Incorporating Self-Serve Technology into Co-Production Designs

Ulrike Schultze (uschultz@mail.cox.smu.edu)
Anita D. Bhappu (abhappu@mail.cox.smu.edu)
Cox School of Business
Southern Methodist University
PO Box 750333
Dallas, TX 75275-0333

Abstract

Co-production, which is the direct involvement of customers in the design, delivery, and marketing of goods and services that they themselves consume, implies customer-firm collaboration. The nature of this collaboration, however, is highly dependent on the organization’s service design, which increasingly includes Internet-based self-serve technology (SST). While there is considerable research on service design, most of this research fails to consider the role and implications of information technology in co-production. In order to remedy this shortcoming, we build on prior research to develop a contingency theory of co-production design. Using cases of Internet-based SST (e.g., online airline reservations and recommendation-based sales systems), we highlight the unintended consequences of implementing SST in the different co-production designs. In this way, our research contributes to our understanding of information technology’s implications for co-production.

1 We would like to thank Wanda Orlikowski, Steve Postrel, and Ned Kock for their helpful comments on earlier versions of this paper.
Introduction

E-collaboration, which entails collaboration among individuals using electronic technologies to complete a common task (Kock, 2005), is prevalent in today’s service sector, where it frequently takes the form of customer co-production enabled by self-serve technology (SST). As the image of the customer as passive audience and consumer is being replaced by one of the customer as an active co-creator of value (Prahalad & Ramaswamy, 2000), organizations increasingly view their customers as resources that contribute both knowledge and labor to the production process (Larsson & Bowen, 1989). This form of collaboration (or co-laboring) is not only evident in business-to-business (B2B) service delivery, where organizations’ supply chains are becoming vertically integrated (e.g., Walmart shares daily sales information with Procter & Gamble), but also in business-to-customer (B2C) service environments where customers’ actions not only trigger but also complete a transaction (e.g., customers assembling their IKEA furniture themselves) (Normann & Ramirez, 1993).

Internet technologies (e.g., email, WWW, SST) present both opportunities and challenges for the design and delivery of services (Bitner, Ostrom & Meuter, 2002). For instance, in a recent study, Hogg et al. (2003) analyzed the impact of the Internet on the delivery of health care, a service environment typically associated with high information asymmetry. By providing access to information in a cost-effective way (Malone, Yates, & Benjamin, 1987), Internet technologies erode the information asymmetry that has traditionally characterized many customer-firm interfaces (Kulkarni, 2000). Hogg et al’s (2003) results highlight that the healthcare consultation process changed significantly as customers increasingly went online to learn about medical conditions and treatment
options. Whereas the consultation process used to be one in which the health care provider acted as the primary source of information, advice, and decision-making, in the Internet-enabled service environment, the health care provider acts as interpreter and evaluator of information that patients gather from multiple online sources. As patients became more active co-creators of their own treatment, there were changes in the way that patients and health care providers interacted and in the power dynamics of their relationship.

According to Malone et al (1987), networked technologies such as the Internet favor the creation of market structures, where exchanges are instantiated through arm’s-length relationships. So, are Hogg et al’s (2003) findings suggestive of a general shift in the nature of co-production and the demise of customer-provider relationships in Internet-enabled service environments? Or are there different conditions under which networked technologies, such as Internet-based SST, affect co-production and customer-provider relationships differently? In other words, what are the exogenous variables upon which effective co-production designs are contingent, and how are these contingencies impacted by Internet-based SST?

In this paper, we attempt to answer these questions. We commence by defining co-production as a multi-dimensional construct. Based on transaction cost economics and the literature on service design, we then identify the exogenous variables that constitute the contingencies of our framework for co-production design. Next, we describe and analyze three SST implementations, highlighting how SST is affecting customer co-production and the exogenous variables of co-production design. We also
identify some of the unintended consequences that each SST example generates. We close with suggestions for future research.

**CO-PRODUCTION**

The notion of co-production has a substantial history in the public sector. As city and state governments struggle with limited resources, they typically look to citizens to volunteer their time and to assist public-sector service providers with labor and information. In the public administration literature, co-production is thus defined as “direct citizen involvement in the design and delivery of city services with professional service agents” (Brudney & England, 1983), and “the degree of overlap between two sets of participants – regular producers and consumers” (Brudney & England, 1983). Co-production, therefore, implies collaboration (or co-laboring) between customers and service providers.

Co-production is generally associated with services rather than goods (Bowen & Jones, 1986; Larssen & Bowen, 1989). This is because services are intangible and need to be consumed at the time of production lest their value be lost (Mills & Margulies, 1980). Thus, the utility of a service depends on the customer who completes the service by consuming it (Ramirez, 1999). For instance, performing a play to an empty house or flying an empty plane is value lost. Also, the consumption of services typically relies on customers interacting with service providers. Indeed, the service production function is considered to be relational, that is, based on the interaction between the service provider and the consumer (Bell, 1976). Lastly, human services such as medical care, education,

---

2 In the strategy literature on co-production, the customer is regarded as supplier (Normann & Ramirez, 1993; Prahalad & Ramaswamy, 2000), while the organizational literature refers to the customers as partial employee (Mills & Morris, 1986). In the public administration literature, in turn, customers are referred to as consumer-producers (Brudney & England, 1983).
and psychiatric counseling, which seek to affect the personal transformation of the customer, require customers to actively participate in the transformation process (e.g., eating less to losing weight and studying to learn a new skill).

This does not mean, however, that co-production is confined to service industries (e.g., banking, retail, transportation, insurance, health care, and education). Instead, as Normann and Ramirez’s (1993) description of IKEA’s co-production strategy highlights, the distinction between goods and services tends to be equivocal and fluid (Gershuny, 1978). By unbundling the assembly service from the finished good, IKEA invites its customers to co-produce the value of their ( unassembled) furniture. In contrast, most furniture manufacturers subsume the service of assembly in the finished, tangible good. This highlights not only the ambiguity of the boundary between goods and services, but also that co-production is not limited to services.

**Figure 1: A Multi-Dimensional Definition of Co-Production**

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Elements</th>
</tr>
</thead>
</table>
| **Activity**                | Design  
Requesting Assistance | Delivery  
Participating in Production | Marketing  
Recommending and Referring |
| **Mode of Cooperation**     | Customer Motivation  
Voluntary & Compliant | Customer Participation  
Active & Passive |
| **Type of Interdependence** | Sequential  
Planned Coordination | Pooled  
Standardized Coordination | Reciprocal  
Mutual Adjustment |

Whitaker (1980) defines co-production in terms of three dimensions: (i) requesting assistance, (ii) cooperation, and (iii) mutual adjustment. This categorization scheme provides us with a starting point for developing a multi-dimensional definition of co-production. The different dimensions of co-production in our framework, Figure 1, are activity, mode of cooperation, and type of interdependence. The activity dimension
describes what customers do to co-produce, whereas the cooperation and interdependence dimensions respectively describe customers’ mode of cooperation and the nature of coordination between customers and providers.

**Activity**

Whitaker’s (1980) “requesting assistance” reflects the *design* activity. When service options are well defined and customers have a clear understanding of their needs, the act of requesting assistance merely triggers a service (e.g., requesting telephone service). However, when the information required to determine the features of products or services is highly equivocal, the task of requesting assistance becomes more complex and more difficult to structure. Customers have to articulate their problems, needs and expectations, and service providers have to match them to the organization’s service capabilities. Argote (1982) suggests that this latter form of requesting assistance is particularly prevalent in environments with high input uncertainty. An example of high input uncertainty (and high demand diversity) is a hospital emergency room where the need for medical treatment cannot be accurately forecast. Customers’ requests for assistance design not only the service at the time of request but also future service offerings. To the extent that service organizations treat customers’ requests for assistance as feedback on their extant services, requesting assistance shapes the design of both current and future goods and services.

**Delivery**, on the other hand, consists of the production of goods and services that customers consume. The degree of customer involvement in delivery varies depending on the performance ambiguity associated with the product being consumed (Bowen &

---

3 Input uncertainty and demand diversity (Larsson & Bowen, 1989) reflect different perspectives (operations and marketing respectively), but refer to the same concept, namely the qualitative and quantitative variability of the demands customers place on a service.
Incorporating Self-Serve Technology into Co-Production Designs

Jones, 1986; Siehl, Bowen, & Pearson, 1992). For instance, when it is difficult to assess service quality or attribute service outcomes to provider performance, services are characterized by high performance ambiguity. Examples of services with high performance ambiguity include fitness training and weight loss programs, where it is difficult to fault the trainer or program counselor if a client does not lose weight. This is because successful outcomes are also dependent on customers’ active participation in the production process (e.g., exercise and eat healthier foods).

The marketing activities that customers engage in remain largely invisible to providers and firms (Chervonnaya, 2003). Therefore, they are frequently omitted from definitions of co-production. Nevertheless, marketing is not only beneficial to the provider firm (e.g., word of mouth advertising), but also to the customer, especially if the product or service in question is subject to network effects (e.g., fax machines and instant messaging become more valuable to the individual when more people use it). Therefore, based on the different activities that customers engage in, we define co-production as the direct involvement of customers in the design, delivery, and marketing of goods and services that they themselves consume.

Mode of Cooperation

To cooperate means to work jointly or to unite to produce an effect. Whitaker (1980) uses community policing as an illustrations of cooperation. Since cooperation applies to all co-production activities, we treat it as a separate dimension in our framework. For instance, when customers and providers work jointly on defining the requirements of a good or service (especially under conditions where the customer information required is highly equivocal), they are cooperating with each other.
Based on Brudney and England’s (1983) work, we conceptualize cooperation as a mode, that is, the how and why of a customer’s participation in co-production. For instance, Brudney and England (1983) point out that, ideally, customers should cooperate voluntarily in co-production activities, but that there is a fine line between voluntary and compliant cooperation in the public sector. For instance, for effective sanitation, residents are expected to move their trash to the curb on a given day of the week. Failure to comply with these rules of co-production typically results in fines, suggesting that customer behavior is motivated by compliance rather than voluntary cooperation.

Brudney and England (1983) also distinguish between active and passive cooperation. Active cooperation implies that customers purposefully engage in design, delivery or marketing activities. Passive cooperation implies that customers refrain from disruptive behavior and thereby enable other producers to do their work. For instance, by obeying flight attendants’ orders and refraining from getting up when the seat belt signs are illuminated, passengers cooperate in the production of a safe flight.

Even though Brudney and England (1983) maintain that, in theory, co-production should be limited to voluntary and active cooperation, they acknowledge that, in practice, it is difficult to determine when behavior is motivated by cooperation or by compliance. Furthermore, passive compliance tends to be invisible (Chervonnaya, 2003), suggesting that firms are unlikely to consider it in their co-production design until, for instance, consumers actively engage in disruptive behavior or stop cooperating passively.

**Type of Interdependence**

Our third dimension of co-production encompasses Whitaker’s (1980) notion of mutual adjustment. Mutual adjustment occurs when both customer and provider base
their actions on their joint consideration of the issues. This is accomplished through considerable communication and the reciprocal modification of each party’s expectations and actions. In personal or human services, where the goal is to transform the customer (particularly his/her behavior), developing an understanding of the customer’s problems, needs and/or wishes goes beyond the design activity of the service. Instead, it requires mutual adjustment throughout the service delivery process, and it may even call for mutual transformation (e.g., the teacher learning from the student, and vice-versa).

According to Thompson (1967), mutual adjustment is the most appropriate coordination mechanism in situations where participants’ actions are reciprocally interdependent. For instance, a psychologist’s ability to help patients is dependent on the patient’s ability and willingness to trust him/her. Unless patients voluntarily share their hopes, fears and dreams with the psychologist, it is unlikely that the psychologist can adjust his/her treatment to their needs. In contrast, when customers and providers are sequentially interdependent, the most effective coordination mechanism is planning. For instance, organizations forecast what their stock levels or service capacities need to be in order to satisfy consumers’ demands. In situations of pooled interdependence, coordination is best achieved through standardization. For instance, collaborative filtering technologies, on which recommendation systems like the one used by Amazon.com are built, rely on a pool (or data warehouse) of customers’ ratings of books and music. To make these diverse ratings available for analysis, they have to be in a standardized format. We include all three forms of interdependence and their related coordination mechanisms in our definition of co-production as Larsson and Bowen (1989) highlight that all three are manifest in co-production.
In summary, our multi-dimensional definition of co-production suggests that customers are valuable resources in the design, delivery, and marketing of goods and services because they supply organizations with knowledge, skill, and labor. Depending on the specific combination of activity, mode of cooperation, and type of interdependence, co-production can take different forms. For instance, co-production can be as limited as a customer voluntarily requesting assistance to set up telephone service, which triggers a sequentially interdependent service delivery process that is relatively devoid of customer involvement. Alternatively, co-production can be as extensive as a cancer patient repeatedly interacting with his/her oncologist and voluntarily undergoing a series of radiation treatments that require the patient to actively take medications, to passively endure fatigue, and to keep the oncologist informed of symptoms after each treatment. Therefore, by specifying all three of the dimensions set forth in our framework, we can assess the degree of customer co-production in different service environments.

A Contingency Framework of Co-Production Designs

Similar to prior research on organizational design, we take a contingency approach to specifying the most effective forms of co-production in different service environments. While it is not always clear what the important design contingencies are and how they relate to organizational forms (Argote, 1982), there is considerable consensus that uncertainty is a contingency central to organizational design (Galbraith, 1973). In the context of service design, input uncertainty is a key contingency (Siehl et al., 1992) because it describes a firm’s lack of information about the overall composition of customer inputs. Input uncertainty “is expressed as a function of the number of
choices or alternatives in a given situation, e.g., patient conditions [encountered by a hospital emergency room], and of the probability of various alternatives occurring” (1982). Larsson and Bowen (1989) define input uncertainty as incomplete information about the what, when, where, and how of the customer’s input.

Input uncertainty can be expressed in terms of the amount and equivocality of information needed from the customer. We consider information equivocality the more crucial design contingency because organizations can reduce the amount of information they need by limiting their service offerings, thereby reducing input uncertainty through design. Equivocality describes whether information is of doubtful significance and whether it can be interpreted in different ways. For example, a request for something “hot” to eat could be interpreted as meaning “warm” or “spicy.” Therefore, information equivocality is closely related to the nature of the product or service and cannot be easily reduced through design.

Argote’s (1982) research highlights that flexible means of coordinating work are required under conditions of high input uncertainty. For instance, in a hospital emergency room, where both the diversity of conditions demanding treatment and information equivocality are high, medical staff not only have to carefully establish treatment needs, but also determine on-the-spot what procedures need to be performed. High input uncertainty thus increases coordination costs for service providers who have to dynamically coordinate their work in response to customers’ equivocal requests for assistance. A scripted repertoire of responses cannot be developed in such situations, and programmed coordination mechanisms such as rules and authority structures are not the most effective means of coordinating (Argote, 1982).
**Performance ambiguity** is another key design contingency highlighted in the service literature (Siehl et al., 1992; Bowen & Jones, 1986). As mentioned earlier, performance ambiguity refers to the difficulty of assessing service or product quality (Ouchi, 1980). It is related to the intangibility of the object being exchanged and its complexity, i.e., the difficulty of establishing the role and value of each component’s contribution to the quality of the good or service. For example, performance ambiguity is relatively high in education because it is difficult to assess students’ knowledge before the commencement and after the completion of an educational program. Furthermore, even if one could assess students’ learning, it would be difficult to attribute success or failure to the educational program due to the ambiguity of the cause-effect relationships in the learning process. Given the reciprocal interdependence in most human service settings, assessing the value of contributions made by individual resources is problematic (Bowen & Jones, 1986).

Performance ambiguity is a design contingency frequently applied in transaction cost economics (Ouchi, 1980), a theory concerned with the boundary of firms (Williamson, 1975). High performance ambiguity implies high transaction costs because the product or service being exchanged is not completely contractible, which makes it difficult to assess whether the exchange was equitable. For instance, when a service provider is highly dependent on a customer for the successful delivery of a service, it is difficult to establish the value contributed by the service provider (Ouchi, 1980). Therefore, parties to the exchange must actively monitor the service delivery process to ensure equity. Such monitoring implies high transaction costs.
When transaction costs are high, goods or services should be produced within the firm so that it can rely on its employment relations and, ultimately, fiat to manage incomplete contracts (Williamson, 1991). Where performance ambiguity is high, hierarchical organizational designs are the most effective. If performance ambiguity is low, however, exchange participants can establish complete and enforceable contracts to govern the exchange. Thus transaction costs are relatively low, and markets are the most effective way of managing the exchange.

The literature on service design (e.g., Bowen & Jones, 1986; Jones, 1987) relies on transaction cost economics to determine whether to locate customers inside or outside the firm. In this research, customers are referred to as “partial” (Mills & Morris, 1986) or “transient” (Namasivayam, 2003) employees, and service designs based on long-term customer-provider relationships are labeled “hierarchies” (Bowen & Jones, 1986; Jones, 1987). Given Williamson’s (1991) later work, however, organizational forms that are characterized by long-term customer-provider relationships should be called “hybrids” – or “networks” (Thompson, 2003) – rather than “hierarchies,” as they reside between the two discrete structural alternatives of markets and hierarchies. The label “hierarchy” should be reserved for organizational designs characterized by a legally recognized organizational boundary. Furthermore, the notion of customers becoming part of the firm is foreign to transaction cost economics because contracts with employees are very different from those with customers (Williamson, 1991).
Incorporating Self-Serve Technology into Co-Production Designs

Figure 2: Co-Production Design Framework

<table>
<thead>
<tr>
<th>Performance Ambiguity</th>
<th>Input Uncertainty</th>
<th>Quadrant I</th>
<th>Quadrant II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Low</td>
<td>Scripted Market</td>
<td>Scripted Relationship</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Consumer banking</td>
<td>- Buying IKEA furniture</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Booking airline tickets</td>
<td>- Weight-loss programs</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
<td>Personalized Market</td>
<td>Personalized Relationship</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Hospital emergency rooms</td>
<td>- Post-Graduate education</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Book recommendations</td>
<td>- Cancer treatment</td>
</tr>
</tbody>
</table>

In our contingency framework of co-production designs (see Figure 2), we call the two structural alternatives for exchange relations between providers and customers “markets” and “relationships.” Markets are most effective under conditions of low performance ambiguity, whereas relationships are most effective under conditions of high performance ambiguity. Similarly, we label the two methods of coordination co-production activities as “scripted” and “personalized.” Scripted modes of coordination are most effective under conditions of low input uncertainty, whereas personalized modes of coordination are most effective under conditions of high input uncertainty. The four resulting co-production designs (quadrants) are: Scripted Market, Scripted Relationship, Personalized Market, and Personalized Relationship.

Quadrant I – Scripted Market

The scripted market bears much resemblance to Mills and Margulies’ (1980) *maintenance interactive service organization* and Larsson and Bowen’s (1989) *sequential standardized service design*. The customer-firm interface is characterized by low information equivocality and programmed decisions on the part of service providers. Customer requests are limited to an established set of service options, leaving little room
for customers to influence service design. By standardizing service options and scripting the customer-firm interface, firms can program their service delivery process. For these same reasons, individual service providers are interchangeable, preventing customers and employees from forming interpersonal ties. Performance ambiguity is low due to the tangibility of the product and/or the low complexity of the service production process.

According to Mills and Margulies (1980), this form of co-production is particularly prevalent in consumer banking. Bank tellers, for example, typically handle a predetermined set of customer service requests, which are associated with standardized processes that require the bank tellers to make routine decisions. By scripting the customer-firm interface and programming service delivery processes, banks can offer low-cost (Larssen & Bowen, 1989) yet reliable (Mills & Margulies, 1980) banking services to customers. However, because the bank and its customers are sequentially interdependent, the customer must actively request assistance in order to initiate the delivery of standardized banking services.

**Quadrant II – Scripted Relationship**

Neither Mills and Margulies (1980) nor Larsson and Bowen (1989) identify a service design that closely resembles the scripted relationship. However, Bowen and Jones (1986) do identify an *impersonal hierarchy service design* that bears some resemblance to the scripted relationship. They suggest that standardized, hierarchical exchanges will only take place when organizations, such as monopolistic service providers (e.g., utility companies), can impose agreements unilaterally on their captive customers. Another example of an impersonal hierarchy is the penitentiary, where the process of rehabilitation is located within the prison’s organizational boundaries, and
inmates’ participation is coerced through rules and authority structures. Bowen and Jones (1986) highlight that – in the face of high performance ambiguity – customers would be willing to become partial employees in a highly structured production process only under such supra-normal conditions.

The main difference between our conceptualization of the scripted relationship and Bowen and Jones’ (1986) impersonal hierarchy is that we do not regard organizational forms, which require high customer participation throughout the delivery process, as hierarchies. In our framework, the customers remain outside of the organization’s boundaries even as they form and sustain close, long-term relationships with their service providers.

An example of a scripted relationship is the co-production design at IKEA (Normann & Ramirez, 1993) where customers purchase unassembled furniture, transport it themselves, and then assemble it at home. By voluntarily doing some of the work in an otherwise sequentially interdependent production process, customers are able to purchase furniture at IKEA for lower prices. Information equivocality in this setting is relatively low. Not only is furniture tangible and can be described in precise terms (e.g., size, type of wood), but customers can also see and touch the assembled furniture in the IKEA store before they purchase it. The tangibility of the product also implies that there are a finite number of ways in which the furniture can be assembled. Thus, the process of furniture assembly can be scripted and compiled into a set of instructions.

However, performance ambiguity is relatively high as IKEA’s service delivery process relies on customers to complete the production process. Thus, the respective contributions of and the causal relationships between resources (furniture pieces,
assembly instructions, customer) and outcomes (assembled furniture) are ambiguous. What if the assembled furniture wobbles or is misaligned? Who is responsible for this ostensible failure – the furniture designer or the customer? IKEA’s generous return policy seems to acknowledge this high performance ambiguity, as the organization relies on trust to manage the incomplete contract with its customers. For instance, IKEA seems to trust customers to comply with the furniture assembly instructions, and customers appear to trust IKEA to only sell furniture that they can indeed assemble on their own.

**Quadrant III – Personalized Market**

The personalized market bears much resemblance to Mills and Margulies’ (1980) *task interactive service organization* and Larsson and Bowen’s (1989) *sequential customized service design* where customers’ task specifications precede providers’ production of goods and services. The customer-firm interface is focused on task accomplishment based on customers’ personal specifications. In the personalized market, performance ambiguity is low due to the tangibility and/or low complexity of the good or service. However, information equivocality is high in the personalized market, implying that the identification of the customers’ requirements and expectations is challenging and that the decisions that service providers have to make are non-routine. Furthermore, the uniqueness of each customer’s request means that the coordination of production activities cannot be programmed, but must remain flexible and adaptable.

An example of a personalized market is a hospital emergency room where the range of customers’ medical needs is difficult to anticipate. Treatment activities are therefore defined and dynamically coordinated in response to an incoming patient’s symptoms and conditions. Diagnostic decisions are complex because health care
providers have to accurately evaluate test data in combination with equivocal information from the patient about his/her condition, which may also be difficult to elicit, especially if the patient is unconscious or in trauma. Therefore, input uncertainty is high. However, performance ambiguity is relatively low because stabilizing the patient’s condition is the responsibility of the medical staff on call and the patient’s participation is limited to requesting medical care and responding to questions about his/her condition.

**Quadrant IV – Personalized Relationship**

Personalized relationships bear close resemblance to Mills and Margulies’ (1980) *personal interactive service organization*, whose primary objective is to provide services that transform the individual customer through non-programmed coordination and monitoring. Examples include education and counseling. Personalized relationships also reflect characteristics of Larsson and Bowen’s (1989) *reciprocal service design*, which is marked by reciprocal interdependence and mutual adjustment between service provider and customer. Customer input is required throughout the delivery process to ensure that the final outcome (good or service) is aligned with customers’ needs and expectations.

Performance ambiguity in the personalized relationship is high. To manage this ambiguity, organizations rely on social attachment (Mills & Margulies, 1980) and embedded relationships (Uzzi, 1999) between customers and service providers. Service provider interchangeability is therefore low (Mills and Margulies, 1980). These embedded relationships engender norms of reciprocity and trust, which are key to motivating customers to actively participate in the design, delivery, and marketing of the goods and services they consume.
A good example of personalized relationships is post-graduate education, where performance ambiguity is high due to the intangibility of the product and the high degree of reciprocal interdependence between faculty and graduate students. For instance, students typically generate a lot of new ideas, which faculty then help them refine. In this way, students learn skills and gain a better understanding of a topic area. At the same time, faculty rely on their students’ ideas and labor to further their own careers, for instance, by staying up-to-date with topics and by writing joint publications with their students.

Information equivocality at both the design and delivery stage is high. For instance, it is difficult for students to articulate what they mean by and expect from a “good education” and where their research interests lie. When it then comes to writing the thesis, information equivocality is again high. Ideas, literature reviews and experiments that the candidate believes are moving him/her toward the completion of the thesis, might be interpreted very differently by the thesis committee.

Given the high input uncertainty and high performance ambiguity, both faculty and student tend to rely on long-term relationships and the social capital inherent in it (e.g., sense of obligation and reciprocity) to manage the exchange. Thus, thesis advisors allow their students to graduate even they may believe that the thesis could have been improved still, and graduate students might put their advisors names on publications based on their dissertation, even though their advisor may not have assisted in writing the specific paper.
EXAMPLES OF SST-ENABLED CO-PRODUCTION

Having outlined our contingency framework of co-production designs, we now turn our attention to three examples of SST-enabled service environments. Our purpose in analyzing these cases is to gain insights into the conditions under which networked technologies, such as Internet-based SST, affect changes in customer co-production. In particular, we are interested in understanding how SST impacts co-production design as well as the exogenous variables on which these designs are contingent, i.e., input uncertainty and performance ambiguity. Furthermore, we will highlight some of the unintended consequences generated by the SST-induced changes in co-production. We restrict our SST examples to Internet-based, self-serve applications in order to limit the variability of technological features that we need to consider in our analysis. Furthermore, we rely on SST examples with which readers are likely to be familiar so as to capitalize on their experience with these technologies.

Booking an Airline Ticket Online

When customers want to book an airline ticket online, they are typically confronted with a highly scripted interface into which they enter information about their travel needs, e.g., departure and arrival cities, dates of travel, number of passengers, etc. With this unequivocal request for assistance, the SST then offers the customer available flights sorted either by price or schedule. Some systems give customers the option of looking for a cheaper flight, provided their search parameters (e.g., travel days, number of stops) are somewhat flexible. If one of these options meets the customer’s needs, then s/he must enter personal contact and/or billing information to place the reservation on hold or purchase the ticket.
The co-production design instantiated in this SST example reflects the scripted market. The customer-firm interface is characterized by low input uncertainty because information about flights is relatively unequivocal. Although the customer has a wide variety of options in the form of search parameters, these are all standard and pre-defined, thus limiting customer requests to an established set of options. Neither the flight schedule, nor the functionality of the website is generated in response to a specific customer’s request for service. This low input uncertainty also makes the airline’s delivery process highly programmable.

Comparing the SST-enabled booking scenario with one in which a customer service representative assists that customer, we see that performance ambiguity in the online setting is even lower than in the provider-based setting. This is because the information regarding the flights is more transparent to the customer, and because the task of matching customer needs to the available flights is left completely to the customer. By disintermediating the service provider, the SST allows the customer to interact directly with the firm’s automated back-office operations. This makes it relatively easy to trace who is to blame if a mistake is made. For instance, the airline holds the customer responsible for errors. If customers inadvertently purchases a ticket that does not meet their needs (e.g., wrong dates of travel), they have to pay a change fee to remedy the situation. Nevertheless, it is in the airline’s interest to minimize the likelihood of customers making mistakes during bookings. For this reason, the design of the SST is such that customers are repeatedly asked to confirm their flight selections, thus giving them opportunities to uncover mistakes themselves.
One unintended consequence of having customers interact with a highly scripted SST is that it severely restricts customer input. As highlighted by Whitaker (1980), customers’ requests for assistance shape both current and future service offerings. In provider-mediated service environments, where service providers match customer needs with available service options, the customer-firm interface is less scripted than in SST-enabled service environments and customers may request services that fall outside of the predetermined options. By paying attention to such requests, firms may be able to detect emerging customer requirements and/or expectations, and respond to them by designing new service offerings. Thus, organizations effectively limit their benefits from co-production, which includes learning from customers, by offering customers a highly scripted SST-based interface and having them match their own needs to the available options.

**Book Recommendations at Amazon.com**

The main difference between the scripted and the personalized market design is the degree of input uncertainty. Experiential goods like books, music and movies, which highlight the ambiguous distinction between goods and services, provide a good illustration of how different levels of input uncertainty can be associated with one object. As a good, a book is a product whose characteristics (e.g., title, author, binding) can be readily described, thus making it suitable for a scripted market design. As a service, that is, the learning or entertainment that it renders, however, the book is highly equivocal. Thus, if a customer is looking for a good cookbook, input uncertainty is relatively high. In order to respond to the request for a good cookbook, a service provider would have to gather copious information about the customer, e.g., his/her tastes in cookbooks, before
recommending a suitable book. Therefore, the service of recommending books is an example of a personalized market co-production design.

On the surface, it appears that Amazon.com’s recommendation system is an example of an SST in the personalized market as it bases book recommendations on a customer’s personal specifications and tastes, even though these may be implicitly derived. Taking the example of a customer searching for a good cookbook, we see the Amazon customer entering search terms into a single, general-purpose search box. In addition to listing books that meet the customer’s keyword search, Amazon also makes available reader’s book reviews, as well as recommendations for other, similar books. These recommendations are generated by Amazon’s collaborative filtering technology, which builds a profile of customers’ preferences to personalize its book recommendations. Thus the options offered customers are not scripted, but dynamically assembled in response to a customer’s request for assistance.

However, the dynamic assembly of these recommendations is highly programmed through the collaborative filtering technology. Amazon creates an implicit profile for each customer based on past purchases. By comparing the purchases of customers with similar profiles (i.e., “neighbors”), Amazon then recommends books to a customer based on the books that his/her neighbors bought. Therefore, Amazon’s co-production design relies on the pooling of customer information (pooled interdependence), as well as the sequential ordering of customer requests and recommendation responses (sequential interdependence). Through its reliance on pooled customer information, the performance ambiguity inherent in Amazon’s recommendation service increases.
In order to determine what mode of cooperation is reflected in Amazon’s recommendation system, we need to separate between the two profiles on which the recommendations are based. Amazon relies on both implicit and explicit profiles. Implicit profiles are derived from customers’ purchases. To develop these, Amazon requires neither voluntary nor active cooperation from customers. In contrast, explicit profiles, which are based on customers’ intentional ratings, do require both active and voluntary customer cooperation. For instance, for every book that customers have purchased from Amazon, they can indicate whether they “loved” or “hated” it, and whether it should be included in their personal profile. Books purchased as gifts do not necessarily reflect the customer’s personal taste in books. Thus, explicit profiles refine and improve the accuracy of implicit profiles, and therefore increase Amazon’s ability to create value through recommendations.

It is the involuntary and inactive mode of customer cooperation underlying Amazon’s book recommendation service that potentially gives rise to unintended consequences. Unless customers voluntarily and actively refine their implicit profiles through explicit ratings, the accuracy of Amazon’s recommendations is likely to deteriorate. For instance, if customers do not identify which books were purchased for themselves and which were purchased as gifts, the customer’s profile remains vague. Since there are currently few incentives for customers to explicitly rate their past purchases, it is not clear that the implicit ratings, on which Amazon’s system currently relies for the majority of its recommendations, can be readily augmented by customer’s active and voluntary cooperation. Furthermore, there is nothing preventing customers from submitting explicit ratings that are untruthful. To the extent that Amazon’s
recommendation system takes explicit ratings into account at all, such inaccurate ratings could compromise the effectiveness of Amazon’s book recommendation service.

**Insurance Brokering at WebGA**

Since relationship-based modes of organizing tend to be more prevalent in B-to-B than in B-to-C settings (Anderson, Hakansson & Johanson, 1994), we take our SST example from the B-to-B realm. WebGA (a pseudonym) is a “general agent”\(^4\) in the small group health insurance market (a detailed description of this case can be found in Schultze (2003) and Schultze and Orlikowski (2004)). As a general agent, WebGA mediates between independent insurance brokers and insurance carriers by providing brokers with “proposals” or “quotes” for health insurance plans from multiple carriers, as well as help with selling, e.g., accompanying brokers on sales calls and providing brokers with sales advice and materials such as enrollment forms and benefits packages.

WebGA’s revenues consist of an “override,” a commission paid by the insurance carrier for insurance policies sold through a general agent. This means that there is a high degree of reciprocal interdependence between WebGA and its the customers. While WebGA relies on brokers’ sales success for its income, brokers rely on WebGA for information (both public and privileged) and assistance with selling. Furthermore, the input uncertainty is this realm is high. For instance, what is meant by “the best health plan” is equivocal. Thus, WebGA’s service reflects the conditions of the personalized relationship co-production design.

Given the reciprocal interdependence and the performance ambiguity surrounding the delivery of health insurance sales and service, WebGA has traditionally relied on embedded relationships as a means of coordinating the co-production between its sales

---

\(^4\) Double quotes signify emic or industry-specific terms.
Incorporating Self-Serve Technology into Co-Production Designs

reps and its broker customers. However, in 1999, WebGA complemented its relationship-based customer-firm interface with an Internet-based SST, which allows customers to generate their own insurance quotes. The logic underlying this design of relationship-SST complementarity was that the structured activities associated with *quoting* (i.e., matching the needs of the group seeking insurance with available plans) could be automated, thus freeing the WebGA reps up to engage in *consulting* activities. These consulting activities include less structured tasks such as gathering market intelligence from brokers, and providing them, in turn, with advice and privileged information, as well as hands-on assistance such as accompanying the broker on sales calls. Mutual adjustment is the primarily mode of coordinating consulting activities.

Prior to the introduction of the SST, quoting and consulting activities were inextricably intertwined as WebGA reps exchanged privileged information with their broker customers as part of quoting. In the new SST-enabled interaction, quoting and consulting are separated in space and time. Nevertheless, WebGA reps are alerted via email every time one of their brokers runs an online quote. This alert prompts the WebGA rep to follow up with the broker through a personal phone call or an email. In this way, the WebGA reps try to build or maintain their embedded relationships with their brokers. However, some brokers resist these attempts at relationship building by refusing to answer the WebGA rep’s calls or emails. Such acts of non-cooperation and non-compliance eventually result in the customer having his/her access to the online quoting system revoked. However, “cutting off” the customer in this manner represents a measure of last resort for WebGA.
Incorporating Self-Serve Technology into Co-Production Designs

Since the challenges that this SST-enhanced co-production design generated for both WebGA sales reps and their customers have previously been described in detail (Schultze, 2003; Schultze & Orlikowski, 2004), we will merely highlight here some of the unintended consequences that the SST generated. In the case of WebGA, there was considerable tension between the SST-based and the representative-based customer interfaces. For instance, the quoting engine included insurance plans that some of the WebGA reps could not endorse in good conscience. When brokers presented these plans to their clients and maybe even sold them, the WebGA reps had to resolve the conflict between the SST’s quoting results and their own consulting recommendations. Such inconsistencies between the two service channels undermined the customer-provider relationship.

Furthermore, as brokers relied increasingly on the SST to do their quoting, the reps’ opportunities to personally interact with them declined, making it more difficult to build the rapport and trust needed to add value through consulting activities. This was because the reps relied on their customers for intelligence from the field. Additionally, as brokers used the SST channel more than they did their WebGA rep, they also began to feel less beholden to their rep and less obliged to write their cases through WebGA. Thus, the weakening of the embedded relationships between customers and providers threatened WebGA’s ability to ensure customer cooperation when it came to submitting cases. Indeed, the introduction of the SST highlighted just how much WebGA relied on these relationships – and the social capital and rules of reciprocity inherent in them – to ensure that brokers complied with the unenforceable agreement that they write through WebGA those cases that had benefited from its quoting and consulting services.
Relating the WebGA case back to our contingency framework, we can see that prior to the introduction of its SST, the firm’s co-production design reflected the personalized relationship. Quoting and consulting activities were offered only as a bundle and through personal contact between a broker customer and his/her WebGA representative. The introduction of SST meant that quoting and consulting were separated. While the consulting activity remained in the personalized relationship quadrant, the quoting activity was expected to migrate into the scripted relationship quadrant. Indeed, WebGA wanted the relationship between its customers and its reps to remain intact, and therefore sought to maintain virtual integration between quoting and consulting. Every time a customer quoted online, their WebGA sales rep was alerted and was expected to follow up with the customer.

However, the lack of seamless integration between WebGA’s provider- and SST-based channels began to undermine the customer-provider relationship, suggesting that some of WebGA’s customers regarded WebGA’s quoting service in terms of a scripted market design, which is instantiated through arm’s-length relationships. This presented a problem for WebGA, however, as it did not price its quoting services separately from its consulting services.

**CONCLUDING REMARKS**

Co-production implies collaboration between firms and their customers in the design, delivery, and marketing of goods or services that customers consume. In this paper, we developed a multi-dimensional definition of co-production, as well as a contingency framework of co-production designs based on the exogenous variables of
input uncertainty and performance ambiguity. Using three examples of SST-enabled service environments, we generated insights into the implications of SST for the different co-production designs and into the unintended consequences that this technology generates. We now focus on Internet technology’s impact on the two exogenous variables that constitute the design contingencies in our framework, i.e., input uncertainty and performance ambiguity.

The performance ambiguity of a service is affected by the Internet’s support for the unbundling of goods and services. For instance, in the Internet-enabled environment, goods become increasingly digital, which allows them to be unbundled from their physical storage and distribution mechanisms (Altinkemer & Bandyopadhyay, 2000). Thus the sale of music in the form of a physical CD or tape is challenged by the availability of sound files that are distributed over the Internet (e.g., iTunes). Similarly, the Internet creates the possibility that business processes that typically define a corporation, e.g., customer management, product innovation and infrastructure management, can be distributed across multiple outsourcers whose operations are nevertheless so tightly integrated that the boundaries between the service providers are virtually imperceptible (Hagel & Singer, 1999). This is indicative of Malone et al’s (1987) electronic integration effect, i.e., the use of technology to create joint, interpenetrating processes between exchange partners.

As we saw from the WebGA case, the commercialization of the Internet allowed the firm to unbundle its service and separate it into quoting (moderate performance ambiguity) and consulting (high performance ambiguity). By separating quoting from consulting, WebGA effectively reduced the performance ambiguity of its overall service
offerings, as only consulting could now be considered high in performance ambiguity. Thus, similar to Malone et al (1987), we would suggest that there is an overall expansion of market-based service arrangements as low-ambiguity activities are increasingly unbundled from medium- to high-ambiguity activities and then migrated to the Internet.

Input uncertainty is also affected by the introduction of Internet technology. As our discussion of the Amazon.com recommendation system showed, high information equivocality can be managed through technologies such as collaborative filtering and natural language search engines. Thus, with increased technological sophistication, it becomes increasingly possible to automate the matching of customers’ needs and expectations with extant service options, especially if these are offered by a range of specialized service providers, who each face less input uncertainty. Furthermore, as Malone et al (1987) highlight, the Internet facilitates the creation of online brokers who match buyers and sellers, i.e., the electronic brokerage effect. These brokers create value by absorbing the input uncertainty that individual firms would otherwise have had to deal with (Anderson & Anderson, 2002). Thus, we anticipate that Internet technology renders environments that were previously considered high in input uncertainty less so. In this way, more opportunities for programmed means of coordination are created.

It is also important to note that the Internet increases transparency (Sinha, 2000). Not only does the technology make it easier for customers to do price comparisons across multiple online firms, but it is also easier for online firms to observe customers’ behavior. For instance, online firms can collect data on what items the customer purchased, but also on what items the customer considered during his/her purchase decision. To the extent that online firms collect customers’ online behavior and use it in their decision making,
they are leveraging the co-production opportunities that the Internet offers. However, this form of co-production is largely passive and involuntary (e.g., Amazon’s implicit profiling). The long-term implications of this type of co-production especially with respect to consumers’ rights, for instance, represent an important area of research.

In light of these Internet-enabled shifts, another opportunity for future research is to identify empirical evidence of these shifts. In the interest of taking the IT artifact seriously (Orlikowski & Iacono, 2001), such research should examine how the features of the Internet-based technologies in their social context of use affect each of the co-production dimensions (i.e., activity, mode of cooperation and type of interdependence). Particular attention should be paid to the unintended consequences of these technologies. As is evident from the WebGA case, even though the electronic integration effect ensures that no key operational information is lost when customers use Internet-based SST to generate their own quotes, tensions resulting from the lack of integration between the SST- and the provider-based channel were apparent. Such situated, practice-based insights are necessary to provide guidelines for the design of collaborative co-production environments.

Despite the challenges highlighted in the WebGA case, it is noteworthy that, on some measures, the deployment of Internet-based SST seems to have been successful. For instance, two reasons why organizations invest in SST are reducing costs, e.g., wages of customer service representatives, and enhancing customer service, e.g., giving customers 24x7 access to service and information (Meuter, Ostrom, Rountree & Bitner, 2000). Thus, reports that Southwest Airline was able to close three of its call centers at the end of 2003, in part because more and more customers were booking their flights
online, suggests that the implementation of SST has been successful in the scripted market co-production design quadrant. Furthermore, Amazon’s success as an online retailer, evidenced by the fact that in February 2004 it earned the highest customer satisfaction index score ever recorded by the American Society for Quality (Michigan State University), suggests that the deployment of SST for services located in the personalized market has been effective. Nevertheless, it appears that we have fewer examples of SST deployments, their success and their impacts on co-production in situations of high performance ambiguity, i.e., in the scripted and personal relationship quadrant. Future research should therefore focus particularly on these.

The notion that Internet-based technologies favor markets raises questions about the nature of co-production and collaboration in an increasingly Internet- and/or SST-enabled environment. As our earlier discussion of the different degrees of co-production implies, the most complete form of co-production and, therefore, customer-provider collaboration, occurs in the personalized relationship design. Here, the participants are actively engaged in multiple activities (e.g., design and delivery) in a mutually transformative way. In contrast, co-production in the scripted market is limited to requesting assistance in a sequentially interdependent manner. Given that most collaboration research assumes the former context of interpersonal interaction and mutual adjustment, and given that the space in which this form of co-laboring appears to be shrinking with the deployment of Internet technologies, it might behoove us to revisit our implicit and explicit definitions of collaboration.

While implementing SST in all three co-production scenarios is fraught with both technical and organizational challenges (e.g., designing an intuitive customer interface,
allaying customers’ privacy concerns, integrating SSTs with back-office systems), our discussion of the empirical examples of SST implementations leads us to believe that the most challenging SST implementation environment is the personalized relationship. This is partly because of the tensions associated with simultaneously maintaining two types of channels with the provider: a socially embedded, customer-provider relationship and an arm’s-length, customer-firm relationship (through the SST) (Selnes & Hansen, 2001). Furthermore, our case study of WebGA highlighted that the seamless integration between unbundled services – an assumption on which Hagel and Singer’s (1999) model of the unbundled corporation is based – is difficult to realize in practice.

Thus, for practitioners seeking to implement Internet-based SST into co-production environments, our analysis of the impact of SST on WebGA’s co-production design suggests that firms will need to pay attention to the way in which SST will impact their co-production design. In particular, the WebGA case suggests that if SST is used to unbundle services, that the location of each service in our contingency framework be determined. If the services are located in different quadrants, the likelihood of tensions between them is high. To manage these tensions, the SST-based customer-firm interface can either be subsumed under the customer-provider relationship, or the separation of the two services can be made more complete. Thus, giving service providers control over the SST, i.e., its content and its functionality, promises to minimize the tensions between the two channels. Alternatively, by separating the services that the SST and the service representative provide, the tension between the two channels can be eliminated.
REFERENCES


Incorporating Self-Serve Technology into Co-Production Designs


