Submitted to
manuscript

Service Design for a Holistic Customer Experience: A Process Perspective

(Authors' names blinded for peer review)

Modern service design practices conceptualize services as multi-step processes. At each step, customers derive an uncertain value, which depends on a functional benefit and a subjective experience. The latter may depend on experiences realized at previous steps. Service designs determine the provider effort at each step given that customers prefer less variable experiences, and enable a holistic perspective of the overall experience. We quantify two factors that shape service designs. The *type* of steps: i) *routine* steps, where effort increases the functional benefit and decreases the experience variability, and ii) *non-routine* steps, where effort increases the functional benefit at the expense of higher variability. A *holistic coupling* factor: at each step, the design is determined not only by experience realizations at predecessor steps, but also by how it can shape subsequent experiences. The optimal efforts depend on the combination of these two factors, giving rise to actionable design rules. For a positive coupling factor, step type homogeneity leads to "spread the effort" designs (complementary efforts), whereas a negative coupling factor suggests focusing the effort on a few key steps at the expense of the rest of the service (substitutable efforts). Step type heterogeneity reverses these recommendations. Moreover, when the customer experience unfolds according to a non-stationary process with serial correlation, the effort at each step is determined by an *impact zone* defined by the steps surrounding the focal service step. Stronger correlation always induces higher effort, whereas weaker correlation may induce less effort in services with heterogeneous step types.

Key words: customer experience, customer journey, design thinking, service design, service process, service provider, touchpoints

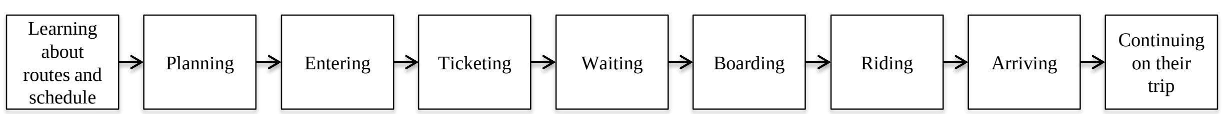
1. Introduction

Design has traditionally been associated with facets of product development, such as architecture (Ulrich 1995), function (Ulrich 2011) and style (Chan et al. 2017). Recently, firms have started broadening their perspective on design beyond the individual product attributes, and towards a holistic and human-centered focus on the needs and actions of the end user of a product. This approach, known as *design thinking* (Brown 2008), has enabled designers to expand beyond tangible goods. In fact, a fast-growing field of industrial design builds on the premises of design thinking, and it is concerned with the design of services.

The field of service design has established practices that define a structured approach to creating new or improving existing services. These practices draw on knowledge and tools from a variety of disciplines (Stickdorn and Schneider 2010). Yet, it has been established that two of the principal values that define the work ethos of service designers are: i) their emphasis on a holistic approach, and ii) their focus on empathy towards the end user (Fayard et al. 2017).[1] The following examples provide insights on how service designers enact these values in practice.

Service design for a better passenger experience: In 1996, Amtrak approached IDEO, a California-based design firm, to identify opportunities to improve the passenger experience in their high-speed rail service between Washington D.C, New York and Boston. At first, IDEO was tasked with solely redesigning the seats of the trains. Soon, the team recognized that for a successful project the customer experience would have to be addressed holistically. During a lengthy observation phase, which included riding trains with customers, IDEO discovered that the actual train ride was a small determinant of the overall customer experience. In fact, the initial focus on the seating comfort during the train ride ignored, and potentially jeopardized, the totality of the customer experience. This realization led to the conception of a *customer journey* map (Figure 1), which captured all the service steps that customers undertake when they use Amtrak's service, and which determined their overall experience. This visualization tool allowed IDEO to identify the steps of the customer journey where Amtrak needed to improve their efforts and then to propose design changes. This project, known as the Acela project (Brown 2009), is recognized by the broader service design community as the first systematic service design project (Fayard et al. 2017).

Figure 1 Service design for a better passenger experience; the Acela project.

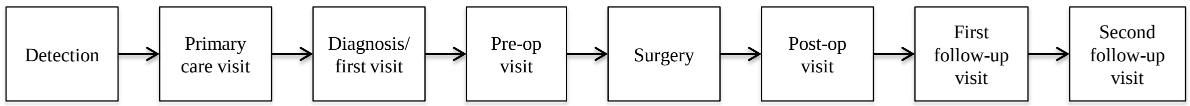


Note. Adapted from Bhavnani and Sosa (2006).

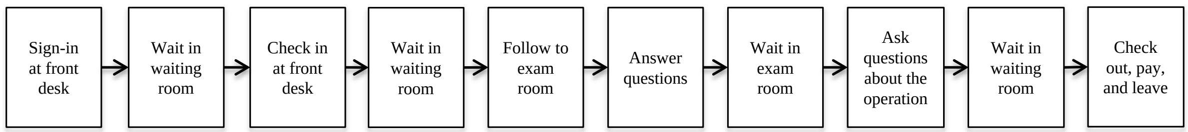
Service design for a better patient experience: Almost a decade later, in 2007, a team from the Carnegie Mellon's School of Design was asked to improve the patient experience in the neurosurgery clinic of the University of Pittsburgh Medical Center. Similar to the Acela project, the design team identified the major steps that patients undertake before and after they receive the surgical services of the clinic (Figure 2a) by shadowing both patients and medical staff. They also developed a more granular outline of the steps that patients go through during their first visit to the clinic (Figure 2b). Based on these journeys, the designers analyzed the service holistically and provided recommendations to improve the patient experience at the different stages of the service process (Cliver et al. 2007). Soon after, Kaiser Permanente implemented a similar approach to redesign its health care delivery (Arieff 2009).

[1] Fayard et al. (2017) recognize *co-creation*, namely, actively involving the service provider and the relevant stakeholders in the design process, as the third principal value that defines the service design work ethos. In this paper, we treat the service designer and provider as a single entity (i.e., a design team) and refer to them interchangeably.

Figure 2 Service design for a better patient experience; the University of Pittsburgh Medical Center project.



(a) The overall patient journey



(b) The patient journey during the diagnosis/first visit

Note. Adapted from Stickdorn and Schneider (2010) and Cliver et al. (2007).

These examples (for additional cases see Bitner et al. 2008, Stickdorn and Schneider 2010, Kalbach 2016) reveal that service designers employ similar methodologies across different service contexts. They engage in a user-centric observation to comprehend and analyze existing offerings from the customers' point of view. Then, they identify the corresponding customer journey, which they use to build better designs.

The customer journey mapping[2] is a powerful visualization technique. It delineates the different steps that customers undertake throughout the service and presents a structured and holistic view of the customer experience. Oftentimes, designers refer to the different service steps as *touchpoints* because they represent instances when customers interact with the provider. These instances of interaction, are the main means and interfaces through which a provider affects the value that customers derive.

The identification of the service steps and the delineation of the customer journey constitutes an insightful starting point. Yet, it remains unclear how this visualization can be used to derive actionable guidelines that the service designer can follow to establish a successful service offering. In practice, the inception of the customer journey map marks the beginning of a rather fuzzy ideation process during which designers and providers seek tangible interventions on the steps of the service process (Fayard et al. 2017). However, given the absence of actionable guidelines, the outcome of this ideation process relies primarily on the experiences, interpretations and perhaps creativity of the designers. As such, the resulting recommendations become ad hoc interventions to the corresponding customer journey. In this paper, we complement the existing qualitative design

[2] The historical roots of the customer journey mapping trace back to the *service blueprints* first introduced by Shostack (1984, 1987). Customer journey maps typically focus on the customer's view, whereas service blueprints expand the scope to provide additional details regarding support and backstage processes. Other visualization techniques such as *value stream mapping* (Rother and Shook 2003, Martin and Osterling 2014) and *lean consumption maps* (Womack and Jones 2005) encourage "going to the gemba" as a means of adopting a "customer-centric thinking." They emphasize waste reduction and work flow improvement in internal operational processes. For a comprehensive overview of the different diagrams used in practice we refer the reader to Kalbach (2016).

methods with an analytical approach that builds upon the customer journey concept and guides palpable service design choices.

Our approach capitalizes on the two pillars that underpin the value of the customer journey: i) *holism* and ii) *empathy.* We recognize that customers are serviced through a multi-step process. Each step of the service process contributes to the overall value delivered by the service. The value that customers perceive at each step comprises two parts: i) a *functional* part, which describes the tangible customer benefit that the service step is designed to deliver, and ii) an *experiential* part, which is uncertain and may to add or subtract from the functional part. The experiential uncertainty shapes how customers perceive the functional benefit they are inrended to receive, and stems from the inherent variability that uniquely characterizes service environments (Zeithaml et al. 1985, Murray and Schlacter 1990). Naturally, we consider that, all else being equal, customers prefer less variable experiences.

The service provider shapes the design of the service through the determination of effort at each step of the process. Consistent with previous research on services, we account for the interaction between average service value and variability (Sriram et al. 2015). This intuitively reflects our functional and experiential value contributions. More effort at a service step should always increase the functional benefit that customers derive. However, it may inadvertently increase or decrease the variable experiential value of the customers. This consideration lends itself to a natural typology of service steps: i) steps where the provider's effort leads to lower experience variability (hereafter, *routine* steps) and ii) steps where the provider's effort leads to higher experience variability (hereafter, *non-routine* steps). Furthermore, the step-specific experiences may not be independent of each other; they may be positively or negatively correlated.

Our analysis captures interaction effects that extend beyond individual service steps. This echoes service design imperatives that have found application in practice (McKinsey&Company 2016). Fundamentally, the design is driven by the holistic coupling of the experiences at the different service steps. We explicitly quantify this coupling and analyze its properties. Our results reveal how the optimal effort at a step is dependent on the characteristics of the rest of the steps. This represents the holistic consideration necessary for designing services. In fact, the characteristics of one part of the process may affect the customer experience, and as an extension the provider's effort, at steps that bear no obvious connection. The level of this coupling identifies within the context of each service, along with the types of steps, define whether the efforts at any two steps should be complementary or substitutable. Ignoring, or mischaracterizing this effect, can result in significantly different designs, leading the provider to erroneous effort investments.

To build additional insights at the level of the complete service process, we focus on a general class of processes where the experiences at the different steps are serially correlated, and their correlation decreases in the number of steps between two experiences. This allows us to address generalizable structures while retaining analytical tractability. We find that the optimal effort invested at a service step is determined by an "impact zone" of steps defined around it. In other words, designers can approach the service process as *modules* of such zones. Such a design guideline is particularly helpful to providers of services that comprise a large number of steps. Furthermore, we find that the provider exerts more effort in *cohesive* services, that is, services where customers perceive their experiences at the different steps as highly correlated. Finally, we characterize the conditions under which the provider exerts more effort at the first or last step of the service process; the efforts at these steps can differ significantly even if both of them serve similar functions.

2. Literature

Since early on, researchers subscribed to the process-based view of services, which suggests that customer experiences occur over multiple stages. Notably, Karmarkar and Pitbladdo (1995) were among the first to delineate research opportunities on the design of multi-stage service processes. This multi-stage perspective enabled several key dimensions of analysis in service design: i) the allocation of provider effort across the service stages, ii) the sequencing of the stages, and iii) the level of interaction (known as co-production) between the provider and the customer at each stage. We review the literature around these dimensions. Our work is closely related to the first stream of research, and it complements the other two streams.

With respect to the provider's effort allocation, Soteriou and Hadjinicola (1999) are the first to determine the optimal budget allocation across different service stages. Their objective is to minimize the difference between the maximum possible level of service quality perception, and the customers' mean service quality perception. In contrast to our paper, they assume a deterministic setting with independent service stages. Soteriou and Chase (2000) extend this line of work to account for uncertainty, but they do not consider the effect of interdependencies across the customer experiences. Bellos and Kavadias (2019) introduce such an effect, but assume discrete effort allocation. Specifically, they consider a binary decision of whether a service task is offered by the provider or not. Their model seeks to answer a different design question, i.e., the delegation of service tasks to customers. Moreover, they analyze a specific interdependency structure. In this work, we treat service design as a continuous variable of resources allocated across different service steps and we consider broader interdependency structures. Our focus on such structures allows us to derive novel design insights regarding how the optimal effort allocation is defined through the interdependencies of the experiences and their interactions with the different types of service steps. These insights are not offered and cannot be deduced by previous research on service design. In a different context, Arora et al. (2019) focus on the effort that non-profit organizations exert at the advisory or delivery stage of their service process in order to maximize social impact. Our analysis provides a more granular and generalized treatment of the interdependencies present in a multi-step service process. To our knowledge, our paper is the first to provide structural results regarding the effect of such interdependencies on the optimal service design.

A few studies have focused on multi-stage services whereby the provider's effort allocation affects specifically system congestion and an important dimension of customer experience, namely the waiting time. For instance, Carmon et al. (1995) model customers' dissatisfaction in a two-stage service system and its implications on the timing of the service provision. Xue and Harker (2003) also consider a two-stage service where at the first stage customers are self-served, and once they proceed to the second stage, they join a single-server queue. In this context, the service provider determines the optimal level of self-service as approximated by the optimal workload division between the provider and the customer. Tong et al. (2016) study the effect of innovations that reduce the service time at the first stage of a two-step service on the overall service quality and congestion. In similar spirit, Lee et al. (2012) examine the optimal contract parameters (i.e., staffing levels and referral rates) that a provider of a two-stage service uses when outsourcing part of the process to an external provider. These studies focus on specific types of service design interventions i.e., staffing, which aim at mitigating service congestion. In our model, we do not explicitly examine system effects such as congestion. However, we focus on one-to-one customer-provider interactions, and we also allow for service steps where investing effort to improve the functional value at a service step may lead to an increase in the variability of the experience; we term such steps as non-routine.

The second stream examines how the sequencing of different service encounters affects the customer experience. Bitran et al. (2008) call for more attention to the temporal aspects of the service delivery process. To that end, Dixon and Verma (2013) empirically find that the sequencing of different musical events plays an important role in determining customers' decisions to repurchase season subscriptions of performing arts. In a similar context, Dixon and Thompson (2016) develop a computational approach to characterize the optimal sequence of such service encounters. Baucells and Sarin (2007, 2010) determine the optimal sequence of intertemporal consumption in the presence of satiation and acclimation (habit formation). In addition to acclimation, Das Gupta et al. (2016) consider memory decay and develop a model for experiential services that determines the optimal sequence and duration of service encounters. Ely et al. (2015) characterize the optimal way to reveal information in order to maximize expected suspense or surprise. Accounting for surprise and anticipation, Dixon et al. (2017) use scenario-based experiments to identify the optimal design of a sightseeing tour. More recently, Martínez-de Albéniz and Valdivia (2019) study, both empirically and via the use of an optimization framework, the impact of content decisions such as duration and synchronization of exhibitions on museum attendance. For excellent reviews and summaries of insights of this literature we refer the reader to Roels (2019) and Dixon and Victorino (2019). This research primarily focuses on services that comprise homogeneous stages (e.g., music concerts, professional education, sightseeing tours) and continuous experiences. We consider services that comprise heterogeneous types of steps serving different and distinct purposes. Building upon Ariely and Zauberman (2003), we posit that such services are less influenced by sequence effects. Thus, we focus on service settings where the sequence of events cannot be altered; for instance, in health care the stage of diagnosis always precedes the stage of treatment. In that respect, our paper addresses the stage of the service design process where the sequence of activities has been determined. At this design stage, the optimal sequence patterns identified by the aforementioned research serve as input to our model; hence, we view our work as an important complement to that research stream.

Finally, the high degree of customer presence and interaction in many service environments (Sasser 1976, Chase 1978, 1981) has motivated research into the optimal allocation of effort between a service provider and a customer. Xue and Field (2008) focus on knowledge-intensive services, such as consulting and determine the optimal pricing and workload division between a consulting firm and their client under different (incomplete) service contracts. Roels et al. (2010) also study pricing and effort division decisions in collaborative services; in contrast to Xue and Field (2008), they assume substitutability of effort levels and determine the optimal contract selection based on the service output sensitivity to each party's effort. White and Badinelli (2012) and Roels (2014) analyze single-stage co-productive service systems and determine the optimal effort division that guarantees total surplus maximization. White and Badinelli (2012) assume a deterministic setting, whereas Roels (2014) accommodates uncertainty and endogenizes the degree of effort substitutability. We consider service settings where explicit contracts that elicit certain customer effort are difficult to enforce, or even define. In contrast to this stream of literature, our focus is on the provider's effort allocation across the entire service process, as defined by the different steps that the customers go through to receive the service. At the individual service step level, we lie complementary to this stream as the provider's effort at each step determines the division between functional and experiential value, which affects the overall value that a customer derives.

3. The Model

In this section, we detail the context of the service, the determinants of the customer value, the provider's decisions, and our model assumptions. Our objective is to characterize how a service provider should allocate her effort across the different steps of a multi-step service process (i.e., customer journey), given the structural properties of the process.

3.1. Contextual Assumptions

Consider a service process $\mathcal{J}$ comprising a set of n well-defined steps $\{1,\ldots,n\}$ that customers have to undertake (i.e., customers do not make process/step choices). The service steps are identified and delineated by service designers via methodologies such as shadowing and contextual interviews (Fayard et al. 2017). Then, designers aim to determine how much effort $\delta_i \in [0,1]$ should be invested at each step $i \in \mathcal{J} = \{1,\ldots,n\}$ of the service process; different effort allocations across the service steps describe different designs. We assume away any strategic interactions of a competitive market and we adopt the perspective of a monopolist provider who offers a single service. The value $\tilde{\nu}_i$ that a customer derives from each service step i is uncertain. Customers do not factor in explicitly how their participation in the service process may affect the value of other customers. As customers go through a series of steps, their experience realization at one service step may affect the experience realization at another step (e.g., unsatisfying experiences early in the service process may set the stage for unsatisfying experiences in later stages). All else being equal, customers prefer less variable experiences, and they exhibit the same sensitivity to the presence of variability. Our emphasis is on the *design* as opposed to the real-time management of a service process. The provider's efforts are non-discretionary, and we do not analyze how the provider's or customers' decisions may possibly be adjusted during the *execution* of the service.

3.2. Customer Value and Service Process Characteristics

Consider a customer who in order to satisfy a need, goes through a number of steps that form a service process $\mathcal{J}$. At each step i of the process, the customer realizes a value $\tilde{\nu}_i$; we refer to $\tilde{\nu}_i$ as the *perceived* value. The realization of $\tilde{\nu}_i$ is uncertain and given by $\tilde{\nu}_i \doteq \mathcal{V}_i + \tilde{e}_i$, where $\mathcal{V}_i$ denotes the *functional* value that step i is designed to deliver (e.g., average effectiveness of a medical treatment) and $\tilde{e}_i \sim \mathcal{N}(0,\sigma_i)$ the realized *experiential* value. For an overview of practical metrics of the functional and experiential values please see Forbes (2019).

The functional value $\mathcal{V}_i$ is partly determined by the provider's effort at service step i. In particular, the provider invests effort $\delta_i \in [0,1]$, which contributes to $\mathcal{V}_i$ the value $V_i \doteq \delta_i V_i^H + (1-\delta_i)\,V_i^L$, where $V_i^H > V_i^L \geq 0$. The value V_i^L represents the minimum customer value that step i should be designed to contribute so that it is considered an integral part of the service; V_i^H captures the maximum feasible functional value that the provider can generate at step i by investing maximum effort (e.g., a state-of-the art lab equipment).

The remaining part of $\mathcal{V}_i$ is determined by the rest of the steps. The service steps may be subject to technical dependencies; the functional value at step i may be enhanced by the effort invested at other steps of the process. For example, a state-of-the art lab equipment can allow a doctor to reduce the amount of time spent on differential diagnostic procedures during the medical examination. To capture such dependencies, we assume that the provider's efforts at the rest of the steps contribute to the functional value at step i, according to $\sum_{j\neq i}\beta_{j,i}V_j$, where $\sum_{i\neq j}\beta_{j,i}\in[0,1)$; $\beta_{j,i}\in[0,1)$ for all $j\neq i$ captures the extent to which the functional value at step j enhances the functional value at step i. Overall, we posit that $\mathcal{V}_i \doteq \sum_{j=1}^{n}\beta_{j,i}V_j$ (for $j=i$, $\beta_{j,i}=1$).

The experiential value $\tilde{e}_i$ may add or subtract from the functional value. Different sources of uncertainty originating from the provider or the customer may contribute to the fact that the experience at any service step cannot be predicted exactly. For instance, overlooking information on a patient's medical history may challenge the diagnosis and result in negative customer experience. However, even if the functional value delivered at a step is always consistent (as evaluated based on objective measures; e.g., diagnostic accuracy) the eventual customer experience may not be always the same. For instance, the same patient may derive a different experience at a specific health care practice even if every time that he visits he has his information and vitals taken by the same nurse, waits in the same exam room, and is seen by the same doctor to treat the same health condition (see Bowen and Ford 2002 for a thorough discussion).

We formally capture this uncertainty by assuming $\tilde{e}_i$ to be normally distributed with a mean zero and a standard deviation σ_i for all $i\in\mathcal{J}$. Then, the provider's effort invested at a step i, affects the extent of the uncertainty as follows: $\sigma_i \doteq \delta_i\sigma_i^H + (1-\delta_i)\,\sigma_i^L$. The value of σ_i^L captures the inherent variability of the experience at step i when the provider invests the minimum possible effort, whereas σ_i^H captures the variability when the provider invests the maximum possible effort.

We entertain the two possible scenarios for the effect of effort (δ_i) on the variability (σ_i). More effort at a step may lead to higher or lower σ_i, i.e., $\partial\sigma_i/\partial\delta_i = \sigma_i^H - \sigma_i^L \gtrless 0$: i) $\sigma_i^H < \sigma_i^L$ implies that larger values of δ_i result in smaller σ_i and ii) $\sigma_i^H > \sigma_i^L$ implies that larger values of δ_i result in larger σ_i. These scenarios map naturally onto a classification of service steps as either *routine* or *non-routine*, as also noted in recent literature; see (Bellos and Kavadias 2019).

In routine steps, service outcomes are evaluated along objective specifications and metrics (e.g., waiting time until called to the exam room). In such steps, the provider's improvement of the functional value V_i (e.g., check-in in an expeditious manner) also reduces the variability of the experience (i.e., $\sigma_i^H < \sigma_i^L$).

On the contrary, in non-routine steps, the service outcome is evaluated along more intangible dimensions (e.g., personal taste, sense of privacy, empathy, peace of mind). In such steps, improving the functional value V_i may lead to a more variable experience (i.e., $\sigma_i^H > \sigma_i^L$). Consider the example of the gradual introduction of sophisticated diagnostic methods into the detection of small lumps during a medical examination. An upfront approach can be the simple manual examination that doctors perform by trying to "feel" the possibility of a lump. One could credibly argue that the approach bears some diagnostic accuracy but up to a point (e.g., small lumps that are not superficial might not be detected). At the same time a trained physician would carefully approach the process ensuring no major experiential drawbacks. For instance, should the patient feel uncomfortable they would stop, discuss, or perhaps if the patient would be accommodating they could try to go faster and explore more.

Moving to a technology-based approach like ultrasound screenings, one could ensure a better diagnostic accuracy as the ultrasounds can detect further abnormalities. The use of the technology can lead to further patient appreciation due to increased confidence in the diagnostic outcome. However, the increased accuracy (in medical screening terminology, sensitivity; i.e., low rate of false negatives) may also introduce specificity issues (i.e., a higher rate of false positives) leading to unnecessary "health scares." Furthermore, the use of a device and material of some liquid form on top of the skin might create more variable experiences with the "average" patient as certain people may dislike the use of the material or the feeling of the device on them. Therefore, one could claim that the introduction of the technology might have increased the variability of experiences.

Finally, moving to costlier and more sophisticated methods such as MRI systems, one can argue the confidence in the diagnostic outcome is even higher (due to increased accuracy/sensitivity and specificity). Yet, this may be associated with even more negative experience realizations due to the discomfort that many patients experience during the use of the equipment; see OpenIDEO (2013) on how design thinking has been used to account for such negative experience realizations of paediatric MRI patients.

Our distinction between routine and non-routine steps is meant to capture the fact that costly design choices in the functional value may have unintended effects on the customer experience. As our examples indicate, the classification of a step as a routine or non-routine ultimately depends on the step properties and in particular on the dimensions along which customers evaluate its outcome. Our typology is similar in spirit to other categorizations found in the literature. For instance, Teboul (2006) differentiates services that are unique and varied in nature from ordinary/commoditized services, which are more limited and standardized. Similarly, Roels (2014) distinguishes between standard and non-standard service tasks, where standard tasks are completed more consistently. In our context, a service may comprise a mixture of ordinary/standard (i.e., routine) and varied/nonstandard (i.e., non-routine) steps.

Regardless of the type, we also posit that the different service steps may admit interdependent experiences. The experience that a customer derives at one step may be influenced by the experience derived at another step. We use $\rho_{i,j}$ to denote the correlation between the realizations of the experiences at any two steps i and j, and $\mathbf{P}$ to denote the resulting correlation matrix. We consider two structures for $\mathbf{P}$. The first is a general structure where for each i,j element of $\mathbf{P}$ we have $\left[\mathbf{P}\right]_{i,j} = \rho_{i,j} \in (-1,1)$ for $i \neq j$ and $\left[\mathbf{P}\right]_{i,j} = 1$ for $i = j$. Positive values of $\rho_{i,j}$ capture the cases in which a satisfying (or unsatisfying) experience at step i sets the stage for a satisfying (or unsatisfying) experience at step j. Negative values of $\rho_{i,j}$ describe the opposite; a satisfying (or unsatisfying) experience at step i renders a customer more (or less) difficult to gratify at step j. This generic structure allows us to account for cases where customers move through the service process in a non-linear fashion (e.g., possibly revisiting previous steps), or when certain steps run in parallel with other service steps. For instance, when visiting a pediatric dentist, anxious children (and parents) derive value from an appropriately-themed decor and equipment (e.g., jungle, space station, construction site) that may extend from the check-in to the waiting and exam areas. That is, in this case, the choice of decor/ambiance can be thought of as a service step (i.e., an interface through which the provider can affect the value customers derive) running in parallel with the steps of check-in, waiting, and exam, and therefore, may induce correlated experiences.

The second structure of $\mathbf{P}$ we consider, emerges when the customer experience unfolds according to a non-stationary process exhibiting serial correlation, i.e., correlation that decays as the elapsed time (in our case, different sequential service steps) between experiences increases. We consider a generalized autoregressive model known as first-order antedependent ($AD(1)$) transition model (see Gabriel 1962). In an $AD(1)$ process, $\tilde{e}_i \perp \{\tilde{e}_{i-2}, \ldots, \tilde{e}_1\} \mid \tilde{e}_{i-1}$, that is, each experience, *given* exactly one immediately preceding experience, is independent of all further preceding experiences.[3] The $\mathbf{P}$ that characterizes an $AD(1)$ transition model is described by the matrix elements $\left[\mathbf{P}\right]_{i,j} = \prod_{m=j}^{i-1} \rho_{m+1,m} \in (-1,1)$ for $i \neq j$ and $\left[\mathbf{P}\right]_{i,j} = 1$ for $i = j$. The correlation between two non-adjacent service steps i and j with $j > i+1$, is the product of the correlations of all the adjacent steps that lie between i and j (see Gabriel 1962, Zimmerman and Núñez-Antón 2009).[4]

The total value that a customer derives from the entire service process is $\tilde{\nu} = \sum_{i=1}^{n} \tilde{\nu}_i$. Ariely and Zauberman (2003) provide evidence in support of the cumulative experience assumption in settings where the experiences are partitioned in multiple discrete components (i.e., our service steps can be viewed as different experience partitions). We posit that in their valuation of the service, customers account not only for the mean but also for the overall variability of their experience. Rust et al. (1999) provide support for our assumption as they show that, in addition to the expected quality

[3] An $AD(1)$ is a generalization of a stationary first-order autoregressive $AR(1)$ model (e.g., one where the experience evolves according to $\tilde{e}_i = \rho\tilde{e}_{i-1} + \tilde{\varepsilon}_i$, with $\rho \in (0,1)$ and $\tilde{\varepsilon}_i \sim \mathcal{N}(0,\sigma)$) because it relaxes the assumptions of equal variances σ^2 and correlations ρ across the different i periods. A higher p-order $AD(p)$ model can also be considered where $\tilde{e}_i \perp \{\tilde{e}_{i-p-1}, \ldots, \tilde{e}_1\} \mid \{\tilde{e}_{i-1}, \ldots, \tilde{e}_{i-p}\}$, that is, each experience, *given* exactly p immediately preceding experiences, is independent of all further preceding experiences. This is equivalent to $\tilde{e}_i$ having a Markovian dependence of order $p \geq 1$ (Pourahmadi 1999, Diggle et al. 2002). We discuss the design implications of the experience unfolding according to an $AD(p)$ in §4.

[4] The estimation of the structure of $\mathbf{P}$ can be derived from the application of likelihood-based estimation testing typically used to estimate heterogeneous covariance structures for repeated measures (e.g., the restricted/residual maximum likelihood (REML) method; see Wolfinger 1996 for such an application and Gabriel 1962, Zimmerman and Núñez-Antón 2009 for additional references).

(i.e., functional value), customers are also sensitive to the perceived variance of the service, and that larger variability may outweigh the benefit of higher expected quality (see also Kannan and Proença 2010 for a review of the related literature). Furthermore, empirical findings that associate high experiential variability with low customer satisfaction have also been reported by practitioners (McKinsey&Company 2014).

We capture customers' aversion to variability through a mean-variance formulation (i.e., customers are risk averse). In particular, we approximate the customers' expected net utility[5] $\mathbb{E}\left[U\left(\tilde{\nu},\pi\right)\right]$ from the service as $\mathcal{U} \doteq \mathbb{E}\left[\tilde{\nu}\right]-\frac{r}{2}\mathrm{Var}\left[\tilde{\nu}\right]-\pi=\sum_{i=1}^{n}\mathcal{V}_i-\frac{r}{2}\left(\sum_{i=1}^{n}\sigma_i^2+2\sum\sum_{i<j}\rho_{i,j}\sigma_i\sigma_j\right)-\pi$,[6] where $r\geq 0$ is the measure of customers' risk aversion and π the price of the service; see Karmarkar and Pitbladdo (1997) and Kim et al. (2007) for additional applications of the mean-variance approximation in manufacturing and service contexts.

3.3. The Provider's Design Problem: Effort Allocation

The monopolist service provider determines the effort δ_i at each step i and the price π that maximizes $\mathcal{U}\left(\delta_1,\ldots,\delta_n,\pi\right)$. Investing δ_i to improve the functional value V_i at each step i imposes a cost c_iV_i, where $c_i \in \left(\sum_{j\neq i}\beta_{j,i},1\right)$. The provider solves:

$$
\begin{aligned}
\underset{\delta_1,\ldots,\delta_n,\pi}{\text{maximize}}\quad & \mathcal{U}\left(\delta_1,\ldots,\delta_n,\pi\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{j,i}\left(\delta_jV_j^H+\left(1-\delta_j\right)V_j^L\right)\\
& -\frac{r}{2}\quad\left(\sum_{i=1}^{n}\left(\delta_i\sigma_i^H+\left(1-\delta_i\right)\sigma_i^L\right)^2+2\sum\sum_{i<j}\rho_{i,j}\left(\delta_i\sigma_i^H+\left(1-\delta_i\right)\sigma_i^L\right)\left(\delta_j\sigma_j^H+\left(1-\delta_j\right)\sigma_j^L\right)\right)-\pi\\
\text{s.t}\quad & \pi-\sum_{i=1}^{n}\left(\delta_iV_i^H+\left(1-\delta_i\right)V_i^L\right)c_i\geq\pi_0\\
& 0\leq\delta_i\leq 1,\ \text{for all } i\in\mathcal{J}.
\end{aligned}
$$

The first constraint ensures that the chosen design results in profit, which at least satisfies a reservation value $\pi_0\geq 0$. The value of π_0 may subsume fixed costs that the provider incurs at different service steps regardless of the invested effort. We acknowledge that in practice more involved cost structures like economies of scale, convex increasing costs, or fixed costs increasing as a step function with respect to the provider's effort are possible. We believe that analyzing the effect of such structures on the design of service processes is a promising avenue for future

[5] From a technical viewpoint, the mean-variance approximation is exact for a negative exponential utility $U\left(x\right)$ and normally distributed x. Levy and Markowitz (1979) and Kroll et al. (1984) have demonstrated the applicability and practical accuracy of the mean-variance formulation for various utility forms and probability distributions.

[6] The current formulations of V_i and σ_i could be modified to explicitly account for the fact that customers may be more sensitive to the variability or they may assign more weight to the functional value they derive at certain steps over others. For notational parsimony, we consider such effects to be subsumed in the values of V_i^H, V_i^L and σ_i^H, σ_i^L.

research. The constraints on $\delta_i \in [0,1]$ for any $i \in \mathcal{J}$ ensure that $V_i \in [V_i^L, V_i^H]$ and $\sigma_i \in [\sigma_i^L, \sigma_i^H]$ (or $\sigma_i \in [\sigma_i^H, \sigma_i^L]$).[7]

The formulations of V_i and σ_i aim to capture the interplay (through δ_i) between average service value and variability in an analytically tractable manner. Such an interplay has been extensively documented in previous literature (see Sriram et al. 2015 and references therein). Hence, we do not assume V_i and σ_i as separate decision levers. This might be the case if V_i is proxied through performance quality, and σ_i through conformance quality; this distinction is prevalent in manufacturing contexts (Karmarkar and Pitbladdo 1997). In such contexts, the firm can affect V_i through changes in the product attributes and σ_i through changes in the production process. In contrast to a typical service process, customers are not present in those settings. On the contrary, pinpointing the source of variability in a service setting is often more challenging as it may stem from the customer and/or the provider. To obtain first-order insights, we focus on capturing how improving the performance quality (i.e., the functional value V_i) affects the variability (i.e., lack of conformance due to σ_i) of the customer experience. We leave the treatment of V_i and σ_i as separate decision levers for future research.

4. Analysis

In this section, we solve the problem of determining the optimal effort δ_i^* for each step $i \in \mathcal{J}$ and optimal price π^* that maximize customers' net utility. In our setting, the provider does not trade off effort investment with price (i.e., the optimal effort allocation is not affected by π^*). For that reason, the presentation of our results focuses on the optimal effort δ_i^* invested at each step i of the service process. From a technical standpoint, for the rest of the analysis we focus on the cases where $\partial \mathcal{V}_i / \partial \delta_i = V_i^H - V_i^L \in \left(\underline{\Delta V}_i, \overline{\Delta V}_i\right)$; analytical expressions for the $\underline{\Delta V}_i$ and $\overline{\Delta V}_i$ thresholds are provided in the Appendix. This describes the maximum improvement of the functional value at each step and ensures that $\delta_i^* \in (0,1)$ for any $i \in \mathcal{J}$, which allows us to present the richer cases capturing the interaction effects among all the service steps.

Proposition 1. *For each service step $i \in \mathcal{J}$, the optimal effort investment is given by* $\delta_i^* = \sum_{j=1}^{n} \frac{\left(1-\left(c_j - \sum_{m \neq j} \beta_{j,m}\right)\right)\left(V_j^H - V_j^L\right)}{r\left(\sigma_i^H - \sigma_i^L\right)\left(\sigma_j^H - \sigma_j^L\right)} \left[\mathbf{P}^{-1}\right]_{i,j} - \frac{\sigma_i^L}{\sigma_i^H - \sigma_i^L}$ *where* $V_i^H - V_i^L \in \left(\underline{\Delta V}_i, \overline{\Delta V}_i\right)$ *for all* $i \in \mathcal{J}$.

[7] The maximization problem stated above is equivalent to determining the δ_i efforts and the price π that maximize the profit $\pi - \sum_{i=1}^{n} c_i V_i$ subject to a customer participation constraint $\mathbb{E}[\tilde{\nu}] - \frac{r}{2}\mathrm{Var}[\tilde{\nu}] - \pi \geq \pi_0$, where $\pi_0 \geq 0$ is a threshold value for the customers' reservation utility. In either formulation, the price π ensures a necessary condition i.e., the provider derives a non-negative profit and the customer derives a non-negative utility. Similar to Karmarkar and Pitbladdo (1997), the optimal design (i.e., effort allocation) is not affected by π (i.e., the provider does not trade-off effort investment with price). An earlier version of the manuscript presented the $\pi - \sum_{i=1}^{n} c_i V_i$ formulation. We are thankful to the Associate Editor for recommending the current formulation.

Proposition 1 carries theoretical importance. It establishes an analytical result on a managerial problem that to this date has been treated descriptively by design practitioners. The offered solution captures how all different properties of each individual step influence the provider's design at a specific step. More importantly, though, it quantifies the effects of the interplay between the different steps and how this determines the optimal effort allocation across all steps of the service process. This interplay is described by three distinct factors: i) the *technical coupling* of a step i with the rest of the steps, captured by the factor $\sum_{m\neq j}\beta_{j,m}$, ii) the *holistic coupling* of a step i with the rest of the steps, captured by the terms $\left[\mathbf{P}^{-1}\right]_{i,j}$ for all $j\neq i$, and iii) the *step types* captured by the pairs of individual steps via the $\left(\sigma_i^H-\sigma_i^L\right)\left(\sigma_j^H-\sigma_j^L\right)$ terms.

Among the factors mentioned above, the technical coupling has a relative straightforward effect and is subsumed in the service provision cost. Specifically, it reduces the cost of each step from c_i to $\left(c_i-\sum_{j\neq i}\beta_{i,j}\right)$. Given the straightforward role of $\sum_{j\neq i}\beta_{i,j}$ in determining δ_i^* and in order to maintain notational parsimony, in the rest of the analysis we consider $\sum_{j\neq i}\beta_{i,j}\to 0$ for all $i\in\mathcal{J}$.

The rest of the factors exhibit more involved effects. For instance, the step types may have a positive or a negative effect on the provider's effort at step i because $(\sigma_i^H-\sigma_i^L)\left(\sigma_j^H-\sigma_j^L\right)\lesseqgtr 0$ depending on the types of the individual steps. Of particular interest is the quantity $\left[\mathbf{P}^{-1}\right]_{i,j}$. It denotes the element in row i and column j of the *inverse* of the correlation matrix. To understand the role of $\left[\mathbf{P}^{-1}\right]_{i,j}$, it is important to highlight the theoretical differences between $\mathbf{P}^{-1}$ and $\mathbf{P}$. The correlation matrix $\mathbf{P}$, captures the interdependencies of the experiences at the different steps in a multibivariate fashion (Raveh 1985); the inverse $\mathbf{P}^{-1}$ captures such dependencies in a multivariate fashion. For instance, the addition of a service step (e.g., step $n+1$) in the process does not affect the value of any of the $\left[\mathbf{P}\right]_{i,j}$ elements with $i,j\in\{1,\dots,n\}$, but it typically affects the values of most $\left[\mathbf{P}^{-1}\right]_{i,j}$ elements. Therefore, the $\mathbf{P}^{-1}$ matrix, captures an endogenously arising coupling effect where each of the $\left[\mathbf{P}^{-1}\right]_{i,j}$ elements summarizes the complete interactions of the experiences between steps i and j, after having accounted for all their indirect interactions with all the other service steps in $\mathcal{J}\setminus\{i,j\}$. For a detailed discussion on the interpretation of $\mathbf{P}^{-1}$ see Raveh (1985).[8]

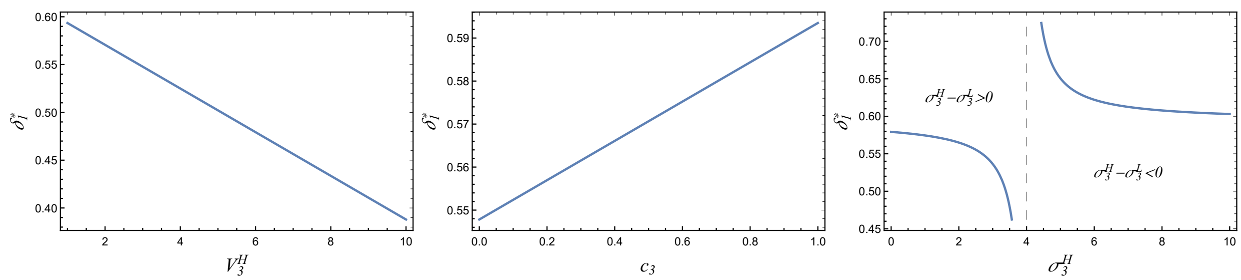
We continue with the general $\mathbf{P}$ structure that comprises the elements $\left[\mathbf{P}\right]_{i,j}=\rho_{i,j}\in(-1,1)$ for $i\neq j$ and $\left[\mathbf{P}\right]_{i,j}=1$ for $i=j$ and we further expound on the role of $\mathbf{P}^{-1}$ in summarizing the independencies of the experiences at the service process level.

[8] To further illustrate the role of $\left[\mathbf{P}^{-1}\right]_{i,j}$ in multivariate analyses, consider the multiple regression model $\tilde{e}_i=\sum_{j\neq i}b_j\tilde{e}_j$. The b_j factors of the regression equation can be estimated as $b_j=-\left[\mathbf{P}^{-1}\right]_{i,j}\Big/\left[\mathbf{P}^{-1}\right]_{i,i}$. Furthermore, the partial correlation $\rho_{i,j\cdot\mathcal{J}\setminus\{i,j\}}$, between steps i and j (i.e., the coupling between steps i and j after accounting for the effect of the remaining steps) is given by $\rho_{i,j\cdot\mathcal{J}\setminus\{i,j\}}=-\frac{\left[\mathbf{P}^{-1}\right]_{i,j}}{\sqrt{\left[\mathbf{P}^{-1}\right]_{i,i}\left[\mathbf{P}^{-1}\right]_{j,j}}}$ (Raveh 1985); note that $\left[\mathbf{P}^{-1}\right]_{i,j}>0(<0)\Leftrightarrow\rho_{i,j\cdot\mathcal{J}\setminus\{i,j\}}<0(>0)$.

Proposition 2. *The provider's effort δ_i^* (δ_j^*) may depend on the characteristics of step j (i) even if the experiences at steps i and j are not directly related as $\rho_{i,j} = 0 \not\Rightarrow \left[\mathbf{P}^{-1}\right]_{i,j} = 0$. Similarly, although the experiences at steps i and j may be positively (negatively) related, δ_i^* and δ_j^* may be negatively (positively) related or not related at all as $\rho_{i,j} > 0 (< 0) \not\Rightarrow \left[\mathbf{P}^{-1}\right]_{i,j} > 0 (< 0)$.*

Proposition 2 sheds more light on the role of the service process in coupling the customer experiences and, as a consequence, the efforts allocated at two different steps. The significance of Proposition 2 stems from the delineation of the importance of $\mathbf{P}^{-1}$ as a key service process metric. Based on that, we find that the provider's effort at a service step may be determined by the characteristics of other steps of the service process even if the experiences at these steps do not directly depend on each other (i.e., $\rho_{i,j} = 0$). Figure 3 offers an illustration of this. It shows that δ_1^* may change considerably with respect to the different characteristics of step 3 despite $\rho_{1,3} = 0$.

Figure 3 Holistic coupling of steps 1 and 3.



(a) δ_1^* vs. V_3^H (b) δ_1^* vs. c_3 (c) δ_1^* vs. σ_3^H

Note. $\rho_{1,2} = 0.15$, $\rho_{1,3} = 0$, $\rho_{1,4} = 0$, $\rho_{2,3} = 0.40$, $\rho_{2,4} = 0$, $\rho_{3,4} = 0.20$, $r = 0.45$, $\boldsymbol{c} = \{0.50, 0.50, c_3, 0.50\}$, $\boldsymbol{V_i^H} = \{6, 5.50, V_3^H, 3.70\}$, $\boldsymbol{V_i^L} = \{1,1,1,1\}$, $\boldsymbol{\sigma_i^H} = \{4, 4, \sigma_3^H, 3.50\}$, and $\boldsymbol{\sigma_i^L} = \{2.50, 2, 4, 4.50\}$. Panel (a): $c_3 = 0.50$ and $\sigma_3^H = 1.50$. Panel (b): $V_3^H = 2$ and $\sigma_3^H = 1.50$. Panel (c): $V_3^H = 2$ and $c_3 = 0.50$.

For the practical implications of Proposition 2, consider the service process shown in Figure 2b. It is reasonable to assume that a patient's experience when inquiring about an upcoming operation (e.g., see the eighth step in Figure 2b) is not directly affected by his experience during check-in at the front desk (e.g., see the third step in Figure 2b), or by waiting to be called in the exam room (e.g., see the fourth step in Figure 2b). However, a patient who experiences a long waiting time, or feels uncomfortable when waiting in a crowded waiting area, may be further frustrated by questions regarding his medical history (e.g., see the sixth step in Figure 2b); he may perceive them to be time consuming and unnecessary. Subsequently, this may affect his ability to ask questions about the operation and receive answers that could alleviate his concerns. From a design point of view, this implies that the efforts determined at the different steps need to account for how the *holistic customer experience* unfolds through the service process. For instance, the effort invested at the check-in can potentially limit the extent to which a prolonged wait creates a negative experience, and subsequently affect the patient's ability to ask questions about the procedure. In that sense, the provider's effort at check-in ought to be determined by the characteristics of the waiting and information inquiry service steps. The reverse influence might also be true; the effort invested at information inquiry and waiting service steps ought to be determined by the effort at check-in. This is the case because, during the design stage, the provider determines the overall effort allocation simultaneously for all steps. In our solution, this process-based (i.e., holistic) coupling of δ_i^* and δ_j^* is captured by the $\left[\mathbf{P}^{-1}\right]_{i,j} = \left[\mathbf{P}^{-1}\right]_{j,i}$ elements of $\mathbf{P}^{-1}$.

In Proposition 2, we also find the opposite effect: although the experiences between two steps i and j may be stochastically related, the corresponding efforts may be independent of the steps' characteristics, implying that the interaction with the rest of the process steps may have a stronger effect in determining δ_i^* and δ_j^* dominating their pairwise interaction. Proposition 1, however, indicates that the pairing of the different step types may have a positive or negative effect on the provider's effort. These findings imply that the effort relationship between δ_i^* and δ_j^* (i.e., whether these efforts are substitutable or complementary) is not straightforward due to the combination of their pairwise interactions and their interactions with the rest of the service steps. Proposition 3 fully characterizes this relationship between δ_i^* and δ_j^* by disentangling these effects. We use $sgn\left[(\cdot)\right]$ to denote the sign of the quantity $(\cdot)$.

Proposition 3. *The characteristics of an individual step i affect δ_i^* according to $\partial\delta_i^*/\partial c_i < 0$, $\partial\delta_i^*/\partial V_i^H > 0$, and $\partial\delta_i^*/\partial\sigma_i^H > 0$ when $\sigma_i^H - \sigma_i^L < 0$ and $\partial\delta_i^*/\partial\sigma_i^H < 0$ when $\sigma_i^H - \sigma_i^L > 0$. They also affect δ_j^* according to $\partial\delta_j^*/\partial c_i < 0$, $\partial\delta_j^*/\partial V_i^H > 0$, $\partial\delta_j^*/\partial\sigma_i^H > 0$ when $\sigma_i^H - \sigma_i^L < 0$ and $\partial\delta_j^*/\partial\sigma_i^H < 0$ when $\sigma_i^H - \sigma_i^L > 0$, iff $sgn\left[\sigma_i^H - \sigma_i^L\right] = sgn\left[\sigma_j^H - \sigma_j^L\right]$ and $\left[\mathbf{P}^{-1}\right]_{i,j} > 0$ or $sgn\left[\sigma_i^H - \sigma_i^L\right] \neq sgn\left[\sigma_j^H - \sigma_j^L\right]$ and $\left[\mathbf{P}^{-1}\right]_{i,j} < 0$. Otherwise, $\partial\delta_j^*/\partial c_i > 0$, $\partial\delta_j^*/\partial V_i^H < 0$, and $\partial\delta_j^*/\partial\sigma_i^H < 0$ when $\sigma_i^H - \sigma_i^L < 0$ and $\partial\delta_j^*/\partial\sigma_i^H > 0$ when $\sigma_i^H - \sigma_i^L > 0$.*

As expected, the provider decreases her step effort investment in the cost of the step in order to ensure the profitability of the service. This result holds for any type of step, routine or non-routine. Similarly, larger maximum functional values (V_i^H) regardless of a step's type, allow for larger value contribution to the customer utility, and therefore, push the provider to invest more effort.

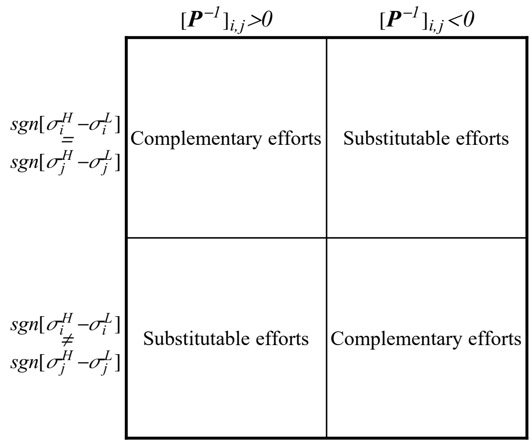
The effect of the variability, captured via σ_i^H, depends on the type of the step. For non-routine steps (i.e., $\sigma_i^H - \sigma_i^L > 0$), larger values of σ_i^H imply that more effort can further increase the variability in the customer experience. For that reason, the provider decreases her effort investment. The opposite is true for routine steps (i.e., $\sigma_i^H - \sigma_i^L < 0$) where larger values of σ_i^H imply that more effort can lead to even lower variability in the experience and further increase customer utility.

Most importantly, Proposition 3 indicates when the different service step and process characteristics, induce efforts that are complementary or substitutable. This is a valuable design insight as it identifies a key directional implication for designers to consider as they determine the emphasis put, from an effort standpoint, on the different steps: it is imperative to recognize whether these efforts work synergistically towards the overall customer experience. If not, effort should be allocated so that it mitigates the antagonistic effects across the steps. In practical terms, Proposition 3 delineates specific conditions under which providers benefit from designs that spread out the effort across steps (i.e., when efforts are complementary) versus placing emphasis on certain steps at the expense of others (i.e., when efforts are substitutable).

The directional relationship between δ_i^* and δ_j^* is determined by the combined effect of $(\sigma_i^H - \sigma_i^L)(\sigma_j^H - \sigma_j^L)$ and $\left[\mathbf{P}^{-1}\right]_{i,j} > 0$; in particular, by how this effect influences the variability of the overall customer experience. The quantity $\left[\mathbf{P}^{-1}\right]_{i,j}$ captures the way that the experiences at two steps relate to each other after controlling for the effect of the remaining steps of the service process; we remind the reader that $\left[\mathbf{P}^{-1}\right]_{i,j} > 0 (< 0) \Leftrightarrow \rho_{i,j \cdot \mathcal{J} \setminus \{i,j\}} < 0 (> 0)$. For instance, $\left[\mathbf{P}^{-1}\right]_{i,j} > 0$ indicates that, after controlling for the remaining steps, the experiences at steps i and j have an inverse relationship in that a positive experience at one step is more likely to be accompanied by a negative experience at the other step. Hence, $\left[\mathbf{P}^{-1}\right]_{i,j} > 0$ brings a "balancing" effect that lowers the overall variability. In order for the provider to fully benefit from this balancing, the effort investments should ensure that the ranges of the possible experience realizations across the different steps are similar. Hence, any changes in δ_i^* that affect the variability at step i towards a specific direction (e.g., they increase σ_i) should be accompanied by changes in δ_j^* so that the variability at step j changes towards the same direction (e.g., they also increase σ_j). Along similar lines, $\left[\mathbf{P}^{-1}\right]_{i,j} < 0$ gives rise to an amplification of the experiences realized during the service; positive experiences drive more positive realizations, whereas negative ones cause a cascading negative effect. The provider mitigates such an amplification by investing effort at one step that offsets the change in variability at other steps. If both steps are of the same type, the provider accompanies a high effort at step i with a low effort at step j, and if they have different types she accompanies a high effort at step i with a high effort at step j too. Figure 4 summarizes these findings.

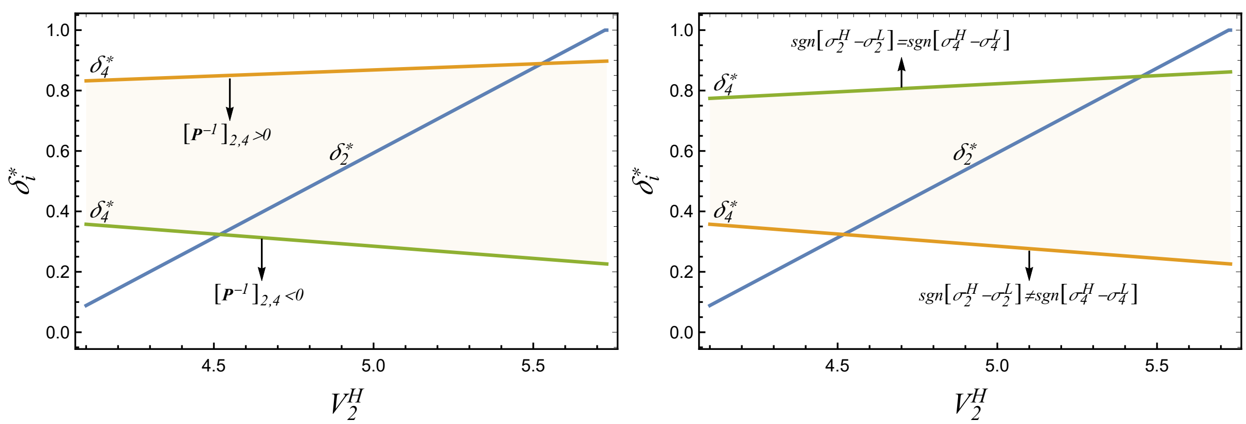
To further illustrate, in Figure 5a we consider a four-step process and we focus on the efforts invested at steps 2 and 4, which are both non-routine. When $\left[\mathbf{P}^{-1}\right]_{2,4} > 0$, larger values of V_2^H not only incentivize the provider to allocate more effort δ_2^* at step 2, but they also increase the relative marginal value of effort at step 4 and therefore, the provider increases δ_4^* too. By doing so, she increases the functional value V_4 that customers derive at step 4 and at the same time contains the overall variability of the experience. On the contrary, when $\left[\mathbf{P}^{-1}\right]_{2,4} < 0$, larger values

Figure 4 **Complementary versus substitutable efforts at steps $i, j \in \mathcal{J}$.**



Note. $sgn\left[\sigma_i^H - \sigma_i^L\right] = sgn\left[\sigma_j^H - \sigma_j^L\right]$ implies same step types, whereas $sgn\left[\sigma_i^H - \sigma_i^L\right] \neq sgn\left[\sigma_j^H - \sigma_j^L\right]$ implies different types. We remind the reader that $\left[\mathbf{P}^{-1}\right]_{i,j} > 0 (< 0) \Leftrightarrow \rho_{i,j \cdot \mathcal{J} \setminus \{i,j\}} < 0 (> 0)$.

Figure 5 **Optimal effort δ_i^* with respect to different holistic and step type couplings.**



(a) $sgn\left[\sigma_2^H - \sigma_2^L\right] = sgn\left[\sigma_4^H - \sigma_4^L\right]$ (b) $\left[\mathbf{P}^{-1}\right]_{2,4} > 0$

Note. $\rho_{1,2} = 0.10$, $\rho_{1,3} = 0.50$, $\rho_{1,4} = 0.10$, $\rho_{2,3} = 0.40$, $\rho_{3,4} = 0.20$; selected $\rho_{i,j}$ values ensure positive definite $\mathbf{P}$, $r = 0.45$, $\boldsymbol{c} = \{0.20, 0.70, 0.80, 0.62\}$, $\boldsymbol{V_i^H} = \left\{3, V_2^H, 3, 3\right\}$, $\boldsymbol{V_i^L} = \{1,1,1,1\}$, $\boldsymbol{\sigma_i^H} = \left\{4.50, \sigma_2^H, 1.45, 2.90\right\}$, and $\boldsymbol{\sigma_i^L} = \{2.50, 2.80, 2.70, 2.20\}$. Panel (a): $\rho_{2,4} = 0.04$ resulting in $\left[\mathbf{P}^{-1}\right]_{2,4} = 0.05$ ($\rho_{2,4} = 0.16$ implies $\left[\mathbf{P}^{-1}\right]_{2,4} = -0.102$), and $\sigma_2^H = 4$. Panel (b): $\rho_{2,4} = 0.04$, and $\sigma_2^H = 1$.

of V_2^H decrease the relative marginal value of effort at step 4 as higher δ_4^* results in higher overall variability. Hence, the provider decreases δ_4^*. Similar observations can be made in Figure 5b.

It is evident from Figure 5 that the types of the steps, and/or their holistic coupling can lead the designers to propose significantly different designs. For instance, in Figure 5b we see that for large values of V_2^H the provider exerts high effort when both steps are routine or non-routine. However, if the steps have different types, the optimal efforts are diametrically different. This finding exposes the intricate subtleties of service design. Often in service systems, providers strive to achieve operational performance (e.g., waiting time), implicitly assuming that most steps are routine. However, even the presence of one non-routine step might be enough to warrant a fundamentally different design.

The previous results were derived by considering a generic structure of $\mathbf{P}$. Hereafter, we characterize the optimal design considering more detailed structural properties of the service process. We consider service settings where the customer experience unfolds according to a generalized first-order autoregressive model known as first-order antedependent ($AD(1)$) transition model. Under this model, the customer experience unfolds in a serial manner and each experience, given exactly one immediately preceding experience, is independent of all further preceding experiences. Such structures tend to capture most typical serial service settings like the ones depicted in Figure 1 and 2. The corresponding $\mathbf{P}$ comprises the elements $\left[\mathbf{P}\right]_{i,j} = \prod_{m=j}^{i-1} \rho_{m+1,m} \in (-1,1)$ for $i \neq j$ and $\left[\mathbf{P}\right]_{i,j} = 1$ for $i = j$ (see Gabriel 1962, Zimmerman and Núñez-Antón 2009).

Proposition 4. *When the customer experience unfolds according to an $AD(1)$ process, the optimal effort at step i is determined by its own characteristics and the characteristics of only the immediately adjacent steps.*

When the customer experience unfolds according to an $AD(1)$ process, the provider's effort at a service step is determined by its own characteristics and the characteristics of the steps immediately before and after it. Although Proposition 4 points to the sufficiency of the local input to determine the optimal effort at a service step, this result is the global optimal outcome from a holistic treatment of the customer experience. From a technical point of view, this happens because $\mathbf{P}^{-1}$, which summarizes a service process effect, is banded with $\left[\mathbf{P}^{-1}\right]_{i,j} = 0$ if $j < i-1$ or $j > i+1$.

From a managerial point of view, service designers can tackle the complexity of the multi-dimensional design challenge through more manageable three-dimensional challenges at a time. This is particularly useful in service processes with a large number of steps. Proposition 4 also implies that deficiencies or limitations at a step are best to be addressed through the effort exerted at the immediately adjacent steps. For instance, in the context of the patient service shown in Figure 2b, issues affecting the customer experience during waiting (e.g., limited waiting space in the fourth step of Figure 2b) can be compensated through more effort in ensuring a streamlined check-in process (e.g., see the third step in Figure 2b) or in the design of the exam room (e.g., see the fifth step in Figure 2b); Arieff (2009) provides details on how Kaiser Permanente implemented such design changes in order to improve patients' "Total Health Journey."

The $AD(1)$ process implies that, given all the experiences that a customer may realize at steps 1 through $i-1$ of a service, the experience at step i depends only on the experience at step $i-1$. Nonetheless, when determining δ_i^*, the service provider factors in the characteristics of both steps $i-1$ and $i+1$. As we have noted before, at the design stage the provider determines her effort allocation holistically. That is, when determining δ_i^* the provider accounts for the way the experience may unfold both before and after step i. This impact zone around step i is bounded by the immediately adjacent steps. For higher-order $AD\,(p)$ models with $p>1$, this zone expands accordingly; we elaborate in the following corollary.

COROLLARY 1. *When the customer experience unfolds according to an $AD\,(p)$ process, the optimal effort at step i is determined by its own characteristics and the characteristics of the p steps immediately before and after it.*

The $AD\,(p)$ process implies that, given the experiences that a customer may have realized at steps 1 through $i-1$ of a service, the experience at step i depends only on the experiences at steps $i-p$ through $i-1$. Although, a generic closed-form expression is not available for $\mathbf{P}^{-1}$ when $p>1$, Zimmerman and Núñez-Antón (2009) establish that the covariance matrix maintains the distinct banded structure for all $j<i-p$ and $j>i+p$. Based on our discussion under Proposition 4, this banded structure implies that when determining δ_i^*, the provider should factor in how the experience may unfold from step $i-p$ through step $i+p$.

In the rest of the paper, we maintain our focus on first-order transition models as a means of capturing how the customer experience unfolds throughout the service process. To facilitate the analysis at the service process level, however, we consider the case where $\left[\mathbf{P}\right]_{i,j}=\prod_{m=j}^{i-1}\rho_{m+1,m}$ with $\rho_{m+1,m}=\rho\in(0,1)$ for any $m\in\mathcal{J}$, which simplifies to $\left[\mathbf{P}\right]_{i,j}=\rho^{|i-j|}\in(0,1)$. In this case, $\rho_{i,j}$ decreases geometrically in the number of steps between i and j and $\mathbf{P}$ takes the structure of a Kac-Murdock-Szergö matrix; see Kac et al. (1953). This structure of $\mathbf{P}$ is characteristic of a special case of an $AD\,(1)$, known as heterogeneous first-order autoregressive model ($ARH\,(1)$; see Wolfinger 1996, Zimmerman and Núñez-Antón 2009, IBM Knowledge Center 2019).

We adopt this simplification to analyze how the provider's efforts depend on the perceived *cohesiveness* of the customer experiences (Ariely and Zauberman 2003) across the different service steps. We define cohesiveness as the extent to which customers perceive the different steps of the service as independent of (or dependent on) each other in forming their overall experience. We capture cohesiveness through the magnitude of ρ. Cohesiveness may arise by the nature of the service. For example, customers dining at a family-style restaurant likely view the different steps of the process as serving fairly distinct roles (e.g., providing parking availability, waiting space, or food options that meet different dietary restrictions, payment options, etc.). Such a view translates to smaller values of ρ compared to the view that customers may have when dining at a theme restaurant. In this case, it is more likely that the different steps are considered to be jointly serving the goal of immersing customers to the specific theme. For instance, use of modern equipment (e.g., tablet computers) for order placement at a Wild West-themed restaurant may be perceived negatively and prompt customers to judge the whole service as "lacking authenticity." Dixon and Verma (2013) also link cohesiveness to the thematic similarity of different service encounters.

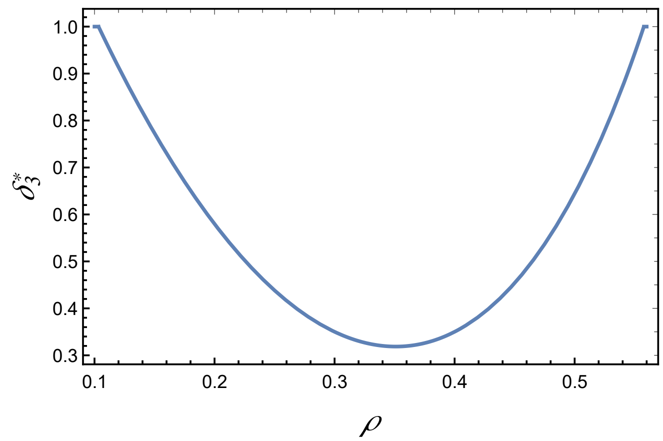
In the following, we define $\Delta_i \doteq \frac{(1-c_i)\left(V_i^H - V_i^L\right)}{r\left(\sigma_i^H - \sigma_i^L\right)}$. The sign of the quantity $\Delta_{i-1} + \Delta_{i+1}$ can be thought of as capturing the effective type of the steps immediately adjacent to step i.

Proposition 5. *When the customer experience unfolds according to an $ARH(1)$ process, the optimal effort at step i depends on the service cohesiveness (ρ) as follows: for service steps with $sgn[\Delta_i] \neq sgn[\Delta_{i-1} + \Delta_{i+1}]$, $\partial\delta^*/\partial\rho > 0$ for any $\rho \in (0,1)$, whereas for steps with $sgn[\Delta_i] = sgn[\Delta_{i-1} + \Delta_{i+1}]$, $\partial\delta^*/\partial\rho \leq 0$ for any $\rho \in (0, \bar{\rho}]$ and $\partial\delta^*/\partial\rho > 0$ for any $\rho \in (\bar{\rho}, 1)$.*

In an $ARH(1)$ process, the service steps tend to generate similar (e.g., satisfying or dissatisfying) experiential outcomes; this is due to $\left[\mathbf{P}^{-1}\right]_{i,j} < 0$, which implies that $\rho_{i,j\cdot\mathcal{J}\setminus\{i,j\}} > 0$ for all $i, j \in \mathcal{J}$. This property is more pronounced in services with more cohesive experiences (i.e., in services with large ρ). Said differently, the strong dependence between experiences turns the overall experience to be either very satisfying or very dissatisfying. Such uniform pattern of customer experience negatively affects customers' valuation of the service due to their risk aversion. In designing the service, the provider accounts for this effect and decides on an effort allocation, which serves as a means of containing the overall variability, and/or providing more functional value to the customers. When $sgn[\Delta_i] \neq sgn[\Delta_{i-1} + \Delta_{i+1}]$, the effective relationship of step i with steps $i-1$ and $i+1$ is complementary. Based on Proposition 3, by exerting more effort at step i, the provider contributes more functional value without increasing the overall experiential variability. This is not the case when $sgn[\Delta_i] = sgn[\Delta_{i-1} + \Delta_{i+1}]$, as then, the efforts at step i, and steps $i-1$ and i are substitutable. Hence, as per Proposition 3, the provider limits the increase in variability stemming from the stronger cohesiveness through less effort investment. However, beyond a certain threshold of cohesiveness $\bar{\rho}$ (analytical expressions provided in the Appendix; see also Figure 6), the decrease of the functional value as a mitigation design strategy to contain the variability in the customer experience no longer pays off; instead the provider should increase the customer utility by increasing the offered functional value despite the loss from the more variable experience. Said differently, the provider's design approach switches from variability mitigation to core value delivery.

From a managerial viewpoint, Proposition 5 prescribes that the design of services with highly interdependent experiences does not imply autopilot approaches where the provider invests effort only at certain steps assuming that the experiences at the rest of them will follow accordingly. On the contrary, we find that for such services the provider might need to invest more effort across all service steps. For services with less interdependent experiences, the provider's effort at each step

Figure 6 **Optimal effort with respect to the cohesiveness (i.e., correlation ρ) of the experiences.**



Note. $\boldsymbol{c} = \{0.50, 0.50, 0.50, 0.50\}$, $\boldsymbol{V_i^H} = \{3, 3, 3, 3\}$, $\boldsymbol{V_i^L} = \{1, 1, 1, 1\}$, $\boldsymbol{\sigma_i^H} = \{4.50, 4, 3.30, 3\}$, $\boldsymbol{\sigma_i^L} = \{2.50, 2.80, 2.70, 2.20\}$, $r = 0.45$, $\Delta_i \doteq \frac{(1-c_i)\left(V_i^H - V_i^L\right)}{r\left(\sigma_i^H - \sigma_i^L\right)}$ result in $\Delta_2 = 1.20$, $\Delta_3 = 0.60$, and $\Delta_4 = 0.80$.

depends on its type, which, once again, points to the importance of identifying and accounting for the different types of steps that the service process comprises.

Our next finding identifies a mechanism that can help the service provider determine whether she should exert more effort at the initial or final moments of the service. In particular, we characterize the conditions under which the provider exerts more effort at the first than the last step, and vice versa.[9] We consider service processes where the first and last steps are of the same type and serve similar functions (i.e., $\Delta_1 = \Delta_n$), and processes where the first and last steps are of different type and/or serve different functions (i.e., $\Delta_1 \neq \Delta_n$); $\Delta_i \doteq \frac{(1-c_i)\left(V_i^H - V_i^L\right)}{r\left(\sigma_i^H - \sigma_i^L\right)}$, all analytical expressions are provided in the Appendix.

Proposition 6. *When the customer experience unfolds according to an ARH*(1) *process and* $\Delta_1 = \Delta_n$*, the provider allocates* $\delta_n^* > \delta_1^*$ *iff: i) both steps are non-routine and* $\Delta_{n-1} < \Delta_2$*, or ii) both steps are routine and* $\Delta_{n-1} > \Delta_2$*. Otherwise,* $\delta_n^* > \delta_1^*$*. For services with* $\Delta_1 = -\Delta_n$*, the provider allocates* $\delta_n^* > \delta_1^*$ *iff: i)* $\sigma_n^H - \sigma_n^L > 0$ *and* $\Delta_n > \bar{\Delta}_n$*, or ii)* $\sigma_n^H - \sigma_n^L < 0$ *and* $\Delta_n < \bar{\Delta}_n$*. Otherwise, she allocates* $\delta_n^* > \delta_1^*$*.*

We find that lack of differentiation (in terms of the Δ_1 and Δ_n values) between the first and last step does not imply that the provider exerts the same effort at both service steps. The optimal

[9] We should note that previous research has found that for highly cohesive services, where the experience may even be perceived as continuous (e.g., during a specific medical procedure; Kahneman et al. 1993, Ariely and Carmon 2000), additional considerations such as the overall trend of the experience, sequence effects or behavioral traits, may affect the customers' valuation of the service. For instance, Crano (1977) focuses on the primacy effect, which advocates that greater weight should be placed on the experiences that take place early on, whereas Kahneman et al. (1993) identify the "peak-end" rule, which advocates for the importance of the final moments. Ariely and Zauberman (2003) find that settings with partitioned experiences (e.g., experiences generated at discrete service steps), which are the main focus of our paper, are less influenced by such effects.

efforts at the two steps are driven by the characteristics of the immediately adjacent steps (i.e., by the values of Δ_2 and Δ_{n-1}). The first and last step of the service may have the same values of Δ_1 and Δ_n when they both serve very similar functions (e.g., check-in/out) and there are no behavioral biases accentuating the importance of one over the other. This may also be the case when the steps have similar characteristics but different types (i.e., similar c_i, V_i^H, V_i^L, $|\sigma_i^H - \sigma_i^L|$, but $sgn\,[\sigma_1^H - \sigma_1^L] \neq sgn\,[\sigma_n^H - \sigma_n^L]$). For instance, although when arriving at a medical facility for an outpatient procedure a customer may evaluate the check-in process along intangible dimensions (e.g., perceived friendliness/professionalism of the front desk personnel), when leaving, he may view the similar (check-out) process on a more transactional basis (e.g., accuracy of the charges on the insurance, speed of the process). In this case, $|\Delta_1| = |\Delta_n|$, but $\Delta_1 \neq \Delta_n$ due to the different step types $\sigma_1^H - \sigma_1^L > 0$, and $\sigma_n^H - \sigma_n^L < 0$. It is straightforward to show that if $\Delta_2 > 0$ and $\Delta_{n-1} < 0$, the provider optimally exerts more effort at the last step despite it being viewed by the customer as more transactional in nature. Our insights in Proposition 6 apply to service contexts with partitioned experiences and as such, they can be viewed as complementary to the insights offered by previous research on sequence effects during continuous experiences.

5. Discussion

Starting in 1996, with IDEO's engagement by Amtrak to conduct what is known as the first service design project, to 2001, when Livework became in London the first design firm to focus exclusively on service design (Fayard et al. 2017), and beyond, there has been a marked increase in the number of firms specializing in the design of services. Consulting companies like Accenture and Deloitte have also expanded their activities in this domain through the acquisition of service design firms (Fjord 2019, Doblin 2019). They have allowed these units to operate independently, recognizing their unique approach to the design of services. In the academic space, a growing number of schools have incorporated the discipline of service design in their curricula as stand-alone degree offerings or as part of graduate programs on design thinking. As of 2019 examples of such schools were Savannah College of Art and Design, Royal College of Art, Carnegie Mellon, MIT Sloan, Insead, and Köln International School of Design.

The nascent practice of service design has been recognized as being multidisciplinary in nature, drawing on knowledge developed in fields as diverse as marketing, human-computer interaction, and anthropology to mention a few. Despite this multidisciplinarity, Fayard et al. (2017) find two important values that define the service design work ethos and differentiate its practice from other professional practices: i) the emphasis on a holistic approach, and ii) the focus on empathy. Practitioners of service design enact these values through a variety methodologies such as shadowing, and contextual interviews (see Fayard et al. 2017 for an overview), which for the most part culminate to the creation of customer journey maps that visualize customers' interactions with the service provider and delineate the service delivery process.

This customer journey mapping recognizes the multi-stage nature of services and treats the customer experience holistically. It also offers a solid conceptual basis to develop an analytical approach for service design. We consider that to receive a service, customers go through a process with several distinct service steps. At each step, they derive: i) a functional value, which is, the tangible benefit that the step is designed to deliver, and ii) a customer experience, which is uncertain and may have positive or negative realizations. Given this context, we account for the adverse effect of variability on the overall customer experience. Our analysis outlines the optimal design in terms of the provider's effort allocation across the steps of the process. We distinguish between two types of steps: i) routine and ii) non-routine. In non-routine service steps, more provider effort increases the functional value but may also lead to a more variable customer experience. In routine steps, more effort increases the functional value and leads to lower experiential variability. More importantly, though, we explicitly account for the structural interdependencies of the service experiences across the different steps. To our knowledge, this is the first such formal treatment.

We generate insights that address the design of single service steps based on the influences of the entire service process. In other words, we analyze the effects of the characteristics of the service process on the optimal design, i.e., how the customer experiences, and as an extension, the provider's efforts interact across the entire service process. Our analysis quantifies the holistic coupling of the experiences across the different steps as a subtle but also critical design factor. This coupling is process-based in the sense that it captures the interplay of the experiences at two different steps after accounting for the effect of the rest of the process. We find that the experiences at two steps may depend on each other even if this is not expected a priori. The relationship between this coupling and the types of the different steps determines whether the provider's efforts across the service steps are complementary or substitutable. Depending on this relationship, the service design may be characterized by a "spread-out of the effort" approach, versus an approach that places emphasis on certain steps at the expense of others. In that light, misidentifying the type of the steps or their coupling can result in significantly different designs, characterized by over- or under-invested efforts.

To offer more structural insights about the effect of the entire process, we consider the broad class of services where the customer experience unfolds according to a stochastic process exhibiting serial correlation that decays as the number of steps between experiences increases. Interestingly, we show that the design of each step is determined by its type and the types of the steps within a certain impact zone around it. This is of great value to the designers as it reduces the design complexity, and allows them to treat the service process as modules of such zones. We also find that services with highly interdependent experiences do not imply an autopilot design approach, where the provider invests effort only at certain steps and assumes that the experiences at the rest of them will follow accordingly. Instead, we find that in such services the provider invests more effort at all service steps. Finally, we characterize the conditions under which the provider prefers to invest more effort at the initial than the final moments of the service and vice versa. Our findings complement previous literature, which has focused primarily on the effect of behavioral phenomena in the design of services with highly continuous experiences.

Our proposed approach is by no means exhaustive or all-encompassing of the challenges pertinent to the design of services. It offers a starting point, which is based on the application of newfound approaches to the practice of service design. Best practices in service design take the iterative form of "observation-visualization-prototyping" (Fayard et al. 2017). Our paper is motivated by the visualization practices; therefore, further analytical or empirical research on the practices of observation and prototyping (Thomke 2003) can offer valuable insights.

As we have noted before, our model relies on a number of assumptions. Relaxing these assumptions presents promising opportunities for future research into the area of service design. For instance, our focus has been on the design of the service process and for that reason, we have treated the provider's efforts as non-discretionary. During the execution of the service, however, and depending on the service context, an unsatisfying experience at a service step may trigger customer abandonment. This may create the need for the provider to strategically place "check-points" where she solicits feedback from the customer and possibly exerts additional effort to ensure that the customer remains in the process. Our model does not account for the role of employees (Tan and Netessine 2019) or employee engagement/satisfaction (Heskett et al. 1994) on the customer evaluation of the service. Determining the extent of latitude that a provider allows her employees to exercise in dealing with unforeseen situations is equivalent to determining the reactive capacity of the service process and directly affects the provider's ability to turn customer disappointment into delight (Thomke 2019). Furthermore, given the multi-stage nature of services, the design of incentive mechanisms is particularly important as the efforts of an employee at one stage may affect the performance of an employee at a different service stage. We also do not consider externalities across customers and how the provider can design a service that enhances the positive externalities and mitigates the negative externalities. Finally, we perform our analysis without any explicit consideration of the strategic interactions under competitive pressure. Accounting for the effects of competition on the design decisions is another promising direction of future research.

We hope that the approach proposed in this paper will spark interest for future research on the holistic design of services; after all, "Everybody is in service" (Levitt 1972).

Appendix

Proof of Proposition 1. We start by deriving the optimal price for a given $\boldsymbol{\delta} = \{\delta_1, \ldots, \delta_n\}$ and then we determine the optimal effort allocation. Customers' net utility is linear decreasing in π. Hence, for a given $\boldsymbol{\delta}$, the optimal price $\tilde{\pi}$ is based on the lower bound we obtain after rewriting the profitability constraint as $\pi \geq \pi_0 + \sum_{i=1}^{n}\Big(\delta_i V_i^H + (1-\delta_i)\, V_i^L\Big) c_i$, that is, $\tilde{\pi}(\boldsymbol{\delta}) \doteq \pi_0 + \sum_{i=1}^{n}\Big(\delta_i V_i^H + (1-\delta_i)\, V_i^L\Big) c_i$. The provider determines the optimal effort allocation by maximizing $\tilde{\mathcal{U}}(\boldsymbol{\delta}) \doteq \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_{j,i}\Big(\delta_j V_j^H + (1-\delta_j)\, V_j^L\Big) - \frac{r}{2}\Bigg(\sum_{i=1}^{n}\Big(\delta_i \sigma_i^H + (1-\delta_i)\,\sigma_i^L\Big)^2 + 2\sum\sum_{i<j} \rho_{i,j}\Big(\delta_i \sigma_i^H + (1-\delta_i)\,\sigma_i^L\Big)\Big(\delta_j \sigma_j^H + (1-\delta_j)\,\sigma_j^L\Big)\Bigg) - \tilde{\pi}(\boldsymbol{\delta})$ subject to $\delta_i \in (0,1)$ for all $i \in \mathcal{J}$. We state the Hessian matrix, $\boldsymbol{H_n}$, of $\tilde{\mathcal{U}}(\boldsymbol{\delta})$ and the i^{th} principal minor of $|\boldsymbol{H_n}|$ $\forall i \in \mathcal{J}$ as

$$\boldsymbol{H_n} = \begin{bmatrix} \frac{\partial^2 \tilde{\mathcal{U}}}{\partial^2 \delta_1} & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_1 \partial \delta_2} & \cdots & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_1 \delta_n} \\ \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_2 \partial \delta_1} & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial^2 \delta_2} & \cdots & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_2 \delta_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_n \partial \delta_1} & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_n \delta_2} & \cdots & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial^2 \delta_n} \end{bmatrix} \text{ and } |\boldsymbol{H_i}| = \begin{vmatrix} \frac{\partial^2 \tilde{\mathcal{U}}}{\partial^2 \delta_1} & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_1 \partial \delta_2} & \cdots & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_1 \delta_i} \\ \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_2 \partial \delta_1} & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial^2 \delta_2} & \cdots & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_2 \delta_i} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_i \partial \delta_1} & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial \delta_i \delta_2} & \cdots & \frac{\partial^2 \tilde{\mathcal{U}}}{\partial^2 \delta_i} \end{vmatrix},$$

respectively. We use i) $\mathbf{P}[1:i;1:i]$ to indicate the principal sub-matrix that comprises rows 1 to i and columns 1 to i of the matrix $\mathbf{P}$ and ii) $\mathbf{P}[-i;-j]$ to indicate the matrix that results after removing row i and column j from $\mathbf{P}$, $\forall i,j \in \{1,\ldots,n\}$. The covariance matrix is always positive definite, therefore, $\mathbf{P}$ is also positive definite. The latter implies that the determinant $\big|\mathbf{P}\big|$ and every principal minor, $\big|\mathbf{P}[1:i;1:i]\big|$, of $\mathbf{P}$ is positive. Hence, we can now show that:

$$\begin{aligned} |\boldsymbol{H_1}| &= -r\left(\sigma_1^H - \sigma_1^L\right)^2 < 0 \\ |\boldsymbol{H_2}| &= r^2\left(\sigma_1^H - \sigma_1^L\right)^2\left(\sigma_2^H - \sigma_2^L\right)^2\big|\mathbf{P}[1:2;1:2]\big| > 0 \\ &\vdots \\ |\boldsymbol{H_i}| &= (-1)^i\, r^i \underbrace{\prod_{j=1}^{i}\left(\sigma_j^H - \sigma_j^L\right)^2\big|\mathbf{P}[1:i;1:i]\big|}_{>0,\text{ since }\mathbf{P}[1:i;1:i]\text{ is positive definite}}, \end{aligned}$$

which establishes that $\tilde{\mathcal{U}}(\boldsymbol{\delta})$ is strictly concave in $\boldsymbol{\delta}$.

Given that $\tilde{\mathcal{U}}(\boldsymbol{\delta})$ is strictly concave in $\boldsymbol{\delta}$ and the constraints are linear in $\boldsymbol{\delta}$, the provider's maximization problem forms a convex program. This formulation gives rise to a total of 3^n different effort allocations. Since all of our constraints are symmetric and each of them involves only one decision variable, δ_i, we can facilitate our analysis by focusing on the case where the optimal solutions lie in the interior that is, $\delta_i^* \in (0,1)$ for all $i \in \mathcal{J}$. To do so, we solve the system of the first-order conditions $\partial\tilde{\mathcal{U}}(\boldsymbol{\delta})/\partial\delta_1 = 0, \ldots, \partial\tilde{\mathcal{U}}(\boldsymbol{\delta})/\partial\delta_n = 0$, where $\partial\tilde{\mathcal{U}}(\boldsymbol{\delta})/\partial\delta_i = \left(1-\left(c_i - \sum_{m\neq i}\beta_{i,m}\right)\right)(V_i^H - V_i^L) - r(\sigma_i^H - \sigma_i^L)\sum_{j=1}^n \left(\delta_j\left(\sigma_j^H - \sigma_j^L\right) + \sigma_j^L\right)\rho_{i,j}$ for any $i \in \mathcal{J}$, with respect to $\boldsymbol{\delta}$, which returns the efforts

$$\hat{\delta}_i \doteq \sum_{j=1}^n \frac{\left(1-\left(c_j - \sum_{m\neq j}\beta_{j,m}\right)\right)\left(V_j^H - V_j^L\right)}{r\left(\sigma_i^H - \sigma_i^L\right)\left(\sigma_j^H - \sigma_j^L\right)}\left[\mathbf{P}^{-1}\right]_{i,j} - \frac{\sigma_i^L}{\sigma_i^H - \sigma_i^L} \text{ for all } i \in \mathcal{J}.$$

By differentiating the $\hat{\delta}_i$ efforts, we obtain $\frac{\partial\hat{\delta}_i}{\partial\left(V_i^H - V_i^L\right)} = \frac{\left(1-\left(c_i-\sum_{m\neq i}\beta_{i,m}\right)\right)\left|\mathbf{P}\left[-i;-i\right]\right|}{r\left(\sigma_i^H-\sigma_i^L\right)^2\left|\mathbf{P}\right|} > 0$, which implies that for each step i, $\hat{\delta}_i \leq 0$ iff $V_i^H - V_i^L \leq \underline{\Delta V}_i \doteq \left\{V_i^H - V_i^L : \hat{\delta}_i = 0\right\} = \frac{r\left(\sigma_i^H-\sigma_i^L\right)\left|\mathbf{P}\right|}{\left(1-\left(c_i-\sum_{m\neq i}\beta_{i,m}\right)\right)\left|\mathbf{P}\left[-i;-i\right]\right|}\left(\sigma_i^L - \sum_{j\in\mathcal{J}\setminus i}^n \frac{\left(1-\left(c_j-\sum_{m\neq j}\beta_{j,m}\right)\right)\left(V_j^H-V_j^L\right)(-1)^{i+j}\left|\mathbf{P}\left[-j;-i\right]\right|}{r\left(\sigma_j^H-\sigma_j^L\right)\left|\mathbf{P}\right|}\right)$.
Similarly, for each step i, $\hat{\delta}_i \geq 1$ iff $V_i^H - V_i^L \geq \overline{\Delta V}_i \doteq \left\{V_i^H - V_i^L : \hat{\delta}_i - 1 = 0\right\} = \frac{r\left(\sigma_i^H-\sigma_i^L\right)\left|\mathbf{P}\right|}{\left(1-\left(c_i-\sum_{m\neq i}\beta_{i,m}\right)\right)\left|\mathbf{P}\left[-i;-i\right]\right|}\left(\sigma_i^H - \sum_{j\in\mathcal{J}\setminus i}^n \frac{\left(1-\left(c_j-\sum_{m\neq j}\beta_{j,m}\right)\right)\left(V_j^H-V_j^L\right)(-1)^{i+j}\left|\mathbf{P}\left[-j;-i\right]\right|}{r\left(\sigma_j^H-\sigma_j^L\right)\left|\mathbf{P}\right|}\right)$, where $\overline{\Delta V}_i - \underline{\Delta V}_i = \frac{r\left(\sigma_i^H-\sigma_i^L\right)^2\mathbf{P}}{\left(1-\left(c_i-\sum_{m\neq i}\beta_{i,m}\right)\right)\mathbf{P}[-i;-i]} > 0$. Hence, $V_i^H - V_i^L \in \left(\underline{\Delta V}_i, \overline{\Delta V}_i\right)$, ensures that $\hat{\delta}_i \in (0,1)$ and therefore, the optimal effort allocation is given by $\delta_i^* = \hat{\delta}_i$ for all $i \in \mathcal{J}$.□

Proof of Proposition 2. To show that $\rho_{i,j} = 0 \nRightarrow \left[\mathbf{P}^{-1}\right]_{i,j} = 0$ consider $\mathbf{P} = \begin{bmatrix} 1 & \rho_{1,2} & 0 \\ \rho_{1,2} & 1 & \rho_{2,3} \\ 0 & \rho_{2,3} & 1 \end{bmatrix}$, where $\rho_{1,3} = 0$ and $\rho_{1,2}, \rho_{1,3} \in (-1,1)$. In this case, $\mathbf{P}$ is positive definite, and as such a valid correlation matrix, iff $|\mathbf{P}| = 1 - \rho_{1,2}^2 - \rho_{2,3}^2 > 0$ (Rousseeuw and Molenberghs 1994). The inverse of $\mathbf{P}$ is given by $\mathbf{P}^{-1} = \begin{bmatrix} \frac{1-\rho_{2,3}^2}{1-\rho_{1,2}^2-\rho_{2,3}^2} & -\frac{\rho_{1,2}}{1-\rho_{1,2}^2-\rho_{2,3}^2} & \frac{\rho_{1,2}\rho_{2,3}}{1-\rho_{1,2}^2-\rho_{2,3}^2} \\ -\frac{\rho_{1,2}}{1-\rho_{1,2}^2-\rho_{2,3}^2} & \frac{1}{1-\rho_{1,2}^2-\rho_{2,3}^2} & -\frac{\rho_{2,3}}{1-\rho_{1,2}^2-\rho_{2,3}^2} \\ \frac{\rho_{1,2}\rho_{2,3}}{1-\rho_{1,2}^2-\rho_{2,3}^2} & -\frac{\rho_{2,3}}{1-\rho_{1,2}^2-\rho_{2,3}^2} & \frac{1-\rho_{1,2}^2}{1-\rho_{1,2}^2-\rho_{2,3}^2} \end{bmatrix}$ from which we can see that: i) $\left[\mathbf{P}^{-1}\right]_{1,3} = \frac{\rho_{1,2}\rho_{2,3}}{1-\rho_{1,2}^2-\rho_{2,3}^2} > 0$ if $\rho_{1,2}, \rho_{2,3} \in (-1,0)$, or $\rho_{1,2}, \rho_{2,3} \in (0,1)$, and ii) $\left[\mathbf{P}^{-1}\right]_{1,3} = \frac{\rho_{1,2}\rho_{2,3}}{1-\rho_{1,2}^2-\rho_{2,3}^2} < 0$ if $\rho_{1,2} \in (0,1)$ and $\rho_{2,3} \in (-1,0)$, or $\rho_{1,2} \in (-1,0)$ and $\rho_{2,3} \in (0,1)$. To ensure that $|\mathbf{P}| > 0$ (i.e., $\mathbf{P}$ is positive definite) these conditions are modified as follows: i) $\left[\mathbf{P}^{-1}\right]_{1,3} = \frac{\rho_{1,2}\rho_{2,3}}{1-\rho_{1,2}^2-\rho_{2,3}^2} > 0$ if $\rho_{1,2} \in (-1,0)$ and $\rho_{2,3} \in \left(-\sqrt{1-\rho_{1,2}^2}, 0\right)$, or $\rho_{1,2}$ and $\rho_{2,3} \in \left(0, \sqrt{1-\rho_{1,2}^2}\right)$, and ii) $\left[\mathbf{P}^{-1}\right]_{1,3} = \frac{\rho_{1,2}\rho_{2,3}}{1-\rho_{1,2}^2-\rho_{2,3}^2} < 0$ if $\rho_{1,2} \in (0,1)$ and $\rho_{2,3} \in \left(-\sqrt{1-\rho_{1,2}^2}, 0\right)$, or $\rho_{1,2} \in (-1,0)$ and $\rho_{2,3} \in \left(0, \sqrt{1-\rho_{1,2}^2}\right)$. Along similar lines, in the following cases with $\rho_{1,3} \neq 0$: i) $0 < \rho_{1,3} < \rho_{2,3} < 1$ and $\rho_{2,1} = \rho_{1,3}/\rho_{2,3}$ imply $\left[\mathbf{P}^{-1}\right]_{1,3} = 0$, ii) $-1 < \rho_{2,3} < \rho_{1,3} < 0$ and $\rho_{2,1} = \rho_{1,3}/\rho_{2,3}$ imply $\left[\mathbf{P}^{-1}\right]_{1,3} = 0$, iii) $0 < \rho_{1,2} < \rho_{1,3} < 1$ and $\rho_{2,3} \in \left(\rho_{1,2}\rho_{1,3} - \sqrt{\left(1-\rho_{1,2}^2\right)\left(1-\rho_{1,3}^2\right)}, \rho_{1,2}\rho_{1,3} + \sqrt{\left(1-\rho_{1,2}^2\right)\left(1-\rho_{1,3}^2\right)}\right)$ imply $\left[\mathbf{P}^{-1}\right]_{1,3} < 0$, and iv) $-1 < \rho_{1,3} < -\rho_{1,2} < 0$ and $\rho_{2,3} \in \left(\rho_{1,2}\rho_{1,3} - \sqrt{\left(1-\rho_{1,2}^2\right)\left(1-\rho_{1,3}^2\right)}, \rho_{1,2}\rho_{1,3} + \sqrt{\left(1-\rho_{1,2}^2\right)\left(1-\rho_{1,3}^2\right)}\right)$ imply $\left[\mathbf{P}^{-1}\right]_{1,3} > 0$. Cases i)-iv) also imply $|\mathbf{P}| > 0$.□

Proof of Proposition 3. By differentiating δ_i^* we obtain $\partial\delta_i^*/\partial c_i = -\frac{\left(V_i^H - V_i^L\right)}{r\left(\sigma_i^H-\sigma_i^L\right)^2}\left[\mathbf{P}^{-1}\right]_{i,i} < 0$, and $\partial\delta_i^*/\partial V_i^H = \frac{(1-c_i)}{r\left(\sigma_i^H-\sigma_i^L\right)^2}\left[\mathbf{P}^{-1}\right]_{i,i} > 0$. For ease of exposition, define $A \doteq \frac{(1-c_i)\left(V_i^H-V_i^L\right)}{r\left(\sigma_i^H-\sigma_i^L\right)^2}\left[\mathbf{P}^{-1}\right]_{i,i} > 0$ and $B \doteq \sum_{j\in\mathcal{J}\setminus i}\frac{(1-c_j)\left(V_j^H-V_j^L\right)}{r}\left[\mathbf{P}^{-1}\right]_{j,i} \lessgtr 0$. Then, $\delta_i^* = \frac{1}{\sigma_i^H-\sigma_i^L}\left(\frac{A}{\sigma_i^H-\sigma_i^L} + B - \sigma_i^L\right)$ and $\partial\delta_i^*/\partial\sigma_i^H =$

$-\frac{2A+\left(B-\sigma_i^L\right)\left(\sigma_i^H-\sigma_i^L\right)}{\left(\sigma_i^H-\sigma_i^L\right)^3}$. Given that $\delta_i^*>0$, $\sigma_i^H-\sigma_i^L<0$ implies $\frac{A}{\sigma_i^H-\sigma_i^L}+B-\sigma_i^L<0$ and as an extension $2A+\left(B-\sigma_i^L\right)\left(\sigma_i^H-\sigma_i^L\right)>0$, results in $\partial\delta_i^*/\partial\sigma_i^H>0$. Similarly, given that $\delta_i^*>0$, $\sigma_i^H-\sigma_i^L>0$ implies $\frac{A}{\sigma_i^H-\sigma_i^L}+B-\sigma_i^L>0$ and as an extension $2A+\left(B-\sigma_i^L\right)\left(\sigma_i^H-\sigma_i^L\right)>0$, resulting in $\partial\delta_i^*/\partial\sigma_i^H<0$. To characterize the relationship between the optimal efforts at step i and j it suffices to differentiate δ_i^* with respect to V_j^H. Doing so returns $\frac{\partial\delta_i^*}{\partial V_j^H}=\frac{\left(1-c_j\right)}{r\left(\sigma_i^H-\sigma_i^L\right)\left(\sigma_j^H-\sigma_j^L\right)}\left[\mathbf{P}^{-1}\right]_{i,j}\lesseqgtr 0$. The sign of $\frac{\partial\delta_i^*}{\partial V_j^H}$ is determined by the sign of: i) $\left(\sigma_i^H-\sigma_i^L\right)\left(\sigma_j^H-\sigma_j^L\right)$, which is positive if both steps are routine or non-routine (i.e., $sgn\left[\sigma_i^H-\sigma_i^L\right]=sgn\left[\sigma_j^H-\sigma_j^L\right]$) and negative otherwise and ii) $\left[\mathbf{P}^{-1}\right]_{i,j}$. For instance, $\frac{\partial\delta_i^*}{\partial V_j^H}<0$ when $\sigma_i^H-\sigma_i^L>0$, $\sigma_j^H-\sigma_j^L<0$, and $\left[\mathbf{P}^{-1}\right]_{i,j}>0$, which implies that the efforts are substitutable. A similar approach, resulting in the same insights applies in order to determine the signs of $\partial\delta_i^*/\partial c_j=-\frac{\left(V_j^H-V_j^L\right)}{r\left(\sigma_i^H-\sigma_i^L\right)\left(\sigma_j^H-\sigma_j^L\right)}\left[\mathbf{P}^{-1}\right]_{j,i}$ and $\partial\delta_i^*/\partial\sigma_j^H=-\frac{\left(1-c_j\right)\left(V_j^H-V_j^L\right)}{r\left(\sigma_i^H-\sigma_i^L\right)\left(\sigma_j^H-\sigma_j^L\right)^2}\left[\mathbf{P}^{-1}\right]_{j,i}$, and compare them with $\partial\delta_i^*/\partial c_i$, and $\partial\delta_i^*/\partial\sigma_i^H$, respectively.□

Proof of Proposition 4 and Corollary 1. For the purposes of this proof, we draw directly from Gabriel (1962) and Zimmerman and Núñez-Antón (2009). Specifically, Zimmerman and Núñez-Antón (2009) (Theorem 2.2, p.p. 37-38) establish that normally distributed random variables $\tilde{e}_1,\ldots,\tilde{e}_n$ with positive definite covariance matrix $\mathbf{\Sigma}$ follow an $AD\left(p\right)$ process iff $\left[\mathbf{\Sigma}^{-1}\right]_{i,j}=0$ for all $i,j\in I_p^n$, where $I_p^n=\{i,j:i\in\{1:n,\},j\in\{1:n\},|i-j|>p\}$. This is equivalent to $\rho_{i,j\cdot\mathcal{J}\setminus\{i,j\}}=0$ for $|i-j|>p$ (see also Theorem 1 in Gabriel 1962) and therefore, $\left[\mathbf{P}^{-1}\right]_{i,j}=0$ for $|i-j|>p$. To derive $\mathbf{P}$ for the case of $p=1$, we replicate the proof in p. 48 of Zimmerman and Núñez-Antón (2009): Consider the partial covariance $\sigma_{i,j\cdot m}$. Under an $AD\left(1\right)$, $\sigma_{i,j\cdot m}=0$ for any $m\in\left(i,j\right)$. Hence,

$$0=\sigma_{i,j\cdot m}=\sigma_{i,j}-\frac{\sigma_{i,m}\sigma_{j,m}}{\sigma_{m,m}}=\rho_{i,j}\left(\sigma_{i,i}\sigma_{j,j}\right)^{1/2}-\frac{\rho_{i,m}\left(\sigma_{i,i}\sigma_{m,m}\right)^{1/2}\rho_{j,m}\left(\sigma_{j,j}\sigma_{m,m}\right)^{1/2}}{\sigma_{m,m}},$$

which implies $\rho_{i,j}=\rho_{i,m}\rho_{m,j}$ for $i>m>j$, where m is an arbitrary step between steps i and j. After repeatedly substituting $\rho_{i,m}$ and $\rho_{m,j}$ with products of correlations of adjacent steps, we obtain $\left[\mathbf{P}\right]_{i,j}=\prod_{m=j}^{i-1}\rho_{m+1,m}$. To illustrate, consider $\mathbf{P}=\begin{bmatrix}1&\rho_1&\rho_1\rho_2\\\rho_1&1&\rho_2\\\rho_1\rho_2&\rho_1&1\end{bmatrix}$, with $\rho_i\doteq\rho_{i+1,i}\in(-1,1)$. $\mathbf{P}$ is always positive definite and results in $\left[\mathbf{P}^{-1}\right]=\begin{bmatrix}\frac{1}{1-\rho_1^2}&-\frac{\rho_1}{1-\rho_1^2}&0\\-\frac{\rho_1}{1-\rho_1^2}&\frac{1-\rho_1^2\rho_2^2}{\left(1-\rho_1^2\right)\left(1-\rho_2^2\right)}&-\frac{\rho_2}{1-\rho_2^2}\\0&-\frac{\rho_2}{1-\rho_2^2}&\frac{1}{1-\rho_2^2}\end{bmatrix}$.□

Proof of Proposition 5. For notational brevity we define $\Delta_i\doteq\frac{(1-c_i)\left(V_i^H-V_i^L\right)}{r\left(\sigma_i^H-\sigma_i^L\right)}$. In an $ARH\left(1\right)$, $\left[\mathbf{P}\right]_{i,j}=\rho^{|i-j|}\in(0,1)$ for all $i,j\in\mathcal{J}$, implying

$$\mathbf{P}=\begin{bmatrix}1&\rho&\rho^2&\ldots&\rho^{n-1}\\\rho&1&\rho&\ldots&\rho^{n-2}\\\vdots&\ddots&\ddots&\ddots&\vdots\\\rho^{n-2}&\ldots&\rho&1&\rho\\\rho^{n-1}&\ldots&\rho^2&\rho&1\end{bmatrix}\text{ and }\mathbf{P}^{-1}=\frac{1}{1-\rho^2}\begin{bmatrix}1&-\rho&0&\ldots&0\\-\rho&1&-\rho&\ldots&0\\\vdots&\ddots&\ddots&\ddots&\vdots\\0&\ldots&-\rho&1&-\rho\\0&\ldots&0&-\rho&1\end{bmatrix}.$$

Hence, we can rewrite the provider's optimal effort at step i as $\delta_i^* = \frac{-\rho\Delta_{i-1}+\left(1+\rho^2\right)\Delta_i-\rho\Delta_{i+1}}{\left(\sigma_i^H-\sigma_i^L\right)\left(1-\rho^2\right)}$. After differentiating δ_i^* with respect to ρ we obtain $\partial\delta_i^*/\partial\rho = -\frac{\left(\Delta_{i-1}+\Delta_{i+1}\right)\left(1+\rho^2\right)-4\Delta_i\rho}{\left(\sigma_i^H-\sigma_i^L\right)\left(1-\rho^2\right)^2}$. It is easy to show that when: i) $\sigma_i^H - \sigma_i^L > 0$ (i.e., $\Delta_i > 0$) and $\Delta_{i-1} + \Delta_{i+1} < 0$ or ii) $\sigma_i^H - \sigma_i^L < 0$ (i.e., $\Delta_i < 0$) and $\Delta_{i-1} + \Delta_{i+1} > 0$, $\partial\delta_i^*/\partial\rho > 0$ for any $\rho \in (0,1)$. For the rest of the cases with $\sigma_i^H - \sigma_i^L > 0$ (i.e., $\Delta_i > 0$) and $\Delta_{i-1} + \Delta_{i+1} > 0$ or $\sigma_i^H - \sigma_i^L < 0$ (i.e., $\Delta_i < 0$) and $\Delta_{i-1} + \Delta_{i+1} < 0$, we solve $\partial\delta_i^*/\partial\rho = 0$ with respect to ρ and we obtain $\partial\delta_i^*/\partial\rho \le 0$ for all $\rho \in (0,\bar{\rho}]$ and $\partial\delta_i^*/\partial\rho > 0$ for all $\rho \in (\bar{\rho},1)$, where $\bar{\rho} \doteq \frac{2\Delta_i}{\Delta_{i-1}+\Delta_{i+1}} - \sqrt{-1 + \frac{4\Delta_i^2}{\left(\Delta_{i-1}+\Delta_{i+1}\right)^2}}$.$\square$

Proof of Proposition 6. Given that steps 1 and n have the same characteristics we can define $\Delta_1 = \Delta_n \doteq \Delta$, where $\Delta_i = \frac{(1-c_i)\left(V_i^H-V_i^L\right)}{r\left(\sigma_i^H-\sigma_i^L\right)}$, and $\Delta = \frac{(1-c)\left(V^H-V^L\right)}{r\left(\sigma^H-\sigma^L\right)}$. Hence, δ_1^* and δ_n^* can be expressed as $\delta_1^* = \frac{\left(1+\rho^2\right)\Delta-\rho\Delta_2}{\left(\sigma^H-\sigma^L\right)\left(1-\rho^2\right)}$, and $\delta_n^* = \frac{\left(1+\rho^2\right)\Delta-\rho\Delta_{n-1}}{\left(\sigma^H-\sigma^L\right)\left(1-\rho^2\right)}$ based on which we obtain $\delta_1^* - \delta_n^* = -\frac{\left(\Delta_2-\Delta_{n-1}\right)\rho}{\left(\sigma^H-\sigma^L\right)\left(1-\rho^2\right)} > 0$ iff: i) $\sigma^H - \sigma^L > 0$ and $\Delta_2 < \Delta_{n-1}$, or ii) $\sigma^H - \sigma^L < 0$ and $\Delta_2 > \Delta_{n-1}$. For services where $\Delta_1 \neq \Delta_n$, we obtain $\delta_1^* - \delta_n^* = \frac{\Delta_1\left(\sigma_n^H-\sigma_n^L\right)-\Delta_n\left(\sigma_1^H-\sigma_1^L\right)}{\left(\sigma_1^H-\sigma_1^L\right)\left(\sigma_n^H-\sigma_n^L\right)} - \frac{\Delta_2\left(\sigma_n^H-\sigma_n^L\right)-\Delta_{n-1}\left(\sigma_1^H-\sigma_1^L\right)}{\left(\sigma_1^H-\sigma_1^L\right)\left(\sigma_n^H-\sigma_n^L\right)}\frac{\rho}{1-\rho^2} > 0$ iff: i) $\sigma_n^H - \sigma_n^L > 0$ and $\Delta_n < \bar{\Delta}_n$, or ii) $\sigma_n^H - \sigma_n^L < 0$ and $\Delta_n > \bar{\Delta}_n$, where $\bar{\Delta}_n \doteq \Delta_1\frac{\sigma_n^H-\sigma_n^L}{\sigma_1^H-\sigma_1^L} + \frac{\Delta_{n-1}\left(\sigma_1^H-\sigma_1^L\right)-\Delta_2\left(\sigma_n^H-\sigma_n^L\right)}{\sigma_1^H-\sigma_1^L}\frac{\rho}{\left(1+\rho^2\right)}$.$\square$

References

Arieff, A. 2009. A Breath of Fresh Air for Health Care. `http://opinionator.blogs.nytimes.com/2009/12/13/a-breath-of-fresh-air-for-health-care/`. Accessed January 22, 2019.

Ariely, D., Z. Carmon. 2000. Gestalt Characteristics of Experiences: The Defining Features of Summarized Events. *J. Behav. Decision Making* **13** 191–201.

Ariely, D., G. Zauberman. 2003. Differential Partitioning of Extended Experiences. *Organ. Behav. Hum. Dec.* **91**(2) 128–139.

Arora, P., M. Rahmani, K. Ramachandran. 2019. Service Design of Nonprofits that Serve Distressed Individuals. Georgia Tech Scheller College of Business Research Paper No. 17-33.

Baucells, M., R. Sarin. 2007. Satiation in Discounted Utility. *Oper. Res.* **55**(1) 170–181.

Baucells, M., R. Sarin. 2010. Predicting Utility Under Satiation and Habit Formation. *Management Sci.* **56**(2) 286–301.

Bellos, I., S. Kavadias. 2019. When Should Customers Control Service Delivery? Implications for Service Design. *Production and Oper. Management* **28**(4) 890–907.

Bhavnani, R., M. Sosa. 2006. IDEO: Service Design (A&B). INSEAD Case Study 11/2006-5276.

Bitner, M., A. Ostrom, F. Morgan. 2008. Service Blueprinting: A Practical Technique for Service Innovation. *Calif. Manage. Rev.* **50**(3) 66.

Bitran, G., J. Ferrer, P. Rocha e Oliveira. 2008. OM Forum–Managing Customer Experiences: Perspectives on the Temporal Aspects of Service Encounters. *Manufacturing Service Oper. Management* **10**(1) 61.

Bowen, J., R. Ford. 2002. Managing Service Organizations: Does Having a "Thing" Make a Difference? *J. Management.* **28**(3) 447–469.

Brown, T. 2008. Design Thinking. *Harvard Bus. Rev.* **86**(6) 84.

Brown, T. 2009. *Change by Design*, chap. 4; p.p. 94-95. HarperCollins Publishers, 195 Broadway, New York, NY 10007.

Carmon, Z., J. Shanthikumar, T. Carmon. 1995. A Psychological Perspective on Service Segmentation Models: The Significance of Accounting for Consumers' Perceptions of Waiting and Service. *Management Sci.* **41**(11) 1806–1815.

Chan, T., J. Mihm, M. Sosa. 2017. On Styles in Product Design: An Analysis of US Design Patents. *Management Sci.* **64**(3) 1230–1249.

Chase, R. 1978. Where Does the Customer Fit in a Service Operation? *Harvard Bus. Rev.* **56**(6) 137–142.

Chase, R. 1981. The Customer Contact Approach to Services: Theoretical Bases and Practical Extensions. *Oper. Res.* **29**(4) 698–706.

Cliver, M., J. Hegeman, K. Lee, L. Libert, K. Tennant. 2007. Designing for the Clinic Experience: Service Design for UPMC Presbyterian Neurosurgery Clinic. `http://www.slideshare.net/jaminhegeman/upmc-neuro-clinic-service-design`. Accessed January 23, 2019.

Crano, W. 1977. Primacy versus Recency in Retention of Information and Opinion Change. *J. Soc. Psych.* **101**(1) 87–96.

Das Gupta, A., U. Karmarkar, G. Roels. 2016. The Design of Experiential Services with Acclimation and Memory Decay: Optimal Sequence and Duration. *Management Sci.* **62**(5) 1278–1296.

Diggle, P., P. Heagerty, K. Liang, S. Zeger. 2002. *Analysis of Longitudinal Data*, chap. 5; p. 87. Oxford University Press, Great Clarendon Street, Oxford OX2, 6DP, UK.

Dixon, M., G. Thompson. 2016. Bundling and Scheduling Service Packages with Customer Behavior: Model and Heuristic. *Production and Oper. Management* **25**(1) 36–55.

Dixon, M., R. Verma. 2013. Sequence Effects in Service Bundles: Implications for Service Design and Scheduling. *J. Oper. Management* **31**(3) 138–152.

Dixon, M., L. Victorino. 2019. The Sequence of Service: An Affect Perspective to Service Scheduling. P. Maglio, C. Kieliszewski, J. Spohrer, K. Lyons, L. Patrício, Y. Sawatani, eds., *Handbook of Service Science, Volume II*, chap. 4. Springer International Publishing, 677–690.

Dixon, M., L. Victorino, R. Kwortnik, R. Verma. 2017. Surprise, Anticipation, and Sequence Effects in the Design of Experiential Services. *Production and Oper. Management* **26**(5) 945–960.

Doblin. 2019. Doblin, a Deloitte Business. `https://www.doblin.com/`. Accessed January 19, 2019.

Ely, J., A. Frankel, E. Kamenica. 2015. Suspense and Surprise. *J. of Political Econom.* **123**(1) 215–260.

Fayard, A., I. Stigliani, B. Bechky. 2017. How Nascent Occupations Construct a Mandate: The Case of Service Designers' Ethos. *Adm. Sci. Q.* **62**(2) 270–303.

Fjord. 2019. Design and Innovation from Accenture Interactive. `https://www.fjordnet.com/`. Accessed January 19, 2019.

Forbes. 2019. Four Types Of Metrics That Support Great Customer Experience Measurement. `http://tinyurl.com/yxba2p7a`. Accessed February 20, 2019.

Gabriel, K. 1962. Ante-dependence Analysis of an Ordered Set of Variables. *Ann. Math. Stat.* **33**(1) 201–212.

Heskett, J., T. Jones, G. Loveman, W. Sasser, L. Schlesinger. 1994. Putting the Service-profit Chain to Work. *Harvard Bus. Rev.* **72**(2) 164–174.

IBM Knowledge Center. 2019. Covariance Structures. `http://disq.us/t/2scq5mt`. Accessed January 8, 2019.

Kac, M., W. Murdock, G. Szegö. 1953. On the Eigen-values of Certain Hermitian Forms. *J. Rat. Mech. and Anal.* **2**(6) 767–802.

Kahneman, D., B. Fredrickson, C. Schreiber, D. Redelmeier. 1993. When More Pain is Preferred to Less: Adding a Better End. *Psych. Sci.* **4**(6) 401–405.

Kalbach, J. 2016. *Mapping Experiences: A Complete Guide to Creating Value through Journeys, Blueprints, and Diagrams*. O'Reilly Media, 1005 Gravestein Highway North, Sebastopol, CA 95472.

Kannan, P., J. Proença. 2010. Design of Service Systems Under Variability: Research Issues. *Inf. Syst. E-Bus. Manage.* **8**(1) 1–11.

Karmarkar, U., R. Pitbladdo. 1995. Service Markets and Competition. *J. Oper. Management* **12**(3-4) 397–411.

Karmarkar, U., R. Pitbladdo. 1997. Quality, Class, and Competition. *Management Sci.* **43**(1) 27–39.

Kim, S., M. Cohen, S. Netessine. 2007. Performance Contracting in After-Sales Service Supply Chains. *Management Sci.* **53**(12) 1843–1858.

Kroll, Y., H. Levy, H. Markowitz. 1984. Mean-Variance versus Direct Utility Maximization. *J. Finance* **39**(1) 47–61.

Lee, H., E. Pinker, R. Shumsky. 2012. Outsourcing a Two-Level Service Process. *Management Sci.* **58**(8) 1569–1584.

Levitt, T. 1972. Production-line Approach to Service. *Harvard Bus. Rev.* **50**(5) 41–52.

Levy, H., H. Markowitz. 1979. Approximating Expected Utility by a Function of Mean and Variance. *Amer. Econom. Rev.* **69**(3) 308–317.

Martin, K., M. Osterling. 2014. *Value Stream Mapping: How to Visualize Work and Align Leadership for Organizational Transformation*. McGraw-Hill Education, 2 Pennsylvania Plaza, New York City, NY 10121.

Martínez-de Albéniz, V., A. Valdivia. 2019. Measuring and Exploiting the Impact of Exhibition Scheduling on Museum Attendance. *Manufacturing Service Oper. Management* **21**(4) 761–779.

McKinsey&Company. 2014. The Three Cs of Customer Satisfaction: Consistency, Consistency, Consistency. http://bit.ly/2fQ6GiP. Accessed January 7, 2019.

McKinsey&Company. 2016. The CEO Guide to Customer Experience. *McKinsey Quarterly* **August** 1–10.

Murray, K., J. Schlacter. 1990. The Impact of Services Versus Goods on Consumer' Assessment of Perceived Risk and Variability. *J. Acad. Mark. Sci.* **18**(1) 51–65.

OpenIDEO. 2013. The Story of Doug Dietz: Creative Confidence in the MRI Suite. https://bit.ly/2ud5eBo. Accessed January 2, 2020.

Pourahmadi, M. 1999. Joint Mean-Covariance Models with Applications to Longitudinal Data: Unconstrained Parameterisation. *Biometrika* **86**(3) 677–690.

Raveh, A. 1985. On the Use of the Inverse of the Correlation Matrix in Multivariate Data Analysis. *Am. Stat.* **39**(1) 39–42.

Roels, G. 2014. Optimal Design of Coproductive Services: Interaction and Work Allocation. *Manufacturing Service Oper. Management* **16**(4) 578–594.

Roels, G. 2019. Optimal Structure of Experiential Services: Review and Extensions. P. Maglio, C. Kieliszewski, J. Spohrer, K. Lyons, L. Patrício, Y. Sawatani, eds., *Handbook of Service Science, Volume II*, chap. 6. Springer International Publishing, 105–146.

Roels, G., U. Karmarkar, S. Carr. 2010. Contracting for Collaborative Services. *Management Sci.* **56**(5) 849–863.

Rother, M., J. Shook. 2003. *Learning to See: Value Stream Mapping to Add Value and Eliminate Muda*. Lean Enterprise Institute, One Cambridge Center, Cambridge, MA 02142.

Rousseeuw, P., G. Molenberghs. 1994. The Shape of Correlation Matrices. *Am. Stat.* **48**(4) 276–279.

Rust, R., J. Inman, J. Jia, A. Zahorik. 1999. What You Don't Know About Customer-Perceived Quality: The Role of Customer Expectation Distributions. *Marketing Sci.* **18**(1) 77–92.

Sasser, W. 1976. Match Supply and Demand in Service Industries. *Harvard Bus. Rev.* **54**(6) 133–140.

Shostack, G. 1984. Designing Services that Deliver. *Harvard Bus. Rev.* **62**(1) 133–139.

Shostack, G. 1987. Service Positioning Through Structural Change. *J. Marketing* **51**(1) 34–43.

Soteriou, A., R. Chase. 2000. A Robust Optimization Approach for Improving Service Quality. *Manufacturing Service Oper. Management* **2**(3) 264–286.

Soteriou, A., G. Hadjinicola. 1999. Resource Allocation to Improve Service Quality Perceptions in Multistage Service Systems. *Production and Oper. Management* **8**(3) 221–239.

Sriram, S., P. Chintagunta, P. Manchanda. 2015. Service Quality Variability and Termination Behavior. *Management Sci.* **61**(11) 2739–2759.

Stickdorn, M., J. Schneider. 2010. *This is Service Design Thinking: Basics-Tools-Cases*. BIS Publishers, 1057 DT Amsterdam, The Netherlands.

Tan, T., S. Netessine. 2019. When You Work with a Superman, Will You Also Fly? An Empirical Study of the Impact of Coworkers on Performance. *Management Sci.* **65**(8) 3449–3947.

Teboul, J. 2006. *Service is Front Stage: Positioning Services for Value Advantage*. Palgrave Macmillan, 175 Fifth Avenue, New York, NY 10010.

Thomke, S. 2003. R&D Comes to Services. *Harvard Bus. Rev.* **81**(4) 70–79.

Thomke, S. 2019. The Magic That Makes Customer Experiences STICK. *MIT Sloan Manage. Rev.* **61**(1) 56–63.

Tong, C., M. Nagarajan, Y. Cheng. 2016. Operational Impact of Service Innovations in Multi-Step Service Systems. *Production and Oper. Management* **25**(5) 833–848.

Ulrich, K. 1995. The Role of Product Architecture in the Manufacturing Firm. *Res. Policy* **24**(3) 419–440.

Ulrich, K. 2011. *Design: Creation of Artifacts in Society*. University of Pennsylvania, Philadelphia, PA 19104.

White, S., R. Badinelli. 2012. A Model for Efficiency-Based Resource Integration in Services. *Eur. J. Oper. Res.* **217**(2) 439–447.

Wolfinger, R. 1996. Heterogeneous Variance: Covariance Structures for Repeated Measures. *J. Agric. Biol. Environ. Stat.* 205–230.

Womack, J., D. Jones. 2005. Lean Consumption. *Harvard Bus. Rev.* **83**(3) 58–68.

Xue, M., J. Field. 2008. Service Coproduction with Information Stickiness and Incomplete Contracts: Implications for Consulting Services Design. *Production and Oper. Management* **17**(3) 357–372.

Xue, M., P. Harker. 2003. Service Co-Production, Customer Efficiency and Market Competition. Working Paper, University of Pennsylvania, Philadelphia, PA.

Zeithaml, V., A. Parasuraman, L. Berry. 1985. Problems and Strategies in Services Marketing. *J. Marketing* **49**(2) 33–46.

Zimmerman, D., V. Núñez-Antón. 2009. *Antedependence Models for Longitudinal Data*. CRC Press, 6000 Broken Sound Parkway NW, Boca Raton, FL 33487-2742.