the service process. One’s car is an essential input into the automobile repair service process, and one’s clothing is a necessary input to the dry cleaning service process. Providing tangible inputs often allows the service process to proceed even without customers present, but that would hardly be called “production.”

Customer-provided information is a third type of input to the service process. The tax return preparation process requires that customers provide financial information as process inputs. Without that information input, the service production process cannot begin, even though the tax preparer can prepare for production (such as by studying the tax code or duplicating blank forms).

An important distinction must be made between customer-provided inputs, which are central to the definition of a service, and two types of customer involvement that do not inherently define a production process as a service process. First, as mentioned previously, groups of customers or potential customers might be involved in market research to guide product design. In this case, customers provide opinions about general products destined for future production, with possible consumption by them or by others. Such general feedback from customers can be valuable, but that feedback is not an essential input to the production process for a specific customer. Both manufacturers and service providers conduct this type of market research, and general customer feedback is not exclusive to either domain. So, this type of customer involvement differs from specific customer-provided “inputs” in that customer feedback does not, by itself, define a production process as a service.

The second type of customer involvement that we differentiate from customer inputs is “selecting and consuming the output.” Customers select and consume the output from processes of all types, not just service processes. Although selection processes vary, the act of selection does not define service or nonservice processes. For example, a customer may select a candy bar to buy. That customer act of selecting the product does not make the candy manufacturing process a service process because the customer provided no input into the production of that particular candy bar. However, the vending machine from which he buys the candy bar is indeed executing a service process since, among other things, the customer must enter his selection information into the keypad for the process to proceed. As we will discuss, manufacturing an item and selling that item are two different types of processes when the former does not involve customer inputs but the latter does.

3.2. Customers

The concept of “customer” is so fundamental as to almost elude definition. The American Marketing Association (AMA) defines “customer” as “the actual or prospective purchaser of products or services” (Bennett 1995, p. 73). Such a definition is adequate for most situations, but imprecise in others. For example, who is the customer of a textbook publisher? Is it the students who pick up and purchase the textbooks at the bookstore? Or is it the instructors who require students to purchase specific textbooks for their courses?

The AMA’s definition of “consumer” is a bit broader: “Traditionally, the ultimate user or consumer of goods, ideas, and services. However, the term also is used to imply the buyer or decision maker as well as the ultimate consumer. A mother buying cereal for consumption by a small child is often called the consumer although she may not be the ultimate user” (Bennett 1995). This allows for the separation of purchasing decision-makers and output users.

Consistent with that thought, we define customers as the individuals or entities who determine whether or not the service provider shall be compensated for production (Sampson 2001, p. 28). This captures the concept of purchase decision-making ability but also allows for complex decisions made by more than one entity. For example, who makes the purchasing decision in the
throughput, which is the rate at which the system is consistent with Goldratt’s (1992, p. 60) concept of puts in a way that is valued by customers. This view steps. We consider “production” to be modifying in-
a production process. A “process” is a sequence of
radio broadcasts.

tomers who advertisers desire to be satisfied with the
listeners are important because they are indirect cus-

"produce" (i.e., generate revenue) without advertise-
commercial radio stations are compensated for broadcast-
radio broadcast” process is
ing by advertisers, not by listeners. Advertisers pro-
commercial radio stations are compensated for broadcast-
are customers in that they directly determine whether or
the Federal Aviation Administration’s regulations. So the
radio station shall be compensated for production. This relationship with the
faithful customer of the airline’s safety and other processes.

In some cases, identifying the customer is neither
easy nor obvious, but it is essential to understanding
service nature of the process. For example, con-
sider a commercial broadcast radio station. Its listen-
ers listen to the broadcast but provide nothing in
in determining whether or not the service provider shall be compensated for production, so each, to some
degree, is a customer.

Indeed, all customers are not on equal standing in
the purchase decision. Some may be paying custom-
ners, others, however, may be “indirect customers”—
entities that paying customers may require to be satis-
"prepare for production" by
hear the broadcast from the top of a skyscraper: al-
ames are service businesses,” even though it would
be naïve to assume that all business processes have similar management characteristics and concerns.
When we observe a business process from a thousand
feet up, at the company or industry level, it is difficult

3.3. Production Process
In the Unified Services Theory, the unit of analysis is
a production process. A “process” is a sequence of
steps. We consider “production” to be modifying in-
puts in a way that is valued by customers. This view
is consistent with Goldratt’s (1992, p. 60) concept of
“throughput, [which] is the rate at which the system
generates money through sales.” Goldratt emphasizes
that “if you produce something, but don’t sell it, it’s
not throughput.” We therefore limit “production” to
refer to only productive activities that contribute to
throughput or sales, which is an elevated specification of the term “production.”

There are often process activities that do not directly
lead to sales, but which are nonetheless necessary for
the organization to execute. We define these “support-
ing processes” as processes that normally must be
accomplished for ongoing production activities to oc-
cur. Few would argue against the importance of clean-
ing the equipment at a restaurant, but few customers
would pay to watch the equipment be cleaned, much
less pay to participate in the cleaning process. Other
supporting processes at a restaurant include hiring
employees, ordering food ingredients, and creating
menus. They are important activities even though cus-
tomers do not directly pay for them. More impor-
tantly, we observe that those supporting processes can
be accomplished independently from the customers,
without customer inputs.

Nearly all processes are composed of sub-processes,
smaller sets of tightly linked action steps. One may
decompose a large complicated process into smaller
processes for more detailed study (Sampson 2001, p. 38).
An overall production system may have some
processes that are replete with customer inputs, and
others that are devoid of customer inputs. Studying
service production processes is akin to studying news-
print from a traditional offset press. From a distance it
appears that even black-and-white pictures have
many shades of gray. However, finer observation re-
vels that gray pictures are composed of black dots of
various sizes. Metaphorically, the black dots represent
customer inputs. Some processes have significant cus-
tomer inputs (large dots). Other processes have minor
customer inputs (small dots). Still others have no cus-
tomer inputs whatsoever, and thus are not consist-
tently subject to the unique managerial issues that
concern service processes.

In the past, companies and even entire industries
have been categorized as services (e.g., Kellogg and
Nie 1995; Schmenner 1986), yet such can lead to con-
voluted analysis with unclear or incorrect conclusions
(Verma and Young 2000). It is like trying to study a
sidewalk painting from the top of a skyscraper: al-
though one can make general observations about color
and form, it is difficult to appreciate the details. It is no
wonder some casual observers conclude that “all busi-
nesses are service businesses,” even though it would
be naïve to assume that all business processes have
similar management characteristics and concerns.
When we observe a business process from a thousand
to accurately differentiate between service processes and non-service processes.

For this reason, the Unified Services Theory considers the unit of analysis to be the process, versus the firm or industry. The firm, the total production system itself, is a mixture of service and non-service production processes. It is important to remember that processes involving customer inputs are fundamentally and managerially different from non-service processes. Firms differ widely in composition of service and non-service processes, and therefore differ widely in relevant managerial principles. Further, service processes differ widely in terms of customer inputs, even with processes in similar businesses. Full-service investment brokers and discount brokers share many things in common but differ in intensity of customer inputs. However, processes involving even slight customer inputs differ dramatically from processes devoid of customer inputs.

### 3.4. The Unified Services Theory (UST)

The Unified Services Theory is formally stated as follows (Sampson 2001, p. 16):

> “With service processes, the customer provides significant inputs into the production process. With manufacturing processes, groups of customers may contribute ideas to the design of the product, but individual customers’ only participation is to select and consume the output. All managerial themes unique to services are founded in this distinction.”

The Unified Services Theory is not unique in its foundations but rather in its implications. In fact, we are eager to point out that the founding concepts of the Unified Services Theory have been well documented in research literature. The unique contribution of the Unified Services Theory is in codifying and applying the customer-input concept. The statement, “All other managerial themes unique to services are founded in this distinction” might be said in different ways:

- Service processes are distinguished from non-service processes only by the presence of customer inputs and implications thereof.
- For those familiar with business management in general, understanding those additional issues unique to managing services requires only understanding the implications of customer inputs.
- Customer inputs are the root cause of the unique issues and challenges of services management.

In the next section, we will illustrate this universality of the Unified Services Theory by showing how it explains some widely held services management concepts and frameworks. The purpose is both to show how prior approaches justify the Unified Services Theory and how the Unified Services Theory helps shed insights into those approaches.

### 4. Reconciliation with Prior Service Perspectives

The Unified Services Theory is primarily a generalization of various theories of services that have been previously offered. In that sense, the UST indeed “stands on the shoulders of giants.” In this section, we hope to demonstrate that the UST is easily reconciled with, and is a generalization of, some of the most commonly cited perspectives on services.

#### 4.1. Characteristics of Services

We previously mentioned how researchers have attempted to establish what services are by identifying key characteristics and/or attributes that services share. Over the years, thought has generally focused on five supposed characteristics: intangibility, heterogeneity, simultaneity (inseparability), perishability, and customer participation. Each of these five characteristics is either completely explainable by the presence of customer inputs (the core element of the Unified Services Theory) or is refutable as being characteristic only of service processes (Lovelock 1992; Sampson 2001; Vargo and Lusch 2004).

**Heterogeneity** is the observation that individual units of service production tend to be unique, especially when compared with non-service processes such as mass production (Nie and Kellogg 1999). Accommodating that variability is one of the biggest challenges for service operations. But what is the root cause of this variability? The Unified Services Theory contends that heterogeneity in processing and outcome is primarily caused by heterogeneity in process inputs, specifically customer inputs (Sampson 2001, p. 108). One conclusion we can make is that a viable way to reduce variability in service processes is to reduce variability in customer inputs. For example, researchers have discussed the “service factory” concept in which a service provider limits the range of service offerings by limiting the range of customer inputs (Chase 1978). McDonald’s limits the range of customer-information inputs by providing a limited menu. Federal Express limits the range of customer-belonging inputs by providing standardized containers like the “FedEx envelope.” Shouldice Hospital limits the range of customer-self inputs by limiting surgical offerings to a specific type of hernia. Some service processes and strategies lend themselves to limiting customer inputs but many do not. Instead, services may need to accommodate variability in customer inputs (Shingo 1986).

**Simultaneity**, also called inseparability, refers to the observation that services are generally produced and consumed at the same time (as compared with non-services’ tradition of producing well in advance of demand and consumption). With service processes, significant portions of production cannot begin until
after customer inputs have been presented by the customer, which corresponds with demand (Sampson 2001, p. 52). Because the customer is involved, some aspects of consumption may begin during the production process. We call this concept “inadvertent JIT (Just-in-Time),” implying that JIT production in services is a necessity, not a choice (Karmarkar 1996; Sampson 2001, p. 310). Contrast this with non-service processes where inventory is used to separate production and consumption and where JIT is an option (usually only attempted by manufacturers who are not faint of heart).

**Perishability** alludes to the time-sensitive nature of a service provider’s capacity to produce the service. It is not the service product itself (e.g., the dental work or the tax return) that is perishable (Sampson 2001, p. 82), but rather the capacity (e.g., the empty dentist’s chair or the accountant’s time). Service capacity is time-perishable because significant elements of production cannot begin before customer inputs are present (e.g., the patient’s teeth or the client’s tax information) (Sampson 2001, p. 60).

Related to perishability is the mistaken belief that service processes are unable to produce inventory. We understand inventory to be items of production that are available before needed. Inventory is the result of a mismatch between production and demand. With service processes, we cannot produce before demand due to the reliance on customer inputs. However, there can still be delays in the system if customer inputs arrive in excess of available capacity. In such cases, the customer inputs are in “inventory” either until sufficient server capacity becomes available or until the customer decides to withdraw his or her inputs from the process (Sampson 2001, p. 90). This “customer inventory” is commonly called a “queue” or a “waiting line,” and it experiences a “holding cost” much more time sensitive than traditional manufacturing inventories (Garnett, Mandelbaum, and Reiman 2002). Manufacturing inventory holding costs are calculated over weeks or months, whereas service customer inventory holding costs are typically measured in minutes or hours.

**Intangibility** is the characteristic that has perhaps been most commonly attributed to services. Despite being frequently cited in texts and research literature, it is unfounded and has been discredited in recent literature (Laroche, Bergeron, and Goutaland 2001; Lovelock and Gummesson 2004; Vargo and Lusch 2004). The Unified Services Theory sheds light on the intangibility misconception by considering the tangible items provided by customers (Sampson 2001, p. 100).

“Intangibility” means that something is “incapable of being perceived by the senses” (Dictionary.com 2005). Service processes are capable of being perceived, and service outcomes are often as tangible, or more tangible, than manufacturing outputs (Laroche et al. 2001, p. 27). For example, how might one justify an assertion that a dental root canal procedure eludes perception, especially by the sense of touch? The dental office uses “facilitating goods,” such as needles and mouthwash, as well as a very tangible “supporting facility” (the building), which are common components of the “service package” (Fitzsimmons and Fitzsimmons 2006, p. 19; Porter 1980).

Consider the auto-painting process. If it takes place in an auto factory in Detroit, it is considered a tangible manufacturing process. It is no less tangible if it takes place in your home town and involves your car. In fact, it may be more tangible if it involves your car: you may actually observe the process and you may be more sensitive to tangible process output details (since it is your car that was painted). The confusion about tangibility is rooted in assumptions about ownership of production items (Kotler and Keller 2006, p. 402), with the tangibility of customer-owned inputs apparently being ignored. All production processes have tangible and intangible elements. Intangibility neither defines nor uniquely characterizes services.

In summary, these four “characteristics” of services are not defining characteristics but, when they occur, are simply “symptoms” of the customer inputs. Just as symptoms of the flu can be caused by things besides a flu virus, these symptoms of services can be caused by things other than the presence of customer inputs and, therefore, can also show up in non-services. We must beware of “manufacturing in sheep’s clothing,” which are non-service processes that exhibit customer-input symptoms (Sampson 2001, p. 154). For example, fuel production is a manufacturing process, but electricity is so expensive to store that it is produced “just in time” with demand (not because producing electricity requires customer inputs). Pre-packaged software and recorded music may be intangible products, but they are generally produced without customer inputs through non-service processes (Laroche et al. 2001, p. 27). (Although intangibility is not unique to services, it is a common cause of confusion.)
4.2. Customer Contact Theory and Related Concepts

A valuable tenet of service operations is that the potential efficiency of a service process depends largely on the amount of customer contact involved (Chase 1981; Chase and Tansik 1983). Chase has characterized customer contact as “...the physical presence of the customer in the system” (Chase 1978, p. 138) when “...the customer is in direct contact with the service facility” (Chase 1981, p. 700). These descriptions suggest there is some inherent personal physicality involved in customer contact, and this is indeed how customer contact has been traditionally viewed and measured (e.g., Kellogg and Chase 1995). The Unified Services Theory defines a service process as one involving customer inputs, which inputs may include, among other things, the customer’s physical self. In that light, the traditional notion of customer contact (i.e., physical presence) becomes a special case, or subset, of the Unified Services Theory.

We might extend the traditional definition of customer contact to include service situations where the customer and service provider are not physically co-located during the contact episode (e.g., call centers and distance learning). The Unified Services Theory tells us the inefficiencies arising from face-to-face contact might also exist in these situations due to the presence of significant customer inputs. As mentioned, Lovelock (1992, 1996, p. 29) differentiates inputs associated with the customer’s physical self from those associated with the customer’s mind/consciousness. These latter “virtual self” inputs can be as heterogeneous as the customers themselves. This suggests that removing the customer’s physical body from the service process does not eliminate customer contact, nor does it necessarily reduce the variability of the customer’s inputs (Froehle and Roth 2004). The UST suggests that we would expect operational efficiency implications to arise both in face-to-face and “virtual” customer contact, and that differences arise from the nature and variety of the customer inputs.

A concept rooted in customer contact theory is front-office/back-office differentiation (Metters et al. 2006; Shostack 1984; Shostack 1987). Traditionally, the front-office has been described as “where the customers are,” while the back-office is where processes not directly involving the customer are carried out. Decoupling is the popular practice of removing low-custom contact components of front-office work, standardizing them, and relocating them into the back-office in hopes of increased efficiency (Metters and Vargas 2000). Decoupling is only possible because the nature of the customer inputs is different. For the high-variance, high-contact activities that have to remain in the front-office, the customer inputs generally include the customer’s physical or conscious presence. In contrast, those tasks that can be decoupled, standardized, and moved to the back-office generally rely only on non-self customer inputs (e.g., tangible property, goods, and information) or don’t rely on customer inputs at all. Decoupling essentially splits out service processes that require customer “self” inputs from those service processes that do not.

Custom manufacturing, mass customization, and delayed differentiation are similar concepts from goods production that relate to this idea. In custom manufacturing, the customers trigger production by the presentation of their custom specifications. These customer-information inputs define custom manufacturing as a service process. Indeed, custom manufacturing has many management challenges found in service operations and not present in typical make-to-stock mass production manufacturing settings (Sampson 2001, p. 142). For example, Dell “custom manufactures” a computer when a customer places an order for it, but this simplified description confuses two discrete stages in the production process. First, before the order is placed, Dell executes a non-service process that readsies a variety of components that customers may want, resulting in standard sub-assemblies; this is done without any individual customer’s inputs. Then, once a customer has placed an order (i.e., provided an essential customer input), the service processes associated with completing the assembly and delivering the finished good can be executed. This production approach decouples processes that depend on customer inputs from processes that can be done independently from customer inputs, allowing higher efficiency. This consistent reliance on information about customer inputs to efficiently organize production further substantiates the Unified Service Theory.

4.3. The UST at the Core of Service Classification Schemes

Over the years, various service classification schemes have been proposed as means for generating managerial insights. We observe that all useful schemes have at their core a classification of customer inputs or the treatment of customer inputs.

For example, The Service Process Matrix (Schmenner 1986) was developed to help classify different kinds of services so that operational insights might be gleaned and better management decisions could be made. The variables the Service Process Matrix uses to classify services are (a) service customization and customer interaction, and (b) labor intensity. Customization only occurs as a response to comply with unique customer inputs including specifications. Customer interaction is a primary means for providing customer and labor inputs, including preference specifications. However, customization can occur independently of interaction, such as an audit firm providing a unique
audit process based on the financial records of a client firm. Labor intensity is an antecedent of customer inputs, in that variance in customer inputs often prohibits automation. In other words, the differentiating power of the Service Process Matrix originates from considering customer inputs of differing types.

A second example is the framework proposed by Wemmerlöv (1990), which categorized service processes based on three criteria: technologization (i.e., "rigid" versus "fluid" processes), the degree and nature of customer contact, and the object of the service (what is being acted upon—goods, information, or people). The first dimension examines the technology necessary to accommodate variability in the process. As has been discussed, much of the variability in service operations is customer-induced, so a desire to accommodate variability in customer inputs necessitates more fluid processes and appropriate flexible technologies. The second dimension—the degree and nature of customer contact—directly relates to those customer inputs associated with interaction (i.e., physical presence, indirect technology-mediated communication, or no interactions). The object of the service action, which is Wemmerlöv’s third differentiating dimension, ties directly to the three types of customer inputs identified earlier. Clearly, customer inputs lie at the heart of this valuable typology.

In another classification framework, Kellogg and Nie (1995) suggested the “Service Process/Service Package Matrix,” for deriving operational decisions based on the nature of the service package in a way similar to the well-established Product-Process Matrix (Hayes and Wheelwright 1979). The key element their framework relies upon to define the service process dimension is “customer influence,” which is described as, “the customer, by his/her presence, interaction and/or participation, in some way influences the service process” (p. 325). These authors later empirically found that customer influence is “the most important characteristic in affecting OM strategies and decisions” (Nie and Kellogg 1999, p. 349). This clearly depicts the reliance upon customer inputs, which wholly determine a customer’s participation, presence, and overall impact on the service system.

As a final example, Napoleon and Gaimon (2004) propose a model that categorizes information technology worker systems as either simple or complex, with differences arising from the standardization (or, conversely, unpredictability) of the decision-making elements and the job requirements for service workers. That model provides useful insights, but does not examine the factors that determine whether a system is standardized or unpredictable; we posit those factors inherently involve customer inputs. The Unified Services Theory would therefore suggest that there is additional value in determining how the nature of customer inputs causes system requirements to be simple or complex.

5. Bidirectional Service Supply Chains

The Unified Services Theory also aids our understanding of the relevance of supply chain concepts to service businesses, an important contribution (Roth and Menor 2003). Various publications on supply chain management claim that the principles are applicable to services, yet the vast majority of the stated examples are from manufacturing supply chains (Ellram, Tate, and Billington 2004). Service examples, such as retail and logistics, involve the processing of physical goods and are quite obviously part of manufacturing supply chains (Bozarth and Handfield 2006, p. 4; Wisner, Leong, and Tan 2005, p. 6). But applying supply chain management to “pure services” like consulting and psychology can seem forced and unclear. Two clear exceptions are Anderson and Morrice (2000) and Akkermans and Vos (2003), which studied upstream amplification in mortgage and telecoms service process/supply chains respectively.

Traditional manufacturing supply chains are relatively linear, as depicted in Figure 1 (although obviously more expansive than this simplified figure). Normal production flows are typically unidirectional, with items flowing from suppliers to customers (the solid-line arrows in Figure 1). Reverse logistics aside, the primary thing flowing upstream is information: orders, feedback, and payment information (not depicted in Figure 1). Although customer focus groups may contribute opinions about general product design (the dashed-line arrows in Figure 1), customers are otherwise completely downstream from the manufacturing stages of the supply chain.

The Unified Services Theory indicates that service supply chains have an expanded role for customers,
namely as suppliers of inputs to service provider processes. This means that service processes are bidirectional, as shown in Figure 2. Prior to being recipients of process outputs, customers are suppliers of process inputs. The arrows of Figure 2 represent flows of goods, information, and/or human (e.g., labor) inputs involved in the service production process.

Further, Figure 2 shows that service providers may employ other service providers to supply necessary services. For example, a patient may provide a blood sample (“self” input) to a physician who is the service provider. In turn, the physician may deliver the blood sample (tangible good) to a lab service for analysis. The lab returns the results (information) to the physician who provides diagnosis (information) to the patient. Such a relationship has been referred to as a two-level bidirectional service supply chain (Fitzsimmons and Fitzsimmons 2006, p. 483; Sampson 2000).

Bidirectional service supply chains differ from primarily unidirectional manufacturing supply chains in many respects. First, service supply chains tend to be hubs, not chains (Fitzsimmons and Fitzsimmons 2006, p. 482–486). Service managers are concerned with product flows going in both directions. For two-level bidirectional supply chains, the service provider acts as an agent for the customer when dealing with outside suppliers. A car repair shop may outsource the rebuilding of engines to a machine shop. The car repair shop acts as a hub and represents the customer’s needs to the machine shop.

Second, bidirectional supply chains tend to be short (Sampson 2000, p. 355), as depicted in Figure 2. Service providers tend to interact directly with consumers of the services without the buffer of distributors and retailers. Advantages of short supply chains include reduced complexity and easier sharing of information.

Third, service providers cannot treat suppliers who are customers the same way they would treat suppliers who are not customers. Supply base reduction for customer-suppliers is precarious, since it means reducing the number of customers. Traditional criteria for supplier selection are of little value for customer-suppliers. Managing the quality of supplied inputs is no easy task; even well-defined input quality specifications are likely to be violated by customers. Supplier certification for customer-suppliers can be difficult to enforce.

Fourth, service operations need to be robust enough to handle the stochastic nature of customer-supplied inputs (Sampson 2001, p. 355). This includes being capable of dealing with random arrivals, inconsistent specifications, and varying input quality. Manufacturers can use inventory as a tool for dealing with random demand, but inventories of customer inputs can be very expensive, so other tools of capacity and demand management are used instead. These and other operational issues, such as quality, will be discussed in the next section.

6. Operational Implications of the Unified Services Theory

The Unified Services Theory is, in part, a descriptive theory in that it describes why a service process is a service process. It is also a prescriptive theory in that it guides decision-making; a great many insights about managing services emerge from applying the theory to various operational problems. This section provides examples of this by examining some critical service operations issues through the lens of the Unified Services Theory.

6.1. Capacity and Demand Management

The prior section alluded to the way that the use of capacity in service situations is analogous to the use of inventory in manufacturing situations. A significant proportion of the operational issues encountered in service management are concerned with capacity (e.g., adding, deploying, and scheduling) and accommodating variable demand (Hur, Mabert, and Brethauer 2004; Jack and Powers 2004). A variety of tactics have been devised to address these issues. On the demand management side, reservation systems keep customers from inserting their inputs into the production process until the service provider is likely to be ready for them, thereby minimizing pre-production waits. Price incentives and promotion of off-peak demand work similarly in that they motivate customers to time their inputs so as to correspond with a period of lower demand. For labor-intensive service operations, capacity management is often aided by asking the customer to take on some of the production (i.e., the customer provides labor inputs, such as self-check-out). The use of self-service is increasing as the capabilities of automated service systems improve (Prahalad and Ramaswamy 2004, p. 98).

In services, many of the variable inputs are provided by the customers, who do not charge the service provider for those inputs. This tends to keep variable costs down, but tends to make fixed costs a larger proportion of overall expenses for the service provider. In fact, service labor costs in many situations tend to be largely fixed for two reasons: first, service providers need to schedule labor based on forecasted, rather than actual, customer demand; and second, time-perishable capacity (discussed above) forces the service provider to pay scheduled workers even though there are no customers currently needing service. This high “operating leverage” implies that many service operations will be much more cost-competitive if they have higher utilization levels (Sampson 2001, p. 240) or, alternately, increase their volume.
flexibility (Jack and Powers 2004). Some service operations tend to have higher variable service costs, such as custom home building and retail, so utilization does not have as great an impact on process cost performance in those situations.

We previously discussed how JIT production is mandatory for most services, with extremely time-sensitive “holding costs” for customer inventory. When demand exceeds capacity (even temporarily) and customer inventory results, service providers can shift demand, as just discussed, or they might alternatively focus on reducing the holding cost. When customers are in inventory, the holding cost is primarily psychological, and might be reduced by managing customers’ perceptions (Sampson 2001, p. 319). Maister (1985) pioneered a customer-inventory management approach that has come to be known as the “psychology of queuing.” The idea is to make the wait more acceptable, such as by providing entertainment and a sense of fairness, thus reducing perceived wait times. Manufacturing managers generally do not have the luxury of using psychological means to reduce inventory holding costs.

6.2. Quality Management
Managing quality is difficult in service processes for a variety of reasons, many of which stem from customer inputs. As Fitzsimmons and Fitzsimmons (2006, p. 483) state, “Customer inputs can be incomplete (e.g., tax documents), unprepared (e.g., students), or have unrealistic expectations (e.g., cancer patient). This lack of consistency in the quality of customer-supplied inputs represents a challenge for the service provider to deliver on promises when inputs are questionable.” They emphasize the importance of effective communication in establishing customer expectations before service delivery in order to avoid misunderstandings.

Further, customer-executed process steps can have higher variability because customer-provided labor varies more dramatically in terms of training and experience, and customers make different types of mistakes (Stewart and Chase 1999). Customer-provided property/good inputs can vary dramatically (e.g., a car repair shop sees great variability in terms of makes, models, type of repair, etc.). Customer-provided information inputs are particularly subject to variability, including effects of varying communication skills. Customer moods vary, which can result in assessments of quality that differ from one mood to the next, even with consistent service delivery. Finally, in-process quality correction is difficult because the customer, unlike a physical piece of raw material in a manufacturing process, experiences the service production process as it happens and updates his perception of quality as production occurs (Sampson 2001, p. 368). Service customers remember process failures even if they are corrected.

Even defining service quality can be precarious due to customer-imposed expectations about the process and outcome. Customers of manufacturing processes may also impose expectations about the process outcome (i.e., the product). However, manufacturing customers generally accept that product specifications are dictated by some engineer somewhere who does not have access to each customer’s specific needs and wants. As a result, customers are quite tolerant of manufactured goods. Further, manufacturers manage suppliers to ensure consistently high-quality process inputs. Customers are key suppliers of service processes, providing both process inputs and opinions about how the process and outcome should be. These opinions can make specifying quality like buying cloth based on the ancient “cubit” (distance between a person’s elbow and the tip of the longest finger): The specification varies depending on whose cubit is being used at the time.

Service quality is not only difficult to specify, but it is equally difficult to measure (Prahalad and Krishnan 1999; Sampson 2001, p. 388). Assessing the outcome of service production often requires measuring the customer’s perception of the service experience, which is far from precise (Hays and Hill 2001). The Unified Services Theory suggests that since customer inputs define service processes, we should also attempt to measure the quality of the customer inputs involved in the service. After all, if a customer provides inappropriate or inadequate inputs to service process, that could explain a great deal of his/her dissatisfaction with the service outcome. Anyone who has attempted to teach unprepared students knows this scenario. Interestingly, one of the most commonly-used service quality instruments, SERVQUAL (Parasuraman, Zeithaml, and Berry 1988), does not explicitly include any customer inputs in its measurements.

A related issue that can be problematic is the specification of service productivity. We consider productivity to be the amount of output for a given amount of input, such as labor. Customer inputs confound productivity measurement by introducing heterogeneity, implying that outputs cannot be simply counted (Sampson 2001, p. 400). For example, one attorney wins nine out of 10 court cases and another wins only five out of 10 cases—who is more productive? The first attorney may only take on easy or trivial cases, and the second may only take on high-risk or precedent-setting cases with high potential pay-outs or social implications. Due to the heterogeneity imposed by heterogeneous customer inputs, it is common practice for attorneys to charge not according to some output measurement (such as success of court judgment), but
rather according to an input measure they have direct control over (i.e., billable labor hours).

6.3. Services Strategy

The Unified Services Theory suggests that service organizations can stand out from their competitors by focusing on certain, or different, customer inputs and/or by employing similar customer inputs in different ways. This idea applies to all three of Porter's (1980) generic strategies—cost leadership, focus, and differentiation—which are applicable to services (Fitzsimmons and Fitzsimmons 2006, p. 41; Sampson 2001, p. 176).

For a cost leadership strategy, the goal is reducing costs well below one's competitors. This is often based on low-cost inputs and efficient operating systems; the achievement of both is highly dependent on customer inputs. Many customer inputs can save the firm money (e.g., self-service labor), but not all can. Some customer inputs, such as unusual specifications, drive up process and outcome heterogeneity (even beyond what most customers may require) and increase other costs (such as extra production labor or customer support). Variability in customer inputs is the enemy of efficiency; it diminishes the advantages of learning through repetition, non-divergence, and economies of scale. Thus, if a company wishes to pursue a cost leadership strategy, it must assess the relative merits of each customer input to the production process in order to determine how and when to accept and/or require that input and how to specify/control that input’s variability.

A focus strategy is meeting the exclusive needs of a narrowly defined target customer base. A focus strategy can benefit from critically thinking about customer inputs: the goal should be to identify a set or sets of customers with specific and similar customer inputs and expectations and then design the service process around those inputs and expectations. For example, Shouldice Hospital has so refined this approach that it pre-screens potential customers using a very simple and low-cost process: the patient application survey (Fitzsimmons and Fitzsimmons 2006, p. 140). Before it permits a customer-patient into its primary production process, it determines whether or not he has the right inputs (e.g., type of hernia, overall physical condition, etc.). Once admitted, the customer is propelled through a production process designed to minimize complications and risk while maximizing healing and recuperation. A customer with mismatched inputs would likely cause significant operational problems due to needing special care and accommodation.

Finally, a firm adopting a differentiation strategy strives to provide a unique service or to provide a service in a way that is different from its competitors. By handling customer inputs differently than its competitors, it can generate service outcomes that are distinctive and perhaps hard to duplicate. However, the UST suggests that it can be more difficult to sustain differentiation in services than in manufacturing due to the difficulty in maintaining production trade secrets. With manufacturing, companies can keep secrets in factories by requiring all who enter to sign non-disclosure agreements. With services, competitors posing as customers can have free access to service facilities.

The distinction between a focus strategy and a differentiation strategy, which is often difficult to discern, is made easier by examining customer inputs. While a focus strategy targets a specific niche of customers with similar inputs and needs, a differentiation accommodates a potentially wider array of inputs and employs them in the production process in a unique way. For example, Amazon.com has differentiated itself through the use of customer preferences (customer-provided information inputs), seeking to customize its recommendations and customers’ experiences by tracking customer activity and buying patterns over time.

7. Conclusions and Research Opportunities

The Unified Services Theory states that all managerial issues unique to services stem from the fact that service processes involve customer inputs. Earlier we spoke about Locke’s attributes of a good inductive theory (2005). Examining the Unified Services Theory, we see that it meets those attributes:

1. It is based on a wide variety of research literature and industry observations,
2. It defines services and service concepts in a way that differentiates them from traditional manufacturing concepts,
3. It integrates prior models of services management under a common basis,
4. It shows the cause of various service phenomena (i.e., the requirement of customer inputs),
5. It is based on time-tested research (Although the theory was first formalized in the late nineties, it is founded in the earlier writings of Hill (1977), Chase (1978, 1981, 1983), Lovelock (1983), and others.),
6. It defines services in a way that is very open-ended in terms of implications and applications.

Despite being parsimonious and simple, the UST is not simplistic—its ramifications are significant and far-reaching. We can come to understand that the fundamental reason why service processes can be harder to manage than non-service production processes is the presence of customer inputs. The host of complications arising from this one defining aspect will continue to challenge practitioners and research-
ers alike. However, because we can point to customer inputs as the cause for the unique issues found in services, we are better equipped to begin investigating how to address each problem or take advantage of each opportunity.

Space limitations prevent us from providing more than a cursory review of applications and implications of the Unified Services Theory. Interested readers are directed to Sampson (2001), which is a more complete and detailed recitation on the UST and includes many practical examples and extended discussion on related topics.

There are many rich opportunities for future research involving the Unified Services Theory. One area we have looked at is how to measure customer inputs. We have discussed the problems associated with measuring service production, but similar and additional problems emerge when trying to measure customer inputs. First, different types of inputs are likely to need different measurement approaches; measuring customer-provided labor and measuring customer-provided information are unlikely to use the same metric. Second, can we develop reasonable objective measures (e.g., Soteriou and Chase 1998) or must we rely on subjective means for measuring customer inputs? Measuring inputs consistently from one service environment to another has been a persistent challenge (Metters, Frei, and Vargas 1999). As an example, Silvestro et al. (1992) proposes classifying services according to the number of customers processed per service unit per day. Categorizing and developing metrics for a wide range of service customer inputs would be beneficial to both practitioners and services researchers.

Various planning and analysis tools can also be studied from the perspective of the Unified Services Theory, and new approaches can be derived. For example, the service flowcharting technique known as “service blueprinting” (Shostack 1984; Shostack 1987) can be greatly enhanced by the Unified Services Theory perspective (Sampson 2001, p. 422). Answering important questions raised by the UST can be valuable preparation for blueprinting, such as “What are the key customer inputs that define each process as either a service or a non-service?” and “What inputs to the process are supplied by non-customers and could customer-supplied inputs be used instead, and vice versa?”

Automation and technology in services are increasingly important topics (Davis and Heineke 2003; Roth and Jackson 1995; Walley and Amin 1994) and can also be examined through the lens of the UST. Firms often introduce process technologies in order to produce more with more consistent quality at lower cost. These cost savings often come by allowing technology and customer labor to substitute for expensive paid labor. One problem is that, as yet, process technologies are not as adaptable to high-variance customer inputs as human labor is. Automating technologies in services can significantly damage customer satisfaction because they fail to recognize or accommodate the uniqueness of each customer’s inputs (i.e., they over-standardize) while not providing customers with obvious or adequate cost savings. An example might be automated voice response systems that never let customers talk to a human. Examining the interactions between aspects of customer inputs (e.g., quantity, type, and variety) and aspects of technology (e.g., ease-of-use, customizability, and responsiveness) offers a fertile area to which services researchers can add valuable and innovative insights.

The Unified Services Theory serves as a framework for the study of services management. However, it is hoped that an even greater contribution of the UST will be in providing coherence to the services management discipline, thereby acting as a foundation for future research on managing services.

References


