USING MACHINE LEARNING TO COCREATE VALUE THROUGH DYNAMIC CUSTOMER ENGAGEMENT IN A BRAND LOYALTY PROGRAM

Ajay Aluri
Bradley S. Price
Nancy H. McIntyre
West Virginia University

Hospitality venues traditionally use historical data from customers for their customer relationship management systems, but now they can also collect real-time data and automated procedures to make dynamic decisions and predictions about customer behavior. Machine learning is an example of automated processes that create insights into cocreation of value through dynamic customer engagement. *To show the merits of automation, machine learning was implemented at a major hospitality venue and compared with traditional methods to identify what customers value in a loyalty program. The results show that machine learning processes are superior in identifying customers who find value in specific promotions. This research deepens practical and theoretical understanding of machine learning in the customer engagement-to-value loyalty chain and in the customer engagement construct that uses a dynamic customer engagement model.*

KEYWORDS: *cocreation of value; customer engagement; machine learning; predictive data analytics; customer loyalty*

INTRODUCTION

Decision support systems have been used for a long time in the hospitality industry. Hotels, restaurants, and other venues collect data and information from customers to manage their relationship with those customers. These customer relationship management (CRM) systems allow managers to directly translate historical data into knowledge about the business to improve decision making and performance (McAfee & Brynjolfsson, 2012). However, the use of "big data," or predictive data analytics, which provides the ability to analyze and use real-time data rather than just historical data, is still new in the hospitality industry (Marr, 2016). Rather than just summarizing historical data, predictive data analytics uses techniques such as machine learning to recommend possible

Journal of Hospitality & Tourism Research, 201X, Vol. XX, No. X, Month 2018, 1–23
DOI: 10.1177/1096348017753521

courses of action in real time, as well as the most likely consequences of each course of action (Hastie, Tibshirani, & Friedman, 2009). As the information available in the hospitality industry increases exponentially, it is imperative for hospitality venues to understand and begin to use big data to engage with their customers and cocreate value and ultimately strengthen customer loyalty. This was reiterated in an interview with the former CIO of Wyndham Worldwide, who stated that predictive analytics, machine learning, and artificial intelligence are creating a competitive advantage for those organizations in hospitality who implement them (Price, 2017).

Mitchell (2006) defines the field of machine learning as the convergence of the disciplines of computer science and statistics to study the central question, "How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern these learning processes?" This differs from the defining questions in both computer science and statistics, which focus on manually programming computers and drawing inference from data (Mitchell, 2006). For the purposes of this research, the focus will be on using the tools of statistical machine learning, a specific subfield that automates the estimation and production of methods that have statistical justification. A review of these methods can be found in textbooks such as Hastie et al. (2009) and James, Witten, Hastie, and Tibshirani (2015). Typically, these methods are used to predict an outcome and compare the accuracy of many models. Unlike traditional metrics, algorithms are used to fit and compare many models at one time. This happens automatically, and then the accuracy of the algorithms is compared. Methods such as this can implement algorithms, study the accuracy of the selected models, and draw conclusions based on these results. There are many ways to compare these results, and although these models may be considered nontraditional in the hospitality industry, they have been well studied in other disciplines (Guyon & Elisseeff, 2006; Weisberg, 2014). These models and techniques include comparison of prediction and classification models, and understanding the impact of variables on the ability to effect predictions and decisions, rather than relying on a causal inference model. Though there is work in the machine learning community regarding the effect of causal inference in social science, this research is still in its infancy (Athey, 2015; Grimmer, 2015). The study by Allenby, Bradlow, George, Liechty, and McCulloch (2014) discusses similar issues with big data in marketing and provides perspectives on methodology and the limitations of using big data for causal inference.

Because the application of these techniques in the hospitality and tourism industry is so new, hospitality and tourism literature contains only a few studies that have used big data analytics to offer insights into guest experience and satisfaction from user-generated reviews, and the majority of these studies have used text analytics from social media sites (Liu, Teichert, Rossi, Li, & Hu, 2017; Xiang, Schwartz, Gerdes, & Uysal, 2015). Although several researchers and practitioners have recommended the use of big data analytics to identify guest segments and develop operational strategies that lead to increased revenues

(Jennings, Giorgio, Murali, & Goggin, 2014; Vivek, Beatty, & Morgan, 2012), to our knowledge, none have yet used machine learning to examine behavior at a personalized level. Recently, researchers have conducted empirical studies of customer engagement in brand loyalty, service, and trust, using structural equation modeling, and the authors note that customer engagement has emerged as a predictor of brand loyalty but still lacks empirical justification (So, King, Sparks, & Wang, 2016). However, none of these recent studies have empirically examined how machine learning can be used to increase revenue by enhancing customer engagement.

In the world of big data, data analytics and machine learning can be used to offer targeted services that identify what customers value in real time and increase customer engagement. In this study, we focus on how to engage with customers *during their interaction with the venue* and on using those interactions in a way that cocreates value and customer loyalty. Given the dearth of research in this area, the primary purpose of this study is to examine the predictive impact of customer engagement, to enhance and increase customer engagement as measured by revenue, and to analyze how this *dynamic customer engagement* influences customer loyalty (as measured by annual membership renewal) when machine learning is used to personalize the customer engagement experience.

LITERATURE REVIEW

Hospitality venues worldwide are in a race to create more personalized experiences to build customer loyalty and repeat business. Loyalty programs are considered one of the best tools to accomplish this. The key goal of loyalty programs is to increase customer engagement and, more important, retain customers. Using data collected from these programs, CRM systems create segments based on customer behaviors and engagement with various programs offered by the brand. Customer engagement is a vehicle for enhancing customer relationships that enables sales promotion and discounting, directs product quality improvements, increases customer satisfaction, and builds a competitive advantage for the hospitality venue (Di Gangi & Wasko, 2009). Engaged customers exhibit enhanced customer loyalty, satisfaction, empowerment, connection, emotional bonding, trust, and commitment (Brodie, Ilic, Juric, & Hollebeek, 2013). Customer engagement has also been shown to be a valuable predictor of future business performance, and can drive revenue growth and enhance profitability (Thakur & Summey, 2010).

Loyalty programs are designed to improve customer satisfaction and long-term commitment. They usually give away benefits to members as a token of appreciation for their loyalty to the organization. Bolton, Kannan, and Bramlett (2000) state that when customers are involved in loyalty programs, the perceived benefits they receive will lead to their loyalty. A common practice in many loyalty programs is to offer universal and perpetual discounts to certain populations or to create a points-based rewards system. The goal behind these programs is to drive and incentivize customer engagement. A recent study by Nielsen (2016) looked at how opinions of loyalty programs are changing based on monetary and nonmonetary value and generation gaps. The study indicated that product discounts, rebates, and free products were the three most valued rewards of a loyalty program. To establish a baseline for revenue generation based on discounting for customers in this study, the following hypotheses were formulated. It should be noted that these hypotheses were proposed by the client as a priori hypotheses to test whether randomized discounting would increase customer spending. This gives insight into a baseline of customers' perceived value of discounting.

Hypothesis 1_0 ($H1_0$): Customers who are offered randomized discounts will not spend more using their discount at the hospitality event venue concessions than they would expect to spend when a discount is not offered.

$H1_a$: Customers who are offered randomized discounts will spend more using their discount at the hospitality event venue concessions than they would expect to spend when a discount is not offered.

Customer Engagement and Cocreation of Value

As noted above, the goal of loyalty programs is to increase customer engagement and retain customers. The concept of customer engagement falls within the broader domain of CRM. Patterson, Yu, and DeRuyter (2006) define customer engagement as the level of a customer's physical, cognitive, and emotional presence in the relationship with a service organization. Vivek et al. (2012, p. 127) define customer engagement as "the intensity of an individual's participation and connection with the organization's offerings and activities initiated by either the customer or the organization."

In a recent editorial, noting the "surging academic and managerial interest" in the customer engagement concept, Hollebeek, Conduit, and Brodie (2016) note that there are cognitive, emotional, and behavioral dimensions underlying the customer engagement concept (Brodie, Hollebeek, Ilic, & Juric, 2011), and efforts to operationalize and measure customer engagement have mixed results. Even so, they proposed that all the definitions predominantly reflect an individual's psychologically based willingness to invest in the undertaking of focal interactions with particular engagement objects (e.g., a brand or firm), often beyond just making a purchase (Groeger, Moroko, & Hollebeek, 2016).

The real focus of much of the research on customer engagement is on the consequences, which include satisfaction (Bowden, 2009), commitment (Chan & Li, 2010), trust (Hollebeek, 2011), empowerment (Gruen, Osmonbekov, & Czaplewski, 2006), and loyalty (Bowden, 2009; Schouten, McAlexander, & Koenig, 2007), though the ultimate goal for any hospitality and tourism venue is to increase revenue.

Developing customer engagement by involving customers in value cocreation received special attention from scientists during the first decade of the century (Prahalad & Ramaswamy, 2004; Vargo & Lusch, 2008). The customer value creation process is the process by which producers and customers cocreate value for themselves and each other. It is now believed that to create the highest value, customers should be involved in the process of cocreating what they define as value. Engaged customers become partners who cooperate with the company in the process of value cocreation, to satisfy their own and other customers' needs (Sashi, 2012; Sawhney, Verona, & Prandelli, 2005). Vargo, Maglio, and Akaka (2008) suggest that the locus of value creation should move from exchange to use or context, transforming our understanding of value from one based on *units* of firm output (i.e., points earned) to one based on *processes* that integrate resources. More recently, Fernandes and Esteves (2016) discussed multidimensional concepts of customer engagement, describing links between intention and behaviors of engagement and the linkage of these ideas to customer loyalty through patronage and attitude.

While we concede that multidimensional concepts of engagement, such as word of mouth and other constructs, are useful for studying behavior in a fully dynamic environment, this information is harder to obtain unless, for instance, social media information is available for all customers in real time. Without information for all customers, missing data and bias issues arise in modeling. Therefore, our focus in this study is on behavioral measures that are observed, such as event attendance, loyalty card swipes, discounts offered and used at concessions, and other metrics that can be observed in real time by hospitality venue technologies and analyzed using machine learning techniques. What research has shown is that interaction between a company and a customer is becoming a key element in the process of value cocreation. Studies have suggested that methods that capture cocreation of value and the interactive nature of services, that is, customer cocreation behaviors, are needed in the hospitality and tourism literature (Chathoth, Ungson, Harrington, & Chan, 2016).

To create the type of experience customers are looking for, hospitality venues are turning to "big data" to collect and analyze information to better understand their customers so they can more effectively target them with content, deals, and promotions (Price, 2017). The goal is to maximize the entire experience for customers, from initial reservations or membership purchase to postpurchase, keeping them engaged before, during, and after their visit (Prahalad & Ramaswamy, 2004). Consequently, a few researchers have pointed out that there are very limited studies of customer relationship expansion (dynamic cocreation of value and how this influences loyalty) in the hospitality industry (Hyun & Perdue, 2017).

Machine Learning in Hospitality

Typical CRM systems are created to analyze historical data and identify trends in that data. Machine learning, on the other hand, is centered on the ability of computers to "learn" based on preset models and algorithms, without heavy user interaction. The computer programs that are created for machine learning change when exposed to new data.

In the case of loyalty programs, historical data provide a baseline for the creation of algorithms or models that will “learn” as new data enter the system. Using the data from loyalty programs enables a hospitality venue to identify customer-specific information and create a customer-specific profile. This profile allows for an organization’s marketing efforts to become personalized or customized, rather than relying on traditional segmentation methods that profile a homogenous group of customers. While segmentation clusters customers based on demographics, geography, behaviors, or other variables, targeting or customization drills down to individual customers. Vankalo (2003) presents a survey of the different definitions of personalization and customization, and a large discussion of personalization as it relates to marketing can be found in Vesanen (2007), with an extensive review of definitions of relevant terms. We define personalization using Roberts’s (2003, pp. 470-471) characterization of personalization as “the process of preparing an individualized communication for a specific person based on stated or implied preferences.” One of the best examples of personalization is Amazon.com. Amazon uses collaborative filtering to determine what music or books to recommend to users (Blattberg, Kim, & Neslin, 2008).

One of the main purposes of this study was to demonstrate that the machine learning tools used by the retail industry and other types of businesses could be used to increase customer engagement and have a financial impact on the revenue of hospitality venues, if the data are used in a dynamic way to allow the model to continually learn. Before dynamic implementation can be leveraged, it first must be shown there is a benefit to using machine learning over traditional business decisions that are done in a static manner. If this can be shown, then the advantage of using machine learning in a dynamic procedure is evident, since little to no business or personnel interaction is needed in the process. It was the researchers’ belief that different customers would react differently to discounts offered, and machine learning could be used to identify those individuals. If machine learning algorithms were able to identify those individuals better than intuition-based (traditional business decision) modeling, then dynamic procedures could be implemented to automate this decision making. The following hypotheses were tested:

$H2_0$: The percentage increase of expected customer spending produced by intuition-based discounting at the hospitality event concessions is at least equivalent to the percentage increase of expected customer spending when discount offerings are aided by the use of machine learning.

$H2_a$: The percentage increase of expected customer spending produced by intuition-based discounting at the hospitality event concessions is less than the percentage increase of expected customer spending when discount offerings are aided by the use of machine learning.

Customer Loyalty, Loyalty Programs, and Customer Engagement

Customer loyalty is a customer's attachment to a brand, store, manufacturer, service provider, or other entity, based on favorable attitudes and behavioral responses, such as repeat purchase (Baran, Galka, & Strunk, 2008). Customer loyalty enables companies to retain their current customers, and loyal customers are less price sensitive and more vocal in terms of word-of-mouth advertising for the company or brand (Bowen & Chen, 2001; McCall & McMahon, 2016; Tanford, 2016). Loyalty programs are designed to improve customer satisfaction and commitment. These programs usually give away benefits to members as a token of appreciation for their loyalty to the organization. Bolton et al. (2000) state that when customers are involved in loyalty programs, the perceived benefits they receive will lead to their loyalty.

A dynamic cocreation framework offered by Chathoth et al. (2016) proposed that when businesses create or cocreate customer engagement, they use a value chain function by differentiating or customizing value elements based on customer needs and expectations, resulting in a cocreation of value and loyalty. Kandampully, Zhang, and Bilgihan (2015) reviewed 36 studies in the hospitality and tourism literature and proposed that there is a fundamental shift in the way businesses should now build CRM, using a customer-oriented approach to engage customers in a way that affects loyalty, treating customers as the cocreators of value, and formulating longitudinal examinations of customer loyalty. Several researchers have recommended future studies to examine how technology engagement influences customer relationships, explore the need for longitudinal designs to determine customer financial value, and study the use of technology and its influence on CRM (Baloglu, Zhong, & Tanford, 2014; Hyun & Perdue, 2017; McCall & McMahon, 2016; So, King, & Sparks, 2014).

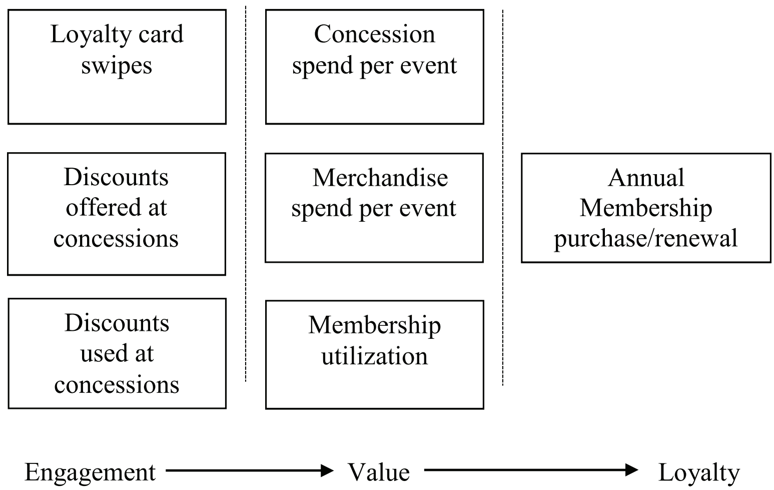
In this study, it was the researchers' belief that when various transactional variables (discounts) based on machine learning were offered, it would increase customer engagement and increase the hospitality venue's ability to predict the customers' likelihood of renewing their annual membership. Those transactional engagement variables differed by customer but included various discounts offered at concessions and discounts offered on merchandise, as well as access to loyalty program events/areas not offered to nonmembers. We refer to these as the collective transactional engagement variables and propose the following hypotheses:

$H3_0$: The collected transactional engagement variables do not increase predictive accuracy of customer loyalty (annual membership renewal).

$H3_a$: The collected transactional engagement variables do increase prediction accuracy of customer loyalty (annual membership renewal).

The engagement-to-loyalty value chain shown in Figure 1 formed the theoretical foundation of the study.

Figure 1
Engagement to Loyalty Value Chain



METHODOLOGY

For this study, the researchers partnered with the management of a large event venue that focuses on hospitality and entertainment. To reach a point where organizations can implement machine learning systems, management must have reason to believe that the offering of discounts and other forms of engagement will result in some benefit. While a system could be implemented immediately, the organization felt that engaging in a trial of discounting before implementing a system would be beneficial in many ways. The management of this organization is revenue focused; therefore, before they would proceed with implementing a system, they wanted evidence of an increase in revenue based on this new program.

Using the investment in a data warehouse that records demographic information for a card-based loyalty program, the researchers and management of the partner organization were able to collect and record purchase behavior and attendance. In this section, the process that led to the implementation of machine learning is described, as well as the data collection and subsequent analysis of results. The partner organization was investigating new methods of customer engagement that were revenue-neutral or revenue-generating. The approach taken was to offer concession discounts to customers. Since the behavior of customers was unknown, a baseline had to be established. Table 1 presents the variables that were collected and used. These variables were selected based on their availability in the CRM system the management team used.

Table 1
Variables Used in Machine Learning Model

Variable Name	Type	Description
Loyalty card swipes	Engagement	Number of times the customer swiped the loyalty card
Discounts offered at concessions	Engagement	The total number of discounted events
Discounts used at concessions	Engagement	Total number of events where discounts offered and used
Concession spend per event	Value	Total amount spent on concessions was discount was offered
Merchandise spend per event	Value	Total amount spent on merchandise when discount was offered
Membership utilization	Value	The number of events for which the customer was ticketed and the ticket was used
Membership renewal	Loyalty	Membership renewal for the upcoming year

These measurements enter the system in real time and can be analyzed by management easily. Because these were the variables management was already using to make decisions, we used these variables for this study. All demographic information described in Table 2 is self-identified from a survey on entry into the loyalty program that is updated annually. Secondary engagement variables and descriptions are also contained in Table 2 and can be derived from the variables contained in Table 1. Other data, such as data collected from social media information, could also be used, but since it was not available for the length of the study, it was not obtained or used in this context.

During the first year, data were collected at four events that were predetermined by management. A card-based loyalty program was used, with a unique barcode for each account, so that when each card was swiped at events, it could be tracked by the CRM. Prior to these four events, each loyalty program member was incentivized to swipe their card to earn rewards through a points reward system (nonmonetary). To create data for the study, during the selected four events, a 20% discount was offered on all food and beverages (excluding alcohol) to a select group of loyalty program members. This created a baseline/treatment group (Adoption Pool). Each of these customers had made concession purchases prior to the study being implemented, and no customer was offered a discount at multiple events. Customers were e-mailed the discount notification on a predetermined weekday prior to the event.

The goal was to see if discounting would increase engagement, specifically revenue-generating engagement. In total, discounts were offered to 3,025

Table 2
Preliminary Data Collected for the Loyalty Program

Variable Name	Type	Description
Income	Demographic	Estimated house hold Income of customer
Gender	Demographic	Customer gender self-identified
Number of users	Demographic	Number of names on the account
Accommodation preference	Demographic	Preferred accommodation in venue
Address	Demographic	Customer home address
Number of subscriptions	Demographic	The number of years a customer has had membership
Discount events attended	Secondary engagement	Total number of events customer attended where discount was offered
Nondiscount events attended	Secondary engagement	Total number of events customer attended where discount was not offered
Total venue spend	Secondary engagement	The total amount spent by customer at the event venue as tracked by the loyalty program
Transaction total	Secondary engagement	The total amount spent in all concession transactions per event

Table 3
Discount for Adopter Pool

Event	% Change in Net Per-User Discount	Sample Size
1-1	–7	100
1-2	20	300
1-3	14	900
1-4	14	1,725

customers in different areas of the venue. The results of this study showed revenue increases for all events and also that over 85% of customers who engaged in the discount (used the discount) spent more at the concession stand than they previously had (gross revenue). The percentage change in net per-user spend after the discount is shown in Table 3.

To test H1, we compared the 85% of customers who engaged in the discount to a 50% baseline that the management of the venue said would be sufficient to continue the program. Using a one sample proportion test with continuity correction in R (Version 3.4.0), the null hypothesis that the true proportion of customers who spend less than their nondiscount average is at most half is rejected ($p < 10^{-16}$, power = 1). The average increase per customer, when compared to their previous average spending, was around a $5 increase, with a standard deviation of around $3. Also, the negative increase seen in Event 1 may have been due to a few customers who drastically underspent, and given the small sample size, this had a large effect.

Given these results, we can reject the null hypothesis $H1_0$, and conclude that randomized discounting encouraged customers who were offered discounts to spend more than their expected spending per event. While these results seemed promising for the discounting program, further investigation indicated that while there was an overall increase in spending, when the management team looked at the increase per loyalty program member and the breakdown of the discount given to increase spending, the spending was not consistent across the venue or even within accommodation areas.

In the second year, the hypothesis that was explored was whether machine learning could be used to personalize the discount offerings and increase customer engagement and spending even more. To further discuss these methods, we need to give the mathematical detail in the most general form possible. The generic form is important for understanding how these models can be used to drive behavior in any setting. Let y_i denote the observed value of the variable that the organization is using as an indicator of engagement (in this case revenue) for the i-th engagement. More generally, we can let x_i be the set of all demographic and engagement history variables for engagement i that we used to predict the engagement in the discounting program. An example of y is the amount of money spent in a discounting program, and if money is spent, the customer had to engage. The predictors for the i-th engagement can be any kind of demographic information or arousal information that the company may have. In event venues, this set of variables will differ for every organization, based on the relevant CRM. We suggest collecting demographic, geographic, and engagement data that could be relevant, because the predictive models will be constrained by the data put in and the ability to predict the response. Contrary to popular belief, there is no one-size-fits-all approach, and the important demographics and engagement information will change, depending on the project and customer insight. The process described below is scalable to many different drivers of engagement that could be identified within any industry, and works on almost any data that could be input.

The data in their simplest form break down to where we observe a (y_i, x_i) pair as a unique engagement for a customer, with demographics and engagement relevant to the time of the transaction. Again, it can't be overstated that these data will vary and change for each business. Assume we have n observations after the end of the last event. The three relevant models to an analysis such as this take on the following form:

1. Score Model:
   $g(y_i) = f(x_i) + \varepsilon_i \quad i = 1, 2, \ldots, n,$
   where g and f are functions and ε_i is an error term for the i-th observation.

2. Probability of Engaging in the Discount:
   p (Person i engaging in Discount Offered) = $(x_i) + \eta_i$ $i = 1, 2, \ldots, n$,
   where P is a probability function, m is a function, and n_i is an error term for the i-th observation.
3. Probability of Selecting an Individual Based on Previous Selections:
   P (Person i engaging in Discount Offered) = $(x_i) + \delta_i$ $i = 1, 2, \ldots, n$,
   where f is a function and δ_i is an error term for the i-th individual.

Given that randomized discounting increased revenue at a per-customer basis, we applied machine learning techniques and personalized marketing to explore the impact of these techniques on spending at a per-customer level. By the end of the fourth event, a data set with over 35,000 concession transactions, each connected to specific loyalty cards, had been created. Using these data, a targeted discounting study was created for the second year of the project, and management carried it out over eight preselected events. The goal was to give decision makers the ability to identify customers who have a pattern that shows they will not only engage in the discount but also spend more when doing so. To evaluate this, we investigated the percentage increase in revenue in each of the four events, and also determined the percentage of customers who spent above their expected amount.

In this study, we define a customer's expected spend to be the average amount the customer has spent on concessions at nondiscounted events. To accomplish this, three models were created: one to predict an increase in revenue, one to predict the probability that a customer would engage, and one to assess the sampling and strategy selection of the discounts. These three models were developed using a proprietary product that implements supervised machine learning algorithms to create a table for decision makers to use in order to formulate the discounting strategy.

From these three models, we created predictions from the most current customer information, and created a multiple-engagement view of how customer behavior and the model has changed. This created the most current prediction of customer behavior based on all relevant and other customer information. From those predictions, decision makers could define what business impact, and what levels of each variable, defined an optimal subgroup.

For comparison purposes, three testing groups were offered discounts for each event. Three customer segments were developed, made up of those groups who machine learning models predicted would generate the highest revenue (Top), those who the models predicted to be the least likely to generate revenue (Bottom), and those who the organization chose by using segmentation and previous spending and/or intuition (Intuition). Customers could appear only in one segment, and if machine learning produced overlap with business intuition, the customer was randomly assigned to that segment.

This strategy was used in each of the first four of the selected events of the second year, and no individual received multiple discounts in the first four

Table 4
Percentage Increase in Revenue Per Event: Top Versus Intuition

	Event	Top Sample Size	Top % Increase	Bottom Sample Size	Bottom % Increase	Intuition Sample Size	Intuition % Increase	*p* For Hypothesis 2	Effect Size	Power
1st Wave of engagement	2-1	265	46	296	6	700	20	.00296	0.322	0.845
	2-2	216	35	249	4	541	13	.00158	0.372	0.899
	2-3	158	27	198	16	414	12	.02746	0.292	0.604
	2-4	155	59	104	19	423	19	0[a]	0.249	0.845
2nd Wave of engagement	2-5	129	13	166	–3					
	2-6	146	9	159	–7					
	2-7	74	26	93	–6					
	2-8	17	22	87	–14					

Note: There are no *p* values during 2nd wave of engagement as there was no intuition study during these events.
[a]50,000 simulations with a *p* value less than 1/50 k = 2×10^{-5}.

events (sampling without replacement from members). After Event 2-4, the Intuition group was dropped because the management team believed it would bias their opinion based on input from the model. This allowed the researchers to test H2.

The groups for Event 2-5 were developed as subsets of those who were offered a discount in Event 2-1, and discount groups were based on the Top, Bottom, and Intuition groups, though all data were used to train the models. Events 1 through 4 were considered the first "wave of engagement," and Events 5 through 8 the second "wave of engagement." The waves describe the sampling strategy for each event. Events 2-6, 2-7 and 2-8 were treated similarly to Events 2-2, 2-3, and 2-4, respectively. We include the results of these events for completeness.

RESULTS

Results from this research indicate that the machine learning methodology–aided discount targeting implemented here is effective in increasing the expected revenue in the Top segment, compared to the Intuition segment. The results of H2 are shown in Table 4 and illustrate the per-event percentage increase in expected revenue. The *p* values were obtained from two-sample randomization testing using 50,000 replications (Lock, Lock, Morgan, Lock, & Lock, 2013). The percentage increase in revenue for the Top segment was larger in comparison to the Intuition segment for all events (*p* values significant at .05 level), implying that machine learning increased the percentage revenue in each event. The results are shown in Table 4 along with power and effect size. All power calculations are presented at the .05 level of significance, and effect size is calculated using Cohen's *d* (Cohen, 1988). We also found that the customers who

Table 5
Percentage Increase in Revenue: Bottom Versus Intuition

	Event	Bottom Sample Size	Bottom % Increase	Intuition Sample Size	Intuition % Increase	*p*	Effect Size	Power
1st Wave of engagement	1	296	6	700	20	.013	0.2181	0.705
	2	249	4	541	13	.086	0.149	0.365
	3	198	16	414	12	.718	–0.075	0.015
	4	104	19	423	19	.514	–0.006	0.055

Table 6
Percentage of Customers With an Increase in Spending Per Event: Top Versus Intuition

	Event	Top Sample Size	Top % Increase	Intuition Sample Size	Intuition % Increase	*p*	Effect Size	Power
1st Wave of engagement	1	265	61	700	56	.07	0.100	0.35
	2	216	59	541	51	.035	0.160	0.468
	3	158	60	414	52	.063	0.160	0.388
	4	155	70	423	52	0[a]	0.370	0.95

[a]*p* value for Event 4 is less than 7.83×10^{-5}

we targeted with machine learning in the Top segment spent $1 (Event 1) to $6 (Event 4) more per transaction when compared to the Intuition segment, showing managerial impact.

A similar study was done to compare the Bottom segment to Intuition. The results are reported in Table 5. Power and effect size are included for completeness.

To evaluate the increase in engagement, we compared the percentages of customers who spent (net spend after discount) more than their average nondiscounted event spend. The results of this two-sample proportion test are presented in Table 6 for Top versus Intuition groups. All results are considered significant at the 0.10 level. Power at the 0.05 significance level and effect size are also included.

A similar study was done to compare the Bottom segment to Intuition. The results are reported in Table 7. No results are significant. Power and effect size are included for completeness. Both results reported in Tables 5 and 6 may suffer from low power due to the observational nature of the study.

Given these results, we were able to reject the null $H2_0$ and conclude that the percentage increase of expected customer spending produced by intuition-based discounting at the hospitality event concessions was less than the percentage

Table 7
Percentage of Customers With an Increase in Spending Per Event: Bottom versus Intuition

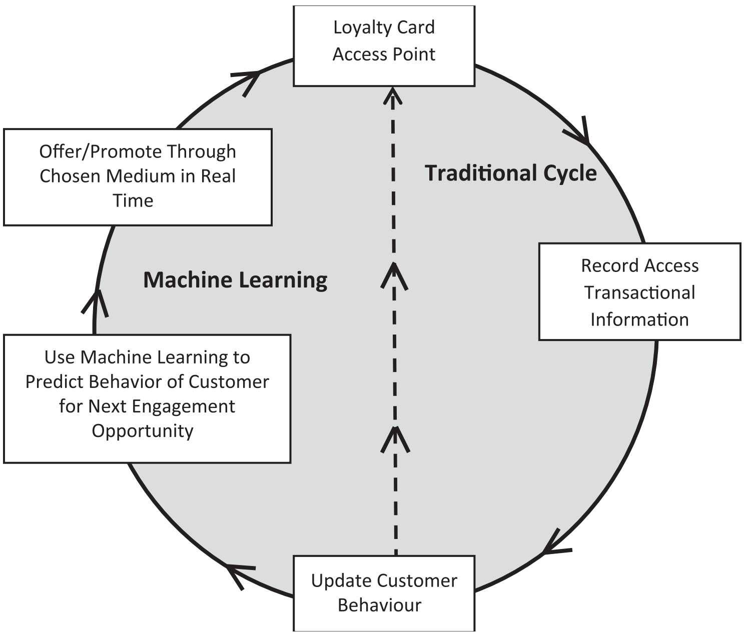
	Event	Bottom Sample Size	Bottom % Increase	Intuition Sample Size	Intuition % Increase	*p*	Effect Size	Power
1st Wave of engagement	1	296	52	700	56	.184	0.080	0.20
	2	249	49	541	51	.278	–0.040	0.14
	3	198	55	414	52	.775	–0.060	0.17
	4	104	55	423	52	.514	–0.060	0.10

increase of expected customer spending when discount offerings are aided by machine learning.

Finally, the renewal information based on engagement data was analyzed to test H3. Although the model used for renewal in this case is proprietary, the variable importance metrics (similar to the resulting *p* values of hypothesis tests in regression), indicated that both membership utility and engagement in the loyalty program (total spend) had a positive effect on predicting annual membership renewals, while variables such as gender and accommodation preference were found to be of no importance in renewal. The final model had a 17% classification error rate on 15,000 loyalty program members, meaning the organization correctly predicted 83% who would renew. After removing the transactional engagement variables discussed, the classification error rate increased to 27%, meaning the variables have predictive importance when other demographic information is taken into account for this organization.

The results indicate a significant difference between the two classification error rates using a one-sided test, with the outcome that the model with machine learning–generated engagement variables outperforms the model without the engagement variables, with a p value less than 10^{-16} and power of 1 (Dietterich, 1998). Thus, we were able to reject the null hypothesis $H3_0$ and conclude that using engagement variables increased the prediction accuracy of membership renewals. Note, we do not conclude that any of these relationships are causal, or that we have driven or created loyalty by driving engagement, since machine learning tools typically lack the ability to imply causation in a nonexperimental setting; however, we are able to say that the inclusion of the engagement variables caused a statistically significant increase in prediction accuracy. As every customer predicted is estimated on average to produce $700 in revenue, the inclusion of the engagement variables was able to provide insights on up to $1.05 million in possible revenue for the organization for the coming year in that venue. These results show a return on investment of the strategy of using machine learning and the loyalty program.

Figure 2
Dynamic Customer Engagement Relationship



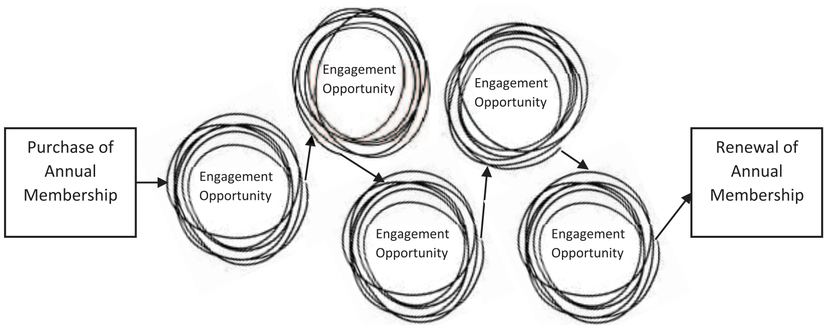
Concerns may exist due to the different sample sizes of the groups listed and the resulting effects on power, but this is common business practice in predictive data analytics. Machine learning methods are developmental, and until they can be tested with full implementation, there is a high risk of decreased revenue. More important, the use of smaller sample sizes allowed management to study the loyalty program and evaluate the best models to use prior to investing significant amounts of money in what could turn out to be the wrong program. This led to small sample sizes and lower than desired power in certain studies.

Theoretical and Managerial Implications

The most significant theoretical contribution of this study has been to demonstrate the concept of *dynamic customer engagement*. By updating the perceived customer value using machine learning each time a transaction occurs, we were able to identify what customers value, in real time. This is shown in Figure 2.

Machine learning can be used to determine what customers actually value in the relationship between the customer and the brand. Machine learning techniques provide a method to capture the cocreation of value in real time in an interactive, natural setting that does not require the measurement of customer perceptions after the fact—perceptions often based on flawed memories of their

Figure 3
Dynamic Customer Engagement Relationship Over Time With Machine Learning at All Points



experience with the hospitality venue. When actual behaviors, both transactional and nontransactional, can be captured and added to the database, and the machine learning algorithms use that data in real time to refine stored knowledge and information on each customer, it allows managers to personalize the loyalty card system to fit the customers, perhaps even while they are interacting with the venue. This is what we term *dynamic customer engagement*. This has the dual goal of increasing revenue as well as increasing managers' ability to better predict future behavior based on increased loyalty. Therefore, this study has both theoretical and managerial contributions.

The dynamic nature of customer engagement over time is shown in Figure 3. Because machine learning models are continuously predicting behaviors as new data enter the database, customer engagement is continuously dynamic. The adaptive nature of machine learning allows for predictions to change when the machine learning model is updated, which will update predictions for all customers. This allows for behavioral trends to be leveraged when the model is updated. If specific customer behavior changes, the predictions for the customer will change as well, possibly creating a new interaction or engagement opportunity.

In traditional CRM systems, customer engagement is still dynamic, but after customer information is updated, offers and promotions cannot be adjusted in real time. Newer CRM offerings have tools that can track customer engagement on social media in real time, acknowledging the dynamic nature of customer engagement, but again offers and promotions need adjusted in real time. Machine learning allows us to leverage real-time behavioral information to cocreate value in the dynamic customer engagement model proposed in this research. In cases where there are many transactions, managers will not have the same capacity to promote (or adjust) personalized engagement opportunities; in contrast, automated processes such as machine learning are capable of identifying these potentially missed engagement opportunities.

Given the dynamic nature of customer engagement, the ability to capture and measure each point of engagement enriches our understanding of the concept. The ability to adjust offerings in real time based on a customer's dynamic behavior enriches our understanding of customer value and creates insights into their loyalty. This is one of the first studies that uses the concept of machine learning in the customer engagement-to-loyalty chain in the hospitality and tourism industry. Moreover, this article illustrates the dynamic nature of customer engagement, showing that machine learning predicts and updates analyses of customer behavior in real-time, an innovative feature that is critical in the dynamic industry of hospitality and tourism. This study bridges the gaps in current knowledge, theory, and traditional industry practices, which have always focused on just the transactional and historic information to engage with customers and to create loyalty-enhancing value. This study changes the way hospitality and tourism venues and businesses—hotels, resorts, theme parks, cruise ships, convention centers, and so on—view customer engagement and use machine learning to predict the behavior of a customer at different engagement opportunities. Furthermore, managers in the hospitality and tourism industry can now use this information to publish offers or promotions via different mediums in real time to engage with the customers in ways that would directly or indirectly influence their loyalty. In the future, researchers can use machine learning in other hospitality venues and implement a similar *dynamic customer engagement* relationship to the one introduced in this article to create customized experiences, understand cocreation of value, and improve customer loyalty.

DISCUSSIONS AND CONCLUSIONS

This study demonstrates how machine learning can be used to increase our understanding of what customers value in the engagement-to-loyalty value chain. It also adds to the theory by proposing a framework for measuring and analyzing *dynamic customer engagement*. This provides methodology for adjusting offers and promotions of products and services to influence customer engagement behaviors in real time, which bridges research gaps on the application of machine learning in the hospitality and tourism literature. We have demonstrated how hospitality venues can cocreate value by offering personalized discounts that influence loyalty. Our methodology gives researchers a way to identify procedures where automation can be implemented, and this dynamic engagement can be exploited. In the future, researchers should also explore the dynamic link between engagement and loyalty. With the capability to gather customer preference data at each engagement opportunity, the link to loyalty can also be explored in a dynamic fashion using machine learning. Using this model, changes in loyalty can also be ascertained in real time.

Regarding hospitality organizations that manage large event venues, such as theaters, concert halls, arenas, convention halls, and theme parks, the use of predictive data analytics and machine learning is new. Machine learning in this study is used as an example for automated processes that create insights into cocreation of value based on dynamic customer engagement. To do this we described how a large hospitality venue engaged customers with discounts in an automated way that increases their spending and revenue, and when these conditions combined, they led to a better understanding of loyalty at an individual customer level.

In the hospitality industry, offering discounts does not fit the culture or strategy of some organizations, and these results may not apply to them. However, when used appropriately, these techniques can be used in other areas of the loyalty system model, such as points, rewards, upgrades, miles, and so on. Also, fully dynamic customer engagement may not fit the strategy of an organization, and managing the number of offerings to customers has not been studied here. In our study, discount offerings were perpetual throughout each event, but this may not be realistic for all companies. Methodology to determine the number of offerings, and the optimal timing of these offerings, in dynamic settings is still an open research problem. The concepts of engagement, cocreation of value, and loyalty may differ between businesses and their brands, so organizations should choose relevant variables when developing their engagement-to-loyalty value chain model.

Finally, in a broader context within the hospitality industry, managers should be involved in the creation of the models to be tested for different loyalty members. They should also be involved in the modification of those models based on customer engagement, perceived value, and loyalty. For instance, in the lodging industry, one of the challenges for managers is to cross-sell engagement opportunities such as restaurants, spas, retail, and other services. Like the within-event comparisons of this study, hospitality organizations may not have the same customer visiting the business, but the same customer may be visiting the same brand or partner brand; therefore, it is important to create these loyalty programs within the brand or with partners to improve guest experiences by engaging with them in ways that lead to increased loyalty. This study reveals that using machine learning techniques can significantly improve an organization's ability to enhance the guest experience and substantially increase revenue.

In the context of big data, the concept of engaging customers in cocreation of value in a way that leads to brand loyalty is still in early development (Xiang et al., 2015). In the future, algorithms to develop highly targeted marketing strategies need to be developed, particularly ones that use the massive amount of customer information available in all loyalty programs. We recommend that academic researchers partner with industry practitioners to continue to develop hospitality-related machine learning methods, and investigate the impact and best practices of big data in the hospitality industry. In future studies, researchers in hospitality should improve the knowledge base of big data application in the industry.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The author(s) received no financial support for the research, authorship, and/or publication of this article.

REFERENCES

Allenby, G. M., Bradlow, E. T., George, E. I., Liechty, J., & McCulloch, R. E. (2014). Perspectives on Bayesian methods and big data. *Customer Needs and Solutions*, *1*, 169-175.

Athey, S. (2015, August). Machine learning and causal inference for policy evaluation. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 5-6). New York, NY: ACM.

Baloglu, S., Zhong, Y. Y., & Tanford, S. (2014). Casino loyalty: The influence of loyalty program, switching costs, and trust. *Journal of Hospitality & Tourism Research*, *41*, 846-868.

Baran, R. J., Galka, R. J., & Strunk, D. P. (2008). *Principles of customer relationship management*. Mason, OH: Thomson South-Western.

Blattberg, R. C., Kim, B. D., & Neslin, S. A. (2008). *Database marketing: Analyzing and managing customers*. New York, NY: Springer.

Bolton, R. N., Kannan, P. K., & Bramlett, M. (2000). Implications for loyalty program membership and service experiences for customer retention and value. *Journal of Academy of Marketing Science*, *28*(28), 95-108.

Bowden, J. (2009). The process of consumer engagement: A conceptual framework. *Journal of Marketing Theory and Practice*, *17*, 63-74.

Bowen, J. T., & Chen, S. L. (2001). The relationship between customer loyalty and customer satisfaction. *International Journal of Contemporary Hospitality Management*, *5*(13), 213-217.

Brodie, R. J., Hollebeek, L. D., Ilic, A., & Juric, B. (2011). Customer engagement: Conceptual domain, fundamental propositions & implications for research in service marketing. *Journal of Service Research*, *14*, 252-271.

Brodie, R. J., Ilic, A., Juric, B., & Hollebeek, L. (2013). Consumer engagement in a virtual brand community: An exploratory analysis. *Journal of Business Research*, *66*, 105-114.

Chan, K. W., & Li, S. Y. (2010). Understanding consumer-to-consumer interactions in virtual communities: The salience of reciprocity. *Journal of Business Research*, *63*, 1033-1040.

Chathoth, P. K., Ungson, G. R., Harrington, R. J., & Chan, E. S. W. (2016). Co-creation and higher order customer engagement in hospitality and tourism services: A critical review. *International Journal of Contemporary Hospitality Management*, *28*, 222-245.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, *10*, 1895-1923.

Di Gangi, P. M., & Wasko, M. (2009). The co-creation of value: Exploring user engagement in user-generated content websites. *Proceedings of JAIS Theory Development Workshop*, *9*(50), 9-50.

Fernandes, F., & Esteves, F. (2016). Customer engagement and loyalty: A comparative study between service contexts. *Services Marketing Quarterly*, *37*, 125-139.

Grimmer, J. (2015). We are all social scientists now: How big data, machine learning, and causal inference work together. *PS: Political Science & Politics*, *48*, 80-83.

Groeger, L., Moroko, L., & Hollebeek, L. (2016). Capturing value from non-paying consumers' engagement behaviors: Field evidence and development of a theoretical model. *Journal of Strategic Marketing*, *24*, 190-229.

Gruen, T. W., Osmonbekov, T., & Czaplewski, A. J. (2006). EWOM: The impact of customer-to-customer online know-how exchange on customer value and loyalty. *Journal of Business Research*, *59*, 449-456.

Guyon, I., & Elisseeff, A. (2006). An introduction to feature extraction. In I. Guyon, M. Nikravesh, S. Gunn, & L. Zadeh (Eds.), *Feature extraction: Vol. 207. Studies in fuzziness and soft computing* (pp. 1-25). Zurich, Switzerland: Springer.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning, data mining, inference, and prediction*. New York, NY: Springer-Verlag.

Hollebeek, L. D. (2011). Demystifying customer brand engagement: Exploring the loyalty nexus. *Journal of Marketing Management*, *27*, 785-807.

Hollebeek, L. D., Conduit, J., & Brodie, R. J. (2016). Strategic drivers, anticipated and unanticipated outcomes of customer engagement. *Journal of Marketing Management*, *32*, 393-398.

Hyun, S. S., & Perdue, R. R. (2017). Understanding the dimensions of customer relationships in the hotel and restaurant industries. *International Journal of Hospitality Management*, *64*, 73-84.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2015). *An introduction to statistical learning with applications in R*. New York, NY: Springer.

Jennings, S., Giorgio, P., Murali, R., & Goggin, S. C. (2014). *Winning the race for guest loyalty*. Boston, MA: Deloitte Development. Retrieved from https://www2.deloitte.com/us/en/pages/consumer-business/articles/winning-the-race-for-guest-loyalty-hotels.html

Kandampully, J., Zhang, T., & Bilgihan, A. (2015). Customer loyalty: A review and future directions with a special focus on the hospitality industry. *International Journal of Contemporary Hospitality Management*, *27*, 379-414.

Liu, Y., Teichert, T., Rossi, M., Li, H., & Hu, F. (2017). Big data for big insights: Investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews. *Tourism Management*, *59*, 554-563.

Lock, R. H., Lock, R. H., Morgan, K. L., Lock, E. F., & Lock, D. F. (2013). *Statistics: Unlocking the power of data*. Hoboken, NJ: Wiley.

Marr, B. (2016, January). How big data and analytics are changing hotels and the hospitality industry. *Forbes*. Retrieved from https://www.forbes.com/sites/bernardmarr/2016/01/26/how-big-data-and-analytics-changing-hotels-and-the-hospitality-industry/

McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*, *90*, 60-66. Retrieved from https://hbr.org/2012/10/big-data-the-management-revolution

McCall, M., & McMahon, D. (2016). Customer loyalty program management, what matters to the customer? *Cornell Hospitality Quarterly*, *57*, 111-115.

Mitchell, T. M. (2006). *The discipline of machine learning* (Technical Report CMU-ML-06-108). Pittsburgh, PA: Carnegie Mellon University.

Nielsen. (2016). Discounts dominate but loyalty program rewards go deeper than dollars. *Nielsen Insights*. Retrieved from http://www.nielsen.com/us/en/insights/news/2016/discounts-dominate-but-loyalty-program-rewards-go-deeper-than-dollars.html

Patterson, P., Yu, T., & DeRuyter, K. (2006, December). *Understanding consumer engagement in services*. *Paper presented at the ANZMAC 2006 conference: Advancing Theory, Maintaining Relevance*, *Brisbane, Queensland, Australia*.

Prahalad, C. K., & Ramaswamy, V. (2004). Co-creation experiences: The next practice in value creation. *Journal of Interactive Marketing*, *18*, 5-14.

Price, B. (2017). An interview with Pete Gibson CEO Datlytics, Ft. Lauderdale Florida, USA. *Journal of Global Information Technology Management*, *20*(2), 1-4.

Roberts, M. L. (2003). *Internet marketing: Integrating online and offline strategies*. Boston, MA: McGraw-Hill/Irwin.

Sashi, C. M. (2012). Consumer engagement, buyer-seller relationships, and social media. *Management Decisions*, *50*, 253-272.

Sawhney, M., Verona, G., & Prandelli, E. (2005). Collaborating to create: The internet as a platform for consumer engagement in product innovation. *Journal of Interactive Marketing*, *19*(4), 4-17.

Schouten, J. W., McAlexander, J. H., & Koenig, H. F. (2007). Transcendent customer experience and brand community. *Journal of the Academy of Marketing Science*, *35*, 357-368.

So, K. K. F., King, C., & Sparks, B. A. (2014). Customer engagement with tourism brands, scale development and validation. *Journal of Hospitality & Tourism Research*, *38*, 304-329.

So, K. K. F., King, C., Sparks, B. A., & Wang, Y. (2016). The role of customer engagement in building consumer loyalty to consumer loyalty brands. *Journal of Travel Research*, *55*, 64-78.

Tanford, S. (2016). Antecedents and outcomes of hospitality loyalty: A meta-analysis. *Cornell Hospitality Quarterly*, *57*, 122-237.

Thakur, R., & Summey, J. H. (2010). Optimizing CRM: A framework for enhancing profitability and increasing lifetime value of customers. *Marketing Management Journal*, *20*, 140-151.

Vankalo, M. (2003). *Internet-enabled techniques for personalizing the marketing program* (Unpublished Master thesis). Swedish School of Economics and Business Administration, Helsinki, Finland.

Vargo, S. L., & Lusch, R. F. (2008). From goods to service(s): Divergence and convergences of logics. *Industrial Marketing Management*, *37*, 254-259.

Vargo, S. L., Maglio, P. P., & Akaka, M. A. (2008). On value and value co-creation: A service systems and service logic perspective. *European Management Journal*, *26*, 145-152.

Vesanen, J. (2007). What is personalization? A conceptual framework. *European Journal of Marketing*, *41*, 409-418.

Vivek, S. D., Beatty, S. E., & Morgan, R. M. (2012). Consumer engagement: Exploring customer relationships beyond purchase. *Journal of Marketing Theory and Practice*, *20*, 127-145.

Weisberg, S. (2014). *Applied linear regression* (4th ed.). Hoboken, NJ: Wiley.
Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, *44*, 120-130.

Submitted June 23, 2017
Accepted December 12, 2017
Refereed Anonymously

Ajay Aluri (e-mail: ajay.aluri@mail.wvu.edu) is an assistant professor and director of the Hospitality Innovation and Technology (HIT) Lab in the Hospitality and Tourism Management program in the Department of Management at the College of Business and Economics at West Virginia University. His research interests are in the areas of hospitality information technology and management and information systems, and how these areas impact international hospitality organizations, society, and education. **Bradley S. Price** (e-mail: brad.price@mail.wvu.edu) is an assistant professor of Business Data Analytics in the College of Business & Economics at West Virginia University. His research interests are in statistical machine learning, statistical computing and data science for business. **Nancy H. McIntyre** (e-mail: nancy.mcintyre@mail.wvu.edu) is an associate professor of Management in the College of Business & Economics at West Virginia University. Her research interests include extrepreneurship, self concept-based motivation, curiosity, and team processes.