Evaluating Causal Relationships Between Customer Satisfaction and Loyalty in Consumer Mobility Services

HENLEY BUSINESS SCHOOL

THE UNIVERSITY OF READING

Doctorate of Business Administration (DBA)

Henley Business School

Christopher Dale McKenzie

May 2022

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Certificate of readiness to be included in the library

I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to me. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract

The study examines the relationships between antecedents of customer satisfaction and customer loyalty, and between satisfaction and loyalty itself, in the mobile phone industry in Canada. This industry is highly competitive, with customer loyalty (measured as percentage of a company's customers that deactivate their service or "churn" each month) a key metric impacting the firm's financial performance and performance within the public investment market. While research over the past 40 plus years provides support for the existence of the relationships noted previously, there are few studies that empirically prove a causal relationship.

As a result, practitioners face the challenge of not always knowing if their investments to improve customer satisfaction are the right investments to drive causal improvements in that metric, or ultimately, in customer loyalty. Mobile service providers spend almost 15% of their revenues on costs associated with retaining customers, and a further 10% replacing those that do leave.

A randomized controlled trial (RCT) is conducted as a field study over twelve months with 35,000 mobile phone subscribers, a study design not seen in prior research. Three treatments are designed and applied to a factorized set of treatment groups; these treatments are based on antecedents identified in prior research, and include treatments based on Perception of Value, Perception of Service, and Frustration with Marketing Communications. The study design measures the impact of these treatments on loyalty on a monthly basis throughout the trial, and the impact on customer satisfaction is measured cumulatively at the end of twelve months. Analysis is conducted primarily through linear regression supplemented by logistic and probit regressions, and other statistical techniques as appropriate.

The results of the study provide support for the positive relationship between the antecedents employed in the RCT and customer satisfaction but does not establish an ability to causally influence those antecedents - or customer satisfaction - through the use of the treatments. Evidence is found to support a causal relationship between the treatment reducing the frequency of marketing messaging on participant Frustration with Marketing Messaging, and of a causal relationship between the combination of two treatments (reduction in marketing messaging and a monetary discount) on loyalty. A cost/benefit analysis of these treatments and the accompanying benefit of increased loyalty indicates that companies can increase the lifetime value of their customers through broader deployment of the indicated treatments.

Contents

List of figures

List of tables

List of abbreviations

ABPU	Average Billing per User
BYOD	Bring Your Own Device
B2C	Business-to-Consumer
CCTS	Commission for Complaints for Telecom-Television Services
CI	Confidence Interval
COA	Cost of Acquisition
COR	Cost of Retention
CRTC	Canadian Radio-television and Telecommunications Commission
CSAT	Customer Satisfaction
DNC	Do Not Contact
DNS	Do Not Survey
DV	Dependent Variable
EBITDA	Earnings Before Interest, Taxes, Depreciation and Amortization
IV	Independent Variable
IVR	Interactive Voice Response
L2R	Likelihood to Recommend
MRC	Monthly Recurring Charge
MTM	Month-to-month
T2B	Top Two Box

1 Background to the Study

Despite an extensive body of research examining consumer behaviours in the areas of customer satisfaction and loyalty, scant proof has been offered by prior research to substantiate causal relationships between various proposed antecedents of, and their impact on, satisfaction and loyalty. Utilizing experimental methodologies for examining causal relationships between variables, this study examines customers within the consumer mobile phone market in Canada to determine if such causal relationships can be established. The study utilizes a randomized controlled trial method of 35,000 customers of a large telecommunications company, in conjunction with three separate treatments applied over a period of twelve months, in a natural field setting.

The following sections of this chapter establish the context for the study through a review of the mobile phone market in Canada, a description of the research problem to be examined, an outline of the scope and aims of the study, and a summary of the potential implications for, and significance to, business practice. The chapter closes with an outline of the remaining sections of the study.

1.1 Study Context

The mobile phone industry in Canada is a mature market, growing from the first cell phone call on July 1, 1985, and becoming a CAN$28.3 billion[1] industry by 2018, comprising 53.3% of the total $53.1 billion telecommunications market (inclusive of wireline data and voice services). In 2018, EBITDA margins for the wireless portion of the market revenue were an average of 41.0%[2], commensurate for capital-intensive industries such as telecommunications, but also the oil and gas, railroad, mining and semiconductor industries.

The retail mobile sector consisted of 33,200,000 subscribers at the end of 2018 (from a Canadian population of 37,238,906[3]), with 90.7% of wireless revenues accounted for by the three largest national mobile service providers (the “Top 3”). These Top 3 providers include Rogers Communications, Inc. (“Rogers”, with 33.3% of 2018 wireless revenue share), Bell Canada Enterprises, Inc. (“Bell”, 29.6% share) and TELUS Corp. (“TELUS”, 27.8% share). Several smaller, regional mobile providers account for the remaining 9.3% of revenue share[2].

[1] Unless otherwise noted, all dollar amounts in this study are expressed in Canadian dollars
[2] CRTC. Communications Monitoring Report (2019)
[3] Statistics Canada. Table 17-10-0009-01 Population estimates, quarterly

The mobile phone market provides two types of services: post-paid and pre-paid. Post-paid service bills a customer at the end of each monthly period and comprised 97.1% of the mobile services revenue of the Top 3 in 2018. Pre-paid service, where the customer pays upfront (for example, by adding funds to a "top-up" account in advance of usage), comprised the remaining 2.9% of revenue[2].

Additionally, each of the Top 3 market their mobile services under a combination of a primary brand (positioned as a premium service under the same brand as the company itself – such as TELUS, Bell or Rogers – and accounting for 75% of revenue share and almost 89% of subscribers) and one or more flanker brands (positioned for more value- and price-conscious consumers – using sub-brand names such as Fido, Koodo, and Virgin Mobile – and accounting for the remaining 25% of revenue/11% of subscribers)[2].

Finally, these Top 3 companies provide mobile phone services on a national basis, with each company also providing landline (Internet, television, home phone) services regionally. For example, TELUS offers landline services in British Columbia and Alberta, while Bell offers landline services in Ontario and Quebec.

Given the dominance of the Top 3 providers, the following sections will focus on the dynamics amongst these three companies with additional focus on the post-paid segment of the market, which comprises over 97% of the revenues for the Top 3.

1.1.1 Competitive Nature of the Market

With three strong national carriers competing in each market (along with one or more regional carriers), there is a high level of competition amongst the companies to acquire and retain customers. As noted in one company's recent annual report:

> "The Canadian wireless market continues to be characterized by high levels of acquisition and retention activities and the associated high costs of device subsidies on two-year contracts, heightened competitive intensity, and the continued adoption of high-value, data-centric smartphones."[4]

Mobile service customers are, generally, free to move amongst carriers. Government-mandated policies such as wireless number portability ("WNP", allowing customers to move their cellular phone numbers for free amongst carriers), device unlocking (allowing smartphones bought from one carrier to work on another carrier's network) and a maximum 24-month contract period facilitates such movement. Each carrier markets their offerings

[4] BCE Inc. 2019 Annual Report

frequently, especially during crucial buying times of the year (for example, Back-to-School, Black Friday and Christmas/Boxing Day) and each carrier provides a wide variety of sales channels in which to purchase a mobile phone and service (including physical stores, call centre and online web stores).

1.1.2 Volumes of Customer Switching Behaviour

Even though each carrier reports year-over-year growth in their total number of wireless subscribers, the volume of switching activity by customers between the carriers eclipses the net increase in the customer base (the net increase being the result of a combination of population growth and increased penetration of services).

As shown in Table 1.1, the 2018 net addition of just under 1.3 million subscribers by the Top 3 required gross additions (new subscriber activations) of almost 4.4 million. The difference, roughly 3.1 million, were subscribers that deactivated (or churned) from one provider to another.[5] These subscribers churn for a variety of reasons, including dissatisfaction with the service received, pricing issues, or to take advantage of an offer from a competing company. Churn is viewed, both by the organizations themselves and the investor community, as the most explicit indicator of customer loyalty in the wireless services market.

Table 1.1 2018 Post-paid Subscriber Metrics for Top 3 Carriers

Metric	TELUS	Bell	Rogers	Top 3 Total
Gross Additions	1,150	1,616	1,632	4,398
Deactivations	794	1,165	1,179	3,141
Net Additions to Subscriber Base	356	453	448	1,257
Ending Subscriber Base	8,311	9,157	8,830	26,298
Average Monthly Churn[a]	0.89%	1.16%	1.10%	1.05%[b]

Note. All figures in 000's except Average Monthly Churn. Results obtained from 2018 annual reports for each carrier.

[a]Monthly Churn = Total Deactivations for Month/[(Beginning Month Subscribers + Ending Month Subscribers)/2]

[b]Non-weighted average

[5] A small percentage of churn can be attributed to non-voluntary deactivations (e.g. for non-payment) or other non-competitive reasons such as customer death, move out of country, etc.

The implication of this volume of switching activity is two-fold: firstly, companies must acquire approximately 3.5 new customers for every one net customer they add to their customer base; and secondly, they must expend substantial effort to contain the switching/deactivation volumes that do occur. Both of these efforts require the expenditure of financial resources.

1.1.3 Financial Implications of Churn

The financial implications of customer churn for companies are substantial and can be examined through simple calculations.

In 2018, the Top 3 carriers reported an Average Billing per User (ABPU) of $66.60 per month. ABPU is the average monthly bill, including recurring charges (such as monthly plan fees) and non-recurring usage charges (including data overages and roaming).

Using a non-weighted average monthly churn rate of 1.05%, the average subscriber stays with a given carrier for approximately 95 months (calculated as 1 / monthly churn, or 1 / 0.0105 = 95.24 months). Ninety-five months of tenure at $66.60/month yields an approximate lifetime revenue of $6,327 per subscriber. The Top 3 carriers each have an average base of 8.8 million subscribers, therefore 1 basis point of monthly churn (0.01%) is equivalent to 10,560 deactivations per year per carrier, with a lifetime revenue value of $66.8 million (10,560 x $6,327 = $66,813,120). A lower churn rate is a strong financial advantage in this industry.

As shown in Table 1.1, despite roughly similar customer base sizes (and overall revenue market share), there is a significant difference in the average monthly churn rates amongst the Top 3: TELUS at 0.89% has a monthly churn rate that is 27 basis points and 21 basis points lower (better) than its competitors Bell and Rogers, respectively. Retaining such a churn advantage or closing the gap with a competitor who holds that advantage, is a primary motivator for companies to improve their customers' satisfaction, and hopefully, their loyalty.

1.1.4 Costs to Acquire or Retain Customers

As discussed in the previous two sections, net subscriber growth from a given carrier can come from two sources: gross additions (new subscribers to the company, often as churn from another carrier) or improved retention of a company's existing subscribers (that is, lower deactivations or churn). Both sources require the expenditure of substantial financial resources, as summarized in Table 1.2.

Table 1.2 2016 Costs of Acquisition and Retention

Cost Metric	TELUS	Bell
Cost of acquisition (COA) per gross addition	$455	$494
Gross Additions (000's)	1,039	1,408
Total COA cost ($millions)	$472.7	$695.6
Cost of retention (COR) spend as % of network revenue	14.7%	13.2%
Network revenue[a] ($millions)	$6,541.0	$6,602.0
Total COR cost ($millions)	$961.5	$871.5
Total COA and COR cost ($millions)	$1,434.2	$1,567.1
Total COA and COR cost as % of network revenue[a]	21.9%	23.7%

Note: Data from 2016 annual reports, the last year that TELUS and Bell reported COA and retention spend in their annual reports; Rogers did not report these metrics.

[a]Network revenue is the revenue derived directly from the wireless services provided to customers and commonly used within the industry as the denominator for cost comparisons

Total COA and COR spend comprised nearly one-quarter of total network revenues for Bell and TELUS (23.7% and 21.9%, respectively). With EBITDA margins of 41.9% (Bell) and 42.2% (TELUS), the combined COA and COR costs represented 40.8% and 37.9% of total operating costs for each company. Although Rogers did not report COA or retention spend in their annual reports, it is reasonable to assume that their spend on these two categories was in the same range as their competitors.

From 2016 to 2019, change in adjusted EBITDA margin was flat for Bell (41.9% in 2016, 42.0% in 2019), slightly higher for TELUS (42.2% in 2016, 45.4% in 2019) and Rogers (45.3% in 2016, 47.0% in 2019). The small improvements in EBITDA margin over four years do not indicate any material reduction in COA or COR costs since the companies ceased reporting these metrics.

With the Top 3 carriers each spending in the range of $1.5 billion annually to acquire or retain subscribers, it is clear to see why they might focus on customer satisfaction as a means to drive customer loyalty, and therefore reduce churn.

1.1.5 Industry Focus on Customer Satisfaction

Given the substantial financial benefits accruing to those companies that can achieve a lower churn rate - relative to their own performance and to that of their peers - it is not surprising that the industry expends great resources on trying to minimize churn. The broadest approach appears to be through their efforts to provide positive customer experiences, and efforts to ensure high levels of customer satisfaction, under the belief that such satisfaction is a prerequisite for greater loyalty (and therefore lower customer churn).

Indeed, each of the Top 3 organizations has expressed their desire to deliver a positive customer experience for some time. As early as 2008, Bell communicated in its annual report that it had "…instituted a strategy to deliver a dramatically better customer experience…".

Perhaps coincidentally, the year prior (2007) had seen the launch of the Commission for Complaints for Telecommunications Services (CCTS), an independent, industry-funded agency to resolve complaints from consumers and small business retail telecom customers. The CCTS also administers the Wireless Code[6], created in June 2013 by the Canadian Radio-television and Telecommunications Commission (CRTC). This Code defines minimum standards that each wireless service provider must meet, including customer rights and responsibilities, and provides a mechanism to submit complaints to the Commission if an issue under the Code is not resolved. These complaints are then reported publicly.

The 2019 annual reports for each of the Top 3 carriers make multiple references to their efforts to improve customer satisfaction and customer experience and prominently record their accomplishments and awards for doing so.

1.1.6 Industry Results in Improving Customer Satisfaction

J.D. Power conducts a series of annual reports within the telecommunications market in Canada. They publish four reports annually, each focused on a different aspect of the customer experience with each carrier: Customer Care, Network Quality, Purchase Experience, and Total Ownership Experience. The results provide a relative ranking of each carrier within and across years.

According to J.D. Power[7], the Wireless Customer Care Study "…measures satisfaction with the entire customer care service experience, from initial contact to problem resolution, based on recent customer experiences across three point-of-contact channels: phone, retail walk-in,

[6] The Wireless Code https://crtc.gc.ca/eng/phone/mobile/codesimpl.htm
[7] J. D. Power, Canada Wireless Customer Care Study https://canada.jdpower.com/tmt/canada-wireless-customer-care-study

and online. The study provides a competitive benchmark to measure customer advocacy and loyalty; identify improvement initiatives; and compare wireless carrier performance across Canada."

Table 1.3 shows the results of the J.D. Power Customer Care Study for the Top 3 carriers and the industry overall.

Table 1.3 J.D. Power Overall Customer Satisfaction Index results (1,000 point scale)[a]

Carrier	2015	2016	2017	2018	2019	2020	Change '20 vs '15
TELUS	745	771	769	756	786	768	+23
Rogers	675	707	713	741	734	746	+71
Bell	690	706	723	726	724	724	+34
Top 3 Average (premium + flanker)	730	749	766	762	770	763	+33
Industry Average	715	738	746	753	761	755	+40

[a]J.D. Power Canada Wireless Customer Care Study (2015 through 2020)

Each carrier, and the industry as a whole, shows improvement from 2015 to 2020 in their Satisfaction Index scores. Rogers leads the improvement with +71 points, followed by Bell at +34 and TELUS at +23; overall industry improved by 40 points.

Based on available churn results from each carrier, it is possible to determine the extent of correlation between this improvement in customer satisfaction scores and churn. Table 1.4 shows the equivalent churn rate for each carrier and the industry (for years available).

Table 1.4 Annual Monthly Average Churn

Carrier	2015	2016	2017	2018	2019	2020
TELUS[a]	0.89%	0.89%	0.82%	0.81%	n/r	n/r
Rogers[a]	1.27%	1.22%	1.20%	1.10%	1.10%	n/r
Bell[a]	1.24%	1.21%	1.15%	1.13%	1.08%	n/r
Top 3 Blended[b] (premium + flanker)	1.46%	1.41%	1.35%	1.34%	n/r	n/r

Industry Average[b]	1.54%	1.51%	1.47%	1.44%	n/r	n/r

[a]Company annual financial reports

[b]CRTC Communications Monitoring Report (2019)

n/r = results not reported

Each of the carriers, and the industry overall, recorded lower churn over the same years that the customer satisfaction index scores improved. The correlation statistics for each carrier are as follows:

TELUS (2015-2018):	-.17
Rogers (2015-2019):	-.96
Bell (2015-2019):	-.90
Top 3 Blended (2015-2018):	-.97
Industry Blended (2015-2018):	-.95

There is a high, inverse correlation between improved customer satisfaction index scores and churn. TELUS is an anomaly, with a weak inverse correlation of -.17; this may be due to the already higher satisfaction scores and lower churn results between TELUS and the rest industry at the start of the data in 2015, possibly indicating reduced opportunity to make churn improvements or customer satisfaction gains.

Despite the strong correlations exhibited between the customer satisfaction index scores and annual churn rates, it cannot be stated from this data that the relationship between these two metrics is *causal*. There may be numerous other factors that might contribute to lower churn, some of which are also measured by J.D. Power (specifically, the other three reports mentioned previously covering aspects such as network quality and purchase experience). Other factors, both endogenous and exogenous, may also play a positive or negative role in churn results.

The strong correlations determined above, along with a common business intuition that “satisfied customers are more loyal”, has led to wireless companies continuing to invest in improvements to drive customer satisfaction in the belief (absent data to the contrary) that this intuition is correct.

1.2 Statement of Problem

Given the significant financial resources expended by the carriers to acquire and retain customers, it might be asked if the focus on improving customer satisfaction is an effective means at improving loyalty and reducing customer defections.

In addition to the COR costs that are directly identified through financial statements, there are other operational costs incurred in the drive to deliver superior customer satisfaction. Examples of these costs, while not enumerated in the annual reports, would include costs to:

- Staff call centres to deliver customer support
- Hire and train call centre and retail store staff to provide customer service
- Develop and maintain digital support (website and apps)
- Improve network quality and reliability

While there is a large body of research and industry experience (explored further in Chapter 2) to suggest that efforts to improve customer loyalty via customer satisfaction are not wrong, the industry and academia have not provided substantive proof that such efforts are causal. Practitioners in the field make business decisions to a large extent on accepted common sense (e.g. “Of course, if customers are more satisfied, they will be more loyal”) or based on strongly correlated decreases in churn that occur when satisfaction increases. However, are these decisions supported by actual customer behaviour?

1.3 Aim and Scope of Study

This study aims to determine i) if a causal linkage between customer satisfaction and loyalty (churn) can be found, and ii) if there exists an ability to influence customer satisfaction and loyalty through specific, practical and scalable treatments.

The study places its emphasis on the primary (premium) post-paid segment of the Canadian wireless market, which as outlined in Section 1.2 comprises the majority of the wireless telecommunications market in Canada (both subscribers and revenue).

1.4 Importance of Study

As explored in Section 1.2, the financial benefits to a mobile services organization to be able to influence customer satisfaction and loyalty, to a degree known to be causal, is of substantial benefit and competitive advantage.

The outcome of the study may yield information and insights that company management can utilize to increase spending in certain areas or programs that might now be known to influence loyalty or decrease spending in those areas where the evidence suggests that there is no causal outcome to be achieved.

1.5 Overview of Study

The study begins in Chapter 2 with a review of previous research and findings in the areas of customer satisfaction and customer loyalty, with particular focus on the methods used (or not used) in this prior research. The opportunity to make a unique contribution through the use of a robust experimental method is identified, and the research hypotheses and specific study methods are then documented in Chapter 3. The results of the experiment are reviewed in Chapter 4, confirming the soundness of the experiment implementation and the outcomes of the experiment itself. The hypotheses proposed in the previous chapter are assessed.

Chapter 5 presents an in-depth discussion of the findings, the implications for practice, and recommendations for further research. The study closes in Chapter 6 with conclusions.

2 Literature review

2.1 Introduction

The aim of the study, as summarized in Section 1.3, is to determine if a causal linkage between customer satisfaction and loyalty exists within mobile phone users, and if there is such a linkage, what practicable and scalable treatments might be employed to increase customer loyalty and improve the financial performance of the organization.

The study encompasses many areas examined by prior research on the relationships between firms and its customers; more specifically, this research includes the examination of customer satisfaction and customer loyalty within Business-to-Consumer (B2C) product and service industries.

The literature reviewed in this chapter seeks to provide an understanding of how previous research has come to define what customer satisfaction is, how it is measured, and what factors or antecedents influence a given customer's level of satisfaction with a company's goods or services. Similarly, prior research is examined to understand the concept of customer loyalty, how it is measured, and how it is influenced both by customer satisfaction and other factors. In both cases, identification will be made where gaps, debates, or contradictions in the existing literature are observed.

Particular attention is paid to understanding the antecedents of satisfaction and loyalty, with a focus on identifying those that might be suitable as treatments in an experimental design to resolve the key aim of this study – the examination of causal relationships between satisfaction and loyalty.

Wherever possible, research conducted in the same industry (telecommunications) as the study is used. Due to limited satisfaction and loyalty research in that specific field, the review is augmented with research covering other consumer product and service industries.

The chapter is organized in four main sections:

- Customer Satisfaction – review of prior research in the consumer products and services markets into the concept of satisfaction, what it is, how it is formed, how it is measured, and why companies focus substantial efforts on improving it.
    - Included in this section is an examination of the antecedents of customer satisfaction, with a focus on antecedents incorporated into this study.
- Customer Loyalty – similar to that of Customer Satisfaction, this section will review research that has been conducted to understand the concept of loyalty.

- Satisfaction:Loyalty Relationship – a detailed examination of prior research examining the role of satisfaction as an antecedent to loyalty.
- Methodologies Employed – an examination of prior research studies in the subject area, and how various researchers have set out to examine the nature of customer satisfaction and loyalty.

Across each of the sections, attention will be paid to the methodologies employed in prior research (especially studies utilizing experimental methodologies) as well as the current debates and gaps in the current research.

2.1.1 Key Definition – "Customer" or "Consumer"

Literature examining the concepts of satisfaction and loyalty appear to use the adjectives "customer" and "consumer" relatively interchangeably. In her review of literature defining satisfaction, Cotiu (2013, p.70) found that "Out of … 37 articles, only 2 expressly referred to the person feeling satisfaction as consumer, 17 used a mix of the two terms, and 18 referred to customers only".

The difference between the two may seem trivial, but the distinction is important in the study of mobile phone users. Across the literature, the most commonly accepted definitions found by Cotiu are:

- "Customers" are persons that have purchased a particular good or service, but may or may not actually use it;
- "Consumers" are always the person using the product or service, regardless of who may have paid for it

In the telecommunications industry, users of a company's mobile phone services are a mixture of these two. For example, a person may purchase and pay for mobile phone service for themselves and one or more other individuals (for example, a parent who pays for a mobile phone for not only for themselves, but also for their spouse and one or more children). In this case, the parent would be both a customer and a consumer, whilst the spouse and children would be considered consumers only.

This study focuses on customers within the mobile phone industry, as the primary owners of the financial relationship with the companies providing mobile phone service. As such, the literature reviewed in this chapter will focus on research examining the concepts of satisfaction and loyalty as they pertain to customers.

2.2 Customer Satisfaction

> "I know it when I see it"
>
> – Justice Stewart, United States Supreme Court, Jacobellis v. Ohio, 1964

Customer satisfaction might seem to be a concept that each individual person "knows when they see it", but researchers have sought for decades to define exactly what customer satisfaction is, while service and product organizations have sought to outdo each other as the best providers of customer satisfaction in their field[8].

But what is meant by the term customer satisfaction?

Many definitions appear in the literature. In an attempt to address the various (and sometimes conflicting) definitions of satisfaction, Giese et al. (2000) undertook a review of the definitions of satisfaction across 30 years of satisfaction literature. Combining the results of this review with data gathered through personal and group interviews (to ensure alignment with the views of consumers themselves), the authors concluded that satisfaction was comprised of the following components:

1. An affective (i.e. emotional) response, that varies in intensity;
2. A satisfaction focus around product choice, expectations, purchase or consumption; and
3. A response at a particular point in time, such as after product use or as a cumulative measure over time.

Each of these three components is explored further in following sub-sections.

2.2.1 Affective Response

The affective or emotional response of a customer to a product or service received is most frequently described (Cotiu 2013) using the "disconfirmation of expectations" or "expectancy-confirmation / disconfirmation" paradigm, introduced by Oliver (1980). Research examining the role of expectation-disconfirmation dates back to at least the early 1970's, where Anderson et al. (1973) found that disconfirmation of product expectations negatively impacts a customer's perception of a product, as well as negatively impacting satisfaction.

An analysis of 373 articles published by the Journal of Customer Satisfaction, Dissatisfaction, and Complaining Behaviour (JCSD&CB) from 1998 to 2014 found that the leading category

[8] Refer to Sections 1.1.5 and 1.1.6 for a discussion of the competitive focus on customer satisfaction within the Canadian mobile telecommunications market.

of articles was the topic of customer satisfaction antecedents, accounting for 25% of all articles (Dahl et al. 2015). Furthermore, this same analysis found expectation-disconfirmation to be the most common variable explored in those articles as an antecedent to customer satisfaction.

This paradigm defines customer satisfaction as an outcome of the customer's evaluation that a given product or service has met his or her expectations (Oliver et al. 1989), and was further extended after observing that satisfaction was not purely outcome dependent; that is, high product performance levels in and of themselves did not necessarily generate satisfaction and likewise low performance levels did not always generate dissatisfaction. Oliver et al. (1985) proposed that an intervening construct – expectation-disconfirmation – was the primary integrating cognition capturing the differences in expectations and actual performance that any individual customer might experience. As an example, a customer with low expectations for the performance of the product, who then receives a level of performance equal to or better than his or her expectations, would experience positive disconfirmation. Another customer receiving the same level of performance, but who had expected a greater level, would experience negative disconfirmation.

Several other authors also support the notion that expectation-disconfirmation plays a role in determining customer satisfaction, however there are variations between them on how this construct is formed. Two of the approaches most commonly cited by other researchers[9] are summarized in Figure 1.

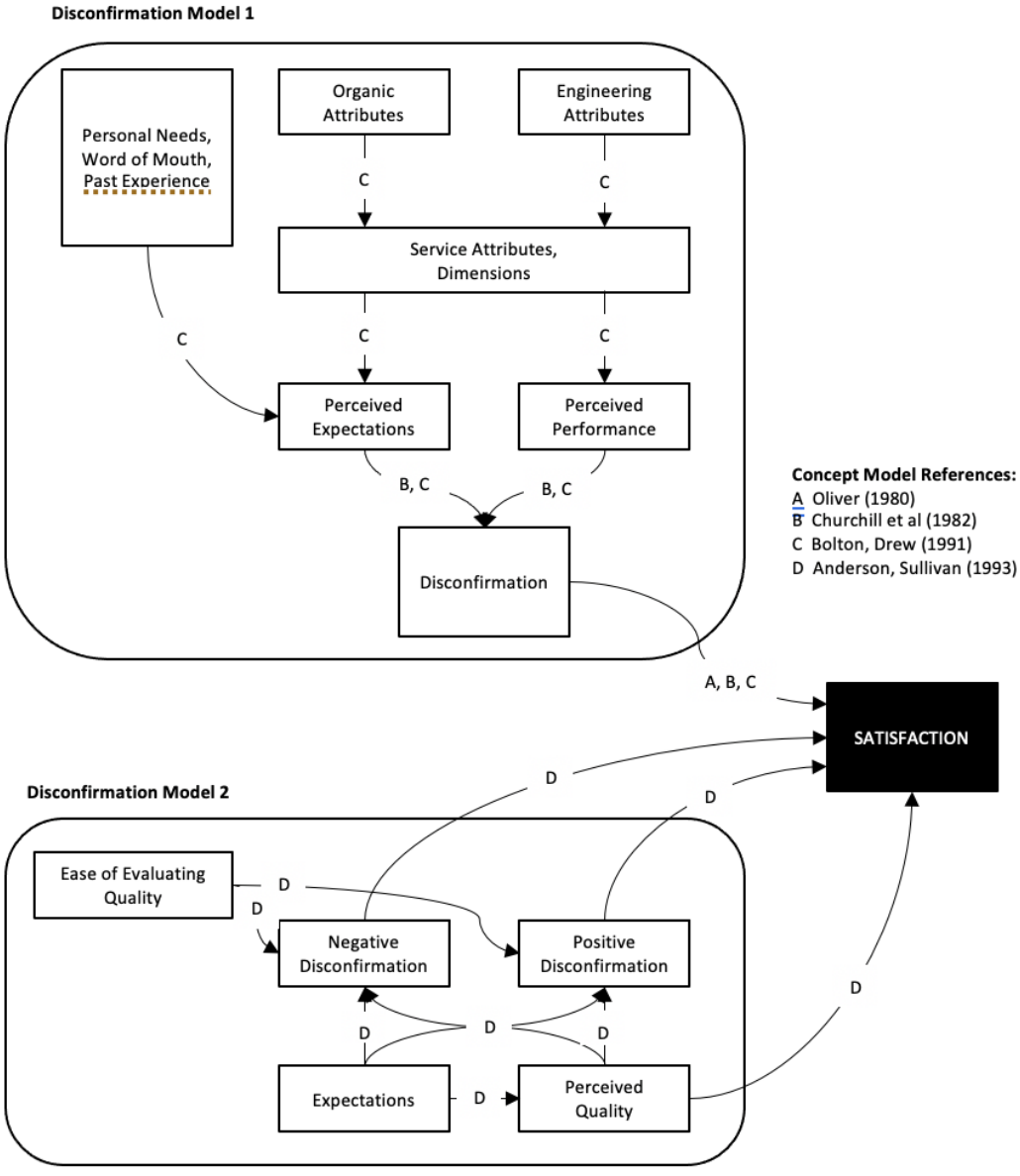
Disconfirmation Model 1 illustrates two studies (Churchill Jr et al. 1982, Bolton et al. 1991) that propose that disconfirmation is an outcome of a customer's perceived expectations versus the perceived performance received, similar to Oliver (1980). The expectation-disconfirmation model is expanded by Bolton et al. (1991) by examining the antecedents (such as personal needs, word-of-mouth, past experiences and organization attributes, among others) of the perceptions that form disconfirmation.

In the empirical research conducted using Disconfirmation Model 1, both sets of authors found support for the role of disconfirmation impacting customer satisfaction. Using an experimental methodology, Churchill Jr et al. (1982, p. 502) found (in a field study involving non-durable products) that "Disconfirmation positively affected satisfaction as is commonly held; when subjects perceived the product performing better than expected, they were more satisfied with it and vice versa".

[9] Collectively, the four studies referenced in Figure 1 have been cited 39,868 times as of 02-May-21. Source: Google Scholar

In the second variation of the model, shown as Disconfirmation Model 2, Anderson et al. (1993) propose that disconfirmation is not a single variable, but two variables: positive disconfirmation and negative disconfirmation, both of which exist simultaneously. The justification put forward for the establishment of the two variables (as opposed to one variable on a continuum) is that there exists an asymmetric quality within the satisfaction framework, hypothesizing that "Satisfaction is more sensitive to negative disconfirmation than positive disconfirmation" (p. 130). The empirical findings support their hypothesis, concluding that satisfaction was, in fact, more sensitive to negative disconfirmation than positive disconfirmation for nearly 80% of the cases in their study.

Figure 1 Disconfirmation Models



Source: Developed for study

2.2.2 Satisfaction Focus

In any relationship between a customer and an organization providing a product or service, there may be many aspects of that relationship subject to an assessment of satisfaction by the customer. The satisfaction focus identifies that part of the relationship under assessment, whether something very specific or more general. When studying customer satisfaction, it is important to determine the focus that is appropriate given the context of the study (Giese et al. 2000).

2.2.3 Timing of Satisfaction

Satisfaction is an outcome that occurs after (post) some action or event experienced by the customer. In her review of customer satisfaction literature, Cotiu (2013) concluded that the timing of satisfaction may result post-consumption (citing Anderson 1994, Giese et al. 2000, Dubrovski 2001, Hung et al. 2007), post product or service choice (citing Zeithaml et al. 1993), post decision (citing Churchill Jr et al. 1982) or at other points in time.

The combination of the dimensions of focus and time yields two general concepts of satisfaction: one that is a transaction-specific approach and on that is a more general, cumulative approach (Bae 2012).

Transaction-specific satisfaction attempts to confine the measure of satisfaction to a single, specific activity at a point in time. Examples of activities that are often measured for such satisfaction include a single interaction with a sales associate or customer service representative, the completion of a transaction on the company's website, or many other discrete activities where a customer can be reasonably prompted to assess their level of satisfaction with that activity.

The cumulative assessment of satisfaction by a customer is defined by a group of activities, transactions or experiences that a customer has with a firm over a duration of time (Johnson et al. 1991, Fornell et al. 1996, Oliver 2010), and is frequently used as the definition of satisfaction when examining the relationship between customer satisfaction and customer loyalty (Bae 2012).

2.2.4 Customer Satisfaction versus Service Quality

Similar to customer satisfaction, service quality is an outcome of an evaluation process by a customer, but focused on the service received (and whether it met the customer's

expectations or not) rather than whether or not the customer is satisfied (Grönroos 1984). Service quality is closely associated with customer satisfaction, with several studies establishing positive relationships between the two measures (Dabholkar et al. 1994, Bansal et al. 2015) and with that of customer loyalty (Parasuraman et al. 1988, Anderson et al. 1993). Despite this close association, it has been recommended that customer satisfaction be viewed and measured separately from service quality, and that researchers make clear which is the focus of their studies, and why (Dabbolkar 1995).

2.2.5 Antecedents of Customer Satisfaction

The specific attributes and antecedents of customer satisfaction have been studied repeatedly by many scholars, across a wide variety of consumer goods and services industries (Patterson et al. 1997). In any given study, the researcher chooses the relevant antecedents to be investigated, based on the particular area of research, the objectives of that research and the industry within which the research is to be conducted (Cotiu 2013).

These antecedents form the basis for multiple points of comparison by the customer within the expectation-disconfirmation framework discussed previously (Bolton et al. 1999).

In order to support the aim of this study and the intended research methodology, it is necessary to identify one or more customer satisfaction antecedents that could be utilized in the proposed experiment. Criteria for the selection of the appropriate antecedents to be used includes:

- Antecedents that are relevant to the industry under study (mobile communications), and are relevant to the subject organization;
- Ability to design and execute a treatment that might be expected to influence the antecedent and can be deployed at the desired scale of the experiment in a real-world natural field setting; and
- Treatments that, if successful in demonstrating the ability to causally influence customer loyalty, can be scaled for deployment across the entire customer base of the organization.

A review of current research into the antecedent of customer satisfaction follows, with a detailed discussion of the specific antecedents intended for inclusion in the study's experiment.

One of the few compilations of the wide variety of customer satisfaction antecedents studied in the literature was completed by Thusyanthy et al. (2017), where they identified 25 unique antecedents across 36 studies. The antecedents found by these authors are shown in Table 2.1. While not all antecedents may be applicable in determining customer satisfaction in all

situations where goods and services are provided, the model developed by Thusyanthy et al. (shown in Figure 2) illustrates the complexity the underlies the construct of customer satisfaction.

In addition to the antecedents shown in Table 2.1, other authors have proposed antecedents of customer satisfaction within the mobile communications market, including customer service, billing systems, network quality and pricing plans (Lim et al. 2006).

Three antecedents were selected from the extant literature on customer satisfaction to be utilized in this study:

- Perceived Value
- Call Centre Service Quality
- Advertising Pressure

All three of these antecedents are areas of concern and focus by the subject organization as they relate to how they impact customer satisfaction and, ultimately, the retention of customers. For example, Perceived Value is measured on a quarterly basis as part of the subject organization's customer satisfaction surveys. Dozens of metrics associated with delivering a quality experience for customers calling into the call centre are monitored on a real-time basis and reviewed by management daily. And finally, there is a growing concern regarding the impact of an increasing volume of direct marketing activities aimed at customers, via email, text messaging, internet advertising etc., and whether such marketing activity is negatively impacting customers.

An examination of prior research for each of these three antecedents is covered in the next three sections of this chapter.

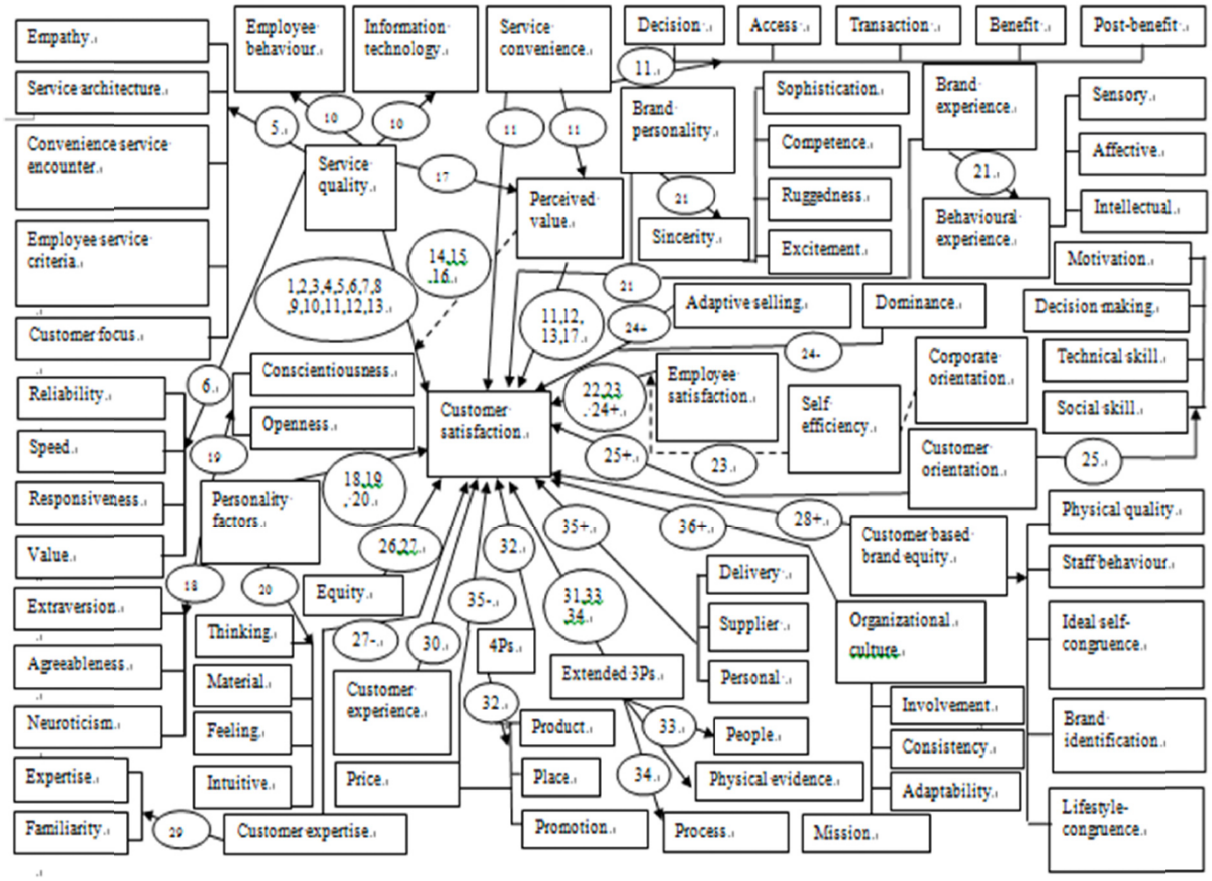
Table 2.1 Antecedents of Customer Satisfaction

Product	Delivery Performance	Customer Orientation
Place	Supplier Know-how	Organizational Culture
Promotion	Personal Interaction	Service Convenience
People	Brand Personality	Perceived Value
Process	Brand Experience	Personality Factors
Physical Evidence	Adaptive Selling	Equity
Employee Interaction	Dominance	Customer Brand Equity

Customer's Experience Employee Satisfaction Customer Expertise

Service Quality

Figure 2 Comprehensive Model of Antecedents of Customer Satisfaction



Source: Thusyanthy et al. (2017)

2.2.5.1 Perceived Value Antecedent

Perceived value has been identified by many researchers as "…the most important and most comprehensive antecedent of customer satisfaction" (Gallarza et al. 2016, p. 167), a position reinforced by El-Adly (2019, p.2), who summarizes research by Cronin et al. (2000), Yang et al. (2004), and Gallarza et al. (2006) establishing perceived value as "…one of the most important means of generating customer satisfaction and loyalty".

Bolton et al. (1991) suggest that perceived value seems "…to be a 'richer', more comprehensive measure of customers' overall evaluation of a service than service quality", while Grönroos et al. (2013) suggest that value is a more satisfactory determinant of how consumers assess a product or service than price alone.

El-Adly (2019) documents over 50 prior studies that examined perceived value across a variety of service industries, including telecommunications, hospitality, retailing and others.

In addition to the role of perceived value as an influence on customer satisfaction, several studies have also found a direct influence between perceived value and customer loyalty itself (Tam 2004, Yang et al. 2004, Wu 2014).

Despite the fairly large body of research on perceived value and its relationship to customer satisfaction, there appears to be few studies examining the role of perceived value on customer loyalty behaviour (Cronin et al. 1997), which the lead author (Cronin 2016) continues to suggest - almost 20 years later - that research is still needed to link customer behaviour with customer perceptions of value (among other attributes).

The definition of perceived value has been a subject of discussion for many researchers, with a large proportion of researchers aligning with the definition maintained by Zeithaml over the past several decades (Zeithaml 1988, Zeithaml et al. 2018). The overall definition that she provides is that "…perceived value is the consumer's overall assessment of the utility of a service based on perceptions of what is received and what is given" (Zeithaml et al. 2018, p. 459). The nature of what is received and what is given varies by product or service category and can even vary between customers of the same product or service. In the end, perceived value is proposed as the net outcome of the trade-offs between these "give" and "get" components of the relationship with a company's product or service.

This ratio of perceived costs and perceived benefits is determined to be something that each consumer determines for his/herself (McDougall et al. 2000). The most frequent method of measuring perceived value is by obtaining the consumer's evaluation of value through the use of survey questions. Examples of survey questions observed in prior research is summarized in Table 2.2. Consistent across the examples provided is the concept of the survey respondent evaluating value based on some aspect of what they have had to give (most often "price" or "cost", but also including "time, energy and effort" or unspecified "sacrifices") along with that they get ("quality", "satisfaction of needs and desires").

The importance ascribed by researchers to perceived value as an antecedent to customer satisfaction, combined with the pricing competitiveness of the mobility market described in Chapter 1, supports the inclusion of this factor in the experiment planned for this study.

Table 2.2 Perceived Value Survey Questions

Authors	Questions Used	Product/Service Context	Measurement Construct
Patterson et al. (1997)	Considering the fee paid and what the consultant delivered, overall I believe we received fair value for money	Consulting services	Seven-point, Likert-type scale
McDougall et al. (2000)	The [provider] offered good value for money	Consumer services (dentist, auto service, restaurant, hair stylist)	Seven-point, Likert-type scale
Wang et al. (2004)	Overall, the chosen offerings are value for money The chosen offerings are worth what is given up such as time, energy and effort Comparing with major competitors, the transaction with the provider is a good choice	Telecommunications	Seven-point, Likert-type scale
Yang et al. (2004)	Compared to alternative companies, the company offers attractive product/service costs. Compared to alternative companies, the company charges me fairly for similar products/services. Compared to alternative companies, the company provides more free services. Comparing what I pay to what I might get from other competitive companies, I think the company provided me with good value. Comparing what I pay to what I might get from other competitive companies, I think the company provides me with good value.	Online banking	Five-point, Likert-type scale

Authors	Questions Used	Product/Service Context	Measurement Construct
Kuo et al. (2009)	I feel I am getting good mobile value-added services for a reasonable price	Telecommunications	Five-point, Likert-type scale
	Using the value-added services provided by this telecom company is worth for me to sacrifice some time and efforts		
	Compared with other telecom companies, it is wise to choose this telecom company		
Lai et al. (2009)	Overall, the service I receive from company X is valuable	Telecommunications	Seven-point, Likert-type scale
	The service quality I receive from company X is worth my time, energy, and efforts		
Edward et al. (2010)	Comparing what I pay to the service I receive, I think my service provider provides me good value.	Telecommunications	Seven-point, Likert-type scale
	This provider's service is a better value for money.		
	The provider charges a reasonable price for the service provided.		
Gallarza et al. (2016)	Reasonable price	Retail stores (grocery, apparel, electronics, home furnishings)	Seven-point, Likert-type scale
	Value for money		
	Worth paying for these products		
	Economical products		
Gallarza et al. (2016)	In general, the value obtained from this tourism experience is high	Tourism	Five-point, Likert-type scale
	Compared to what I have had to sacrifice, the capacity of this tourism experience to satisfy my desires and needs has been high		
	Comparing benefits to sacrifices, I consider this tourism experience to be a good one		

Authors	Questions Used	Product/Service Context	Measurement Construct
Roberts-Lombard et al. (2018)	My cell phone network service provider offers me value for money.	Telecommunications	Seven-point, Likert-type scale
	I consider the rates of my cell phone network service provider as reasonable.		
	Doing business with my cell phone network service provider is a right decision when price and other costs are considered.		
	Doing business with my cell phone network service provider is a right decision when the overall quality of service delivery is considered.		
	Compared to the quality of cellular service reception that I get, I pay a reasonable price.		
Dey et al. (2020)	I get good customer service from my current mobile network provider	Telecommunications	Five-point, Likert-type scale
	My current mobile network provider keeps me informed about things that matter to me such as my bill, out of bundle charges, etc.		
Menidjel et al. (2021)	The product of the store is good value for money	Clothing store	Seven-point, Likert-type scale
	The price of the store is acceptable		
	The product of the store is a good buy		

Note: Descriptors used in the 5- and 7-point Likert-type scales utilized answers ranging from “strongly disagree”, “somewhat disagree”, “slightly disagree”, “neutral”, “slightly agree”, “somewhat agree” and “strongly agree”, or similar variations based on the scale.

2.2.5.2 Call Centre Service Quality Antecedent

The quality of service provided by a call centre as an antecedent to customer satisfaction can be seen as a subset of a larger collection of attributes frequently referred to as “service quality”. One popular model, SERVQUAL, was developed to assess customer perceptions in service and retailing organizations. Through a series of refinements, SERVQUAL ultimately came to define five dimensions of service quality: Responsiveness, Reliability, Tangibles, Assurance and Empathy (Parasuraman et al. 1985, Parasuraman et al. 1988, Parasuraman et al. 1991). Within this model, the dimension of Responsiveness represented, among other things, a customer’s perception of whether or not the “…waiting time to receive service is not extensive” (Parasuraman et al. 1985, p. 47).

While there appears to be some consensus that the level of satisfaction a customer experiences when calling in to a company’s call centre can be a contributor to overall customer satisfaction, consensus on what factors are important in achieving that satisfaction is lacking (Feinberg et al. 2002).

The operational measures employed by call centre managers have become fairly standardized across the industry and are the focus of performance management by these practitioners. A common list of 13 critical measures is shown in Table 2.3 (Anton 1997, as cited in Feinberg et al. 2000). In their study of data from 514 call centres across 15 industries (including 30 call centres for telecommunications companies), Feinberg et al. (2000) found that only seven of these measures were significantly related to customer satisfaction (measures numbered 1 through 7 in the table) and only two (measures 3 and 4, First Call Resolution and Abandonment Rate) were identified as causally related to customer satisfaction.

In a similar study focused on 138 call centres for financial services and banking clients, none of the 13 measures were found to be significantly related to caller satisfaction (Feinberg et al. 2002).

Other researchers, however, have asserted that their empirical results do support a direct linkage between waiting (or queue) time and call satisfaction (Durrande-Moreau 1999, Whiting et al. 2006). Whiting et al. (2009) write that:

> Within the few studies that have been conducted on caller satisfaction, there is one important variable that has been shown to influence callers and that variable is waiting time. This waiting on hold experience has been shown to directly impact satisfaction. (p. 279)

Measures of wait time within the call centre experience must take into account both the “real” wait time (the objective, actual wait time the customer was on hold) and the “perceived” wait time (the subjective perception of how long the customer thinks they waited on hold). Both are shown to influence caller satisfaction (Whiting et al. 2009). The real wait time can be measured and reported by the systems that manage the calls within the call centre environment; perceived wait time is measured by asking callers questions via surveys. These questions are often embedded as part of a broader question related to the ease a customer experiences in accessing a given company. For example, Dey et al. (2020) used the question “My current mobile service provider is easily accessible through a variety of channels such as shops, call centres, their website, etc.” as one of three indicators in their operationalization of a customer service satisfaction construct.

The incorporation of call centre service quality as a factor in this study’s experiment may contribute additional evidence in the arguments for or against the role of call centre satisfaction influencing overall customer satisfaction.

Table 2.3 Critical Call Centre Measures

Measure	Description
1. Average Speed of Answer (ASA)	Average amount of time it takes for calls to be answered in a specific time period
2. Queue Time	Amount of time a given caller is in the line for an answer
3. First Call Resolution	Percentage of callers who have satisfactory resolution on the first call
4. Abandonment Rate	Percentage of callers who hang up prior to call being answered
5. Average Work Time	Time need to finish paper work after call has been completed
6. Percentage Calls Blocked	Percentage of callers that receive a busy signal and could not get into the queue
7. Service Levels	Calls answered within a defined period of time (measured in seconds or minutes) divided by total number of calls handled
8. Average Handle Time (AHT)	Total time caller was connected to an agent
9. Adherence	Percentage of time agents are in seats taking calls as scheduled
10. Time Before Abandonment	Average time caller waited in queue before giving up
11. Inbound Calls per Shift	Number of calls per agent per shift
12. Agent Turnover	Number of call centre agents who quit in a period of time
13. Total Calls	Total calls received in period of time

Source: (Anton 1997, as cited in Feinberg et al. 2000)

2.2.5.3 Advertising Pressure Antecedent

The role of the frequency of marketing communications (e.g. emails, text messages, or phone calls) as an antecedent to customer satisfaction and customer loyalty has been studied in order to determine if there is a level of these communications that result in a sub-optimal outcome (Dreze et al. 2003). The desired outcome of such communications is to increase sales and revenues from consumers that are already customers of an organization, or potential customers that have provided their email or phone number to a company along with permission to be sent marketing communications.

Negative outcomes impacting the ability to generate future sales and revenues can arise through the over-use of this contact information, including deletion of communications without reading them or withdrawal of permission to continue sending communications (referred to as “opting out”). Additionally, customers may also develop negative attitudes towards the company, its brand, and its products (Micheaux 2011).

The concept of advertising pressure was defined by Micheaux (2011, p.46) as “…a state of irritation provoked by the impression of receiving too many e-mail advertisements from a commercial source”, which leaves customers feeling “…inundated with e-mail volume” and “under pressure, in that they are forced to deal with this incoming information”.

Studies of the impact of marketing communication frequency and the associated feeling of advertising pressure have yielded conflicting results. In their experiment sending 2, 4, or 6 emails over an 8-week period, Micheaux (2011) found that while the frequency of emails received negatively impacted the rate at which the emails were opened, it was the content of the emails themselves that had a greater impact on the number of study participants that opted-out of future emails.

In contrast, Dreze et al. (2003) conducted an experiment that sent five million emails over a period of 13 months and found that study participants that received emails more frequently than every nine weeks had increased opt-out rates. The authors advise managers to “…err on the side of longer rather than shorter inter-communication times”.

Customers that receive too many communications may express irritation with the frequency of the messages received from a company, but these sentiments have previously been found to have no impact on the subsequent behaviour of the customer (van Diepen et al. 2009). Further investigation may yield insights into the impact of marketing frequency on customer satisfaction and loyalty.

2.3 Customer Loyalty

The importance of customer loyalty within the mobile services market was explored in Section 1.1, which highlighted that each year Canadian telecommunications providers must acquire approximately 3.5 customers for every one net new customer they add to their subscriber base (in order to offset the 2.5 customers they lose to competitors in the same timeframe). Additionally, mobile service providers spend nearly 15% of their revenue in direct costs to retain customers in an industry where reducing the monthly churn of customers by a single basis point across a given company's subscriber base can result in an additional $66.8 million in incremental lifetime revenue from those customers.

The financial implications of customer loyalty for the mobile phone companies reinforces the findings of researchers that have concluded that "Customer loyalty is the prime determinant of long-term financial performance of firms" and that this is "…particularly true for service firms where increased loyalty can substantially increase profits." (McDougall et al. 2000, p. 392).

Customer loyalty has been linked to positive performance in the product or service marketplace (Fornell 1992, Anderson et al. 2000), strong financial performance (Anderson et al. 1994, Rust et al. 2004, Gupta et al. 2006) and the creation of shareholder wealth as loyalty reduces acquisition and retention costs of customers (Bae 2012).

There are multiple definitions of what constitutes customer loyalty, with the definition by Oliver (2010) used in one form or another by many researchers, namely that customer loyalty is "…a deeply held commitment to rebuy or re-patronize a preferred product or service consistently in the future, despite situational influences and marketing efforts having the potential to cause switching behavior" (p. 432). Zeithaml et al. (1996) adds that in addition to repeat purchasing, a loyal customer is also willing to recommend the product or service to others.

Loyalty itself is understood to be comprised of at least two distinct dimensions: attitudinal loyalty and behavioural loyalty (Day 1969, Jacoby et al. 1973, and Yi 1991, as cited in Bae 2012). These two dimensions can be expanded further into four ascending stages of loyalty, as proposed by Oliver (2010):

1. Cognitive Loyalty – customers are loyal to a company's brand based on information they have about that brand, and is based on brand belief alone
2. Affective Loyalty – loyalty that has evolved, based on cumulative satisfaction with the product or service, into a greater commitment than in the first stage

3. Conative Loyalty – customers have the intention and are motivated to re-buy the product or service
4. Action Loyalty – where the customer puts their intentions into actions, completing the re-purchase action with a company, while simultaneously rejecting alternatives and other obstacles

The evolution from attitudinal loyalty to behavioural loyalty takes place along the continuum of these stages, evolving from a desire to continue purchasing products or services from a given company, to the actual act of re-purchase (Czepiel et al. 1987 and Neal 1999, as cited in Yang et al. 2004).

The simultaneous existence of both attitudinal and behavioural loyalty in a customer is an important concept: the existence of an intention to re-purchase alone may not lead to the actual behaviour of re-buying, and the act of re-buying alone may not reflect an intention to remain loyal (Yang et al. 2004). Customers in the latter situation exhibit what is referred to as "spurious loyalty" and may switch to another company if provided with a more attractive alternative; attitudinal loyalty contributes to ensuring a greater chance of repurchase and resistance to efforts to convince the customer to switch to another company. The intersection of attitudinal and behaviour loyalties can be expressed as four levels, as shown in Figure 3.

In this approach, Loyalty is the most desired outcome representing the highest degree of intent to remain loyal, supported by the behaviour to demonstrate this loyalty in practice; customers that exhibit No Loyalty are at greatest risk to switch to another company's products or services (Dick et al. 1994).

Figure 3 Attitudinal-Behavioural Loyalty Relationship

		Repurchase Behaviour	
		High	Low
Relative Attitude	High	**Loyalty**	**Latent Loyalty**
	Low	**Spurious Loyalty**	**No Loyalty**

Note: Adapted from (Dick et al. 1994)

While behavioural loyalty is the source of the financial benefits ascribed to loyalty, it has been found to be something that is "…difficult to observe and is often equally difficult to measure" (Yang et al. 2004, p. 802). As a result, most researchers rely on measures of attitudinal loyalty (cognitive, affective and conative), where the customer indicates the

strength of their intention to repurchase a company's product or service (Yang et al. 2004, Bae 2012).

2.4 Linkage between Satisfaction and Loyalty

The importance of customer loyalty to firm performance, whether that firm is a mobile services provider or any other goods and services provider, underscores the value of research intended to understand the relationship between customer satisfaction and loyalty. A customer's level of satisfaction is viewed in the current literature as "…the primary driver of customer loyalty" (Bae 2012) and as the most important prerequisite to loyalty (El-Adly 2019). The positive relationship between these two factors has been established by many earlier researchers (Anderson 1996, Anderson et al 1994, Fornell 1992, Fornell et al 1996, Ping 1996, Rust and Zahorik 1993, Rust et al 1995 and Taylor and Baker 1994, as cited by Bae 2012).

Several field studies provide empirical evidence of this relationship (Oliver 1980, Bolton 1998). In a study of participants in a flu vaccination program, Oliver (1980) found that "satisfaction is the primary determinant" of post-usage attitude, and that satisfaction was a determinant of future purchase intentions.

In what is claimed to be the first research of its kind, Bolton (1998) examines the impact of cumulative satisfaction on the subsequent purchase behaviour of wireless customers over a 22-month period. The longitudinal approach to this study is important, as wireless is a recurring service model which requires ongoing payments for service on a monthly basis; the customer is, in essence, making a repurchase decision every 30 days[10]. The findings from this research indicated that there was a positive relationship between satisfaction and loyalty, a relationship that increased when higher degrees of customer satisfaction were present.

In Bolton's study satisfaction was measured in two waves, using a 5-point Likert-like scale (from Very Satisfied to Very Dissatisfied, when asked the question "Overall, how satisfied are you with the services received from [the company]"). Wave 1 measured satisfaction at the beginning of the study period, prior to any decision to stay or leave the company. Wave 2 measurement followed 6 months later for a randomly sampled group, and loyalty was a binary measurement, based on the existence of a deactivation date for the customer's account.

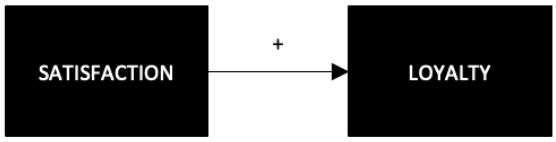
[10] Assuming the customer is not otherwise bound by a service contract

An additional study of telecommunications customers (Gustafsson et al. 2005) also examined the satisfaction-loyalty relationship and supported the findings by Bolton, stating "…customer satisfaction has a consistently negative effect on churn (a positive effect on retention)" (p. 216).

The argument for customer satisfaction as an important (if not the most important) antecedent to customer loyalty is supported by the industry results reviewed in Section 1.1.6, where improvements to customer satisfaction in the wireless industry over six years were shown to be highly correlated to reductions in the monthly churn of subscribers. Yet, there is evidence to suggest that satisfaction on its own is not sufficient to fully explain the existence or lack of loyalty in a given customer-firm relationship.

The research referenced thus far, if taken at face value, might be expressed as "A happy customer is a loyal customer", a simplistic relationship shown in Figure 4.

Figure 4 Satisfaction / Loyalty Relationship



The actual relationship is more complex. High levels of customer satisfaction do not always translate into high levels of loyalty (Oliver 1999), and there is often a lack of explanation in certain studies as to why satisfied customers leave while unsatisfied customers may in fact remain loyal (Jones et al. 1995, Bae 2012).

A variety of factors beyond satisfaction have been identified as playing a role in determining the degree of loyalty exhibited by a customer, several of which are shown in Figure 5. Some factors affect loyalty directly, while others are considered moderators of the relationship between satisfaction and loyalty.

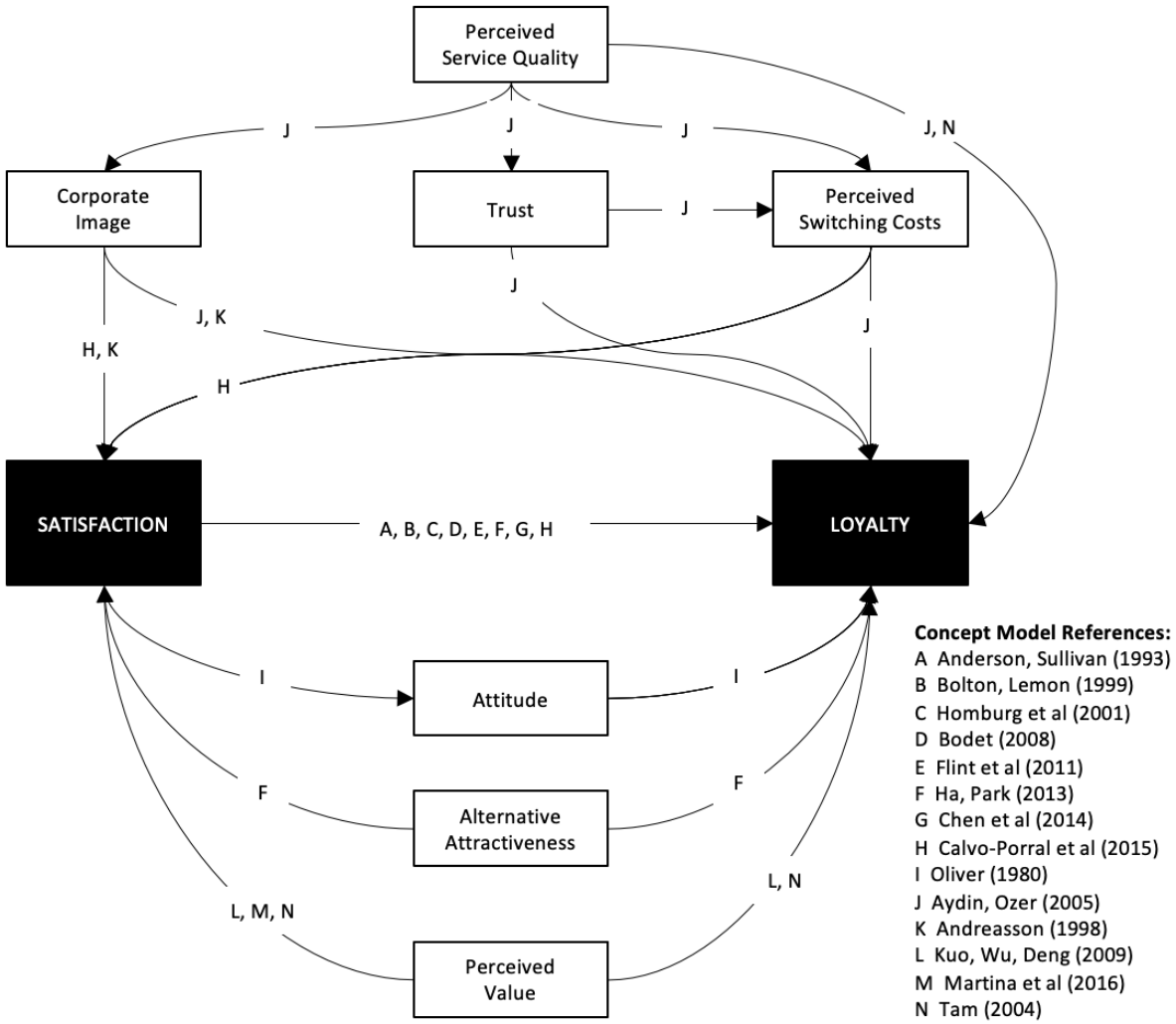
Despite the identification of additional factors beyond satisfaction that influence loyalty, it is clear that satisfaction plays a significant role in almost any model proposed, and perhaps is a "…necessary pre-condition for loyalty…" (Tam 2004, p. 899) that must be met before other factors can contribute their own influence (Dick et al 1994, Strauss et al 1997, as cited in Cotiu 2013). Unresolved, however, is whether satisfaction plays a mediating role between these additional factors and loyalty (McDougall et al 2000, Lin et al 2005, Gallarza et al 2006, Carpenter 2008, Chen et al 2008, as cited in Gallarza et al. 2016), whether these other factors (such as perceived value or perceived service quality) act on loyalty directly (Aydin et al. 2005) or both (Tam 2004, Kuo et al. 2009).

Support for factors other than satisfaction in the determination of loyalty can be seen in an analysis of the entire American Customer Satisfaction Index (ACSI) which found that the R^2 regression coefficient between customer satisfaction and customer loyalty metrics from that entire database was almost 37%, meaning that "…63% of the variance in customer loyalty is unexplained after accounting for customer satisfaction." (Bae 2012, p.2).

The important – but not singular – role that satisfaction plays in determining loyalty may help explain the impact that many factors play in determining both attitudinal and behavioural loyalty (as explored in Section 2.3). Following the model suggested by Dick et al. (1994), a high degree of satisfaction may be sufficient to create a high level of attitudinal loyalty but customers are still subject to factors that may encourage them to buy their goods or services from another provider (referred to as Latent Loyalty in Figure 3, where a customer may state they are satisfied, but are still subject to defection). These other factors can include those outside the incumbent firm's direct control, such as the attractiveness of a competitor's offerings (especially price) or lack of switching costs.

Conversely, a lack of satisfaction that results in low attitudinal loyalty - but is offset by a lack of other factors that might encourage defection such as switching barriers, a lack of a compelling competitive offer or customer apathy - can combine to create the situation of Spurious Loyalty. The risk for a firm whose customers exhibit spurious loyalty is that the influence of the non-satisfaction factors may change over time: switching costs may dissipate (for example, a customer with a service contract fulfils their obligation and is now free to choose another service provider) or the attractiveness of alternatives may increase (such as when a competitor places a product on promotion). The objective of achieving true Loyalty, where a customer's attitudinal and behavioural loyalty are both high, should remain the objective for practitioners.

Figure 5 Model of Selected Factors Influencing Loyalty



Source: Developed for study

2.5 Methods Employed Researching Satisfaction and Loyalty

In order to better understand how research has been conducted in the areas of customer satisfaction and customer loyalty, a detailed review of over thirty studies was conducted with an examination of the following attributes for each study:

- Key dependent variables and their definitions
- Key independent variables
- Methodological approaches employed
- Key findings

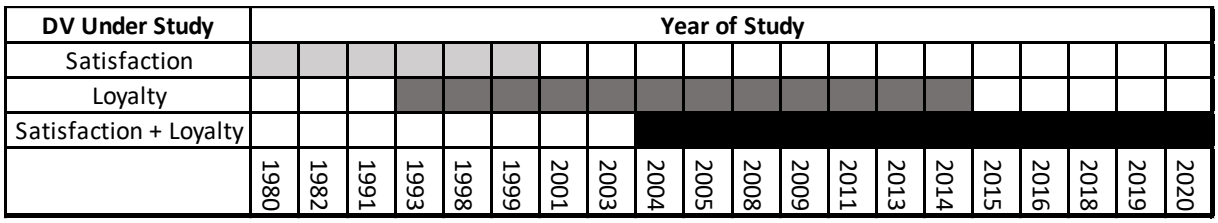
The 31 studies reviewed are not an exhaustive summary of all of the research conducted in this subject area but are a representation of the key authors in this discipline, the methodologies employed and the industries under study.

The purpose of this section is to identify what has been covered in the research to-date, where it leaves off, and where additional research is called for. Each of the points above is covered in the following sections with references to the appropriate studies covered by this detailed review; the complete list of studies and attributes is provided in Appendix A.

2.5.1 Key Dependent Variables and their Definitions

The primary focus of research in the area of customer satisfaction and loyalty has evolved over the past 40 years since the seminal work of Oliver (1980) on customer satisfaction. Research that followed Oliver continued to focus on that same subject area until the 1990's, when loyalty became more prevalent as the focus of researchers. From around 2004, researchers began to investigate customer satisfaction and loyalty jointly, a trend that is observed to continue today. Figure 6 illustrates this transition of research focus across the 31 studies reviewed.

Figure 6 Evolution of Research Focus Areas



This evolution appears to be a result of researchers attempting to link customer satisfaction to firm profitability (Hu et al. 2009, Cotiu 2013). Measures of loyalty appear to be able to be more directly linked to profitability (that is, continued patronage can be more quantitatively

assessed in terms of financial benefit), allowing researchers to therefore link customer satisfaction itself to firm profitability, through the influence on loyalty.

Of the 31 studies summarized in Appendix A, the majority of the studies (26) employed a DV based on the concept of loyalty. Additionally, 15 studies employed satisfaction as a DV (with our without loyalty as well) and 10 studies included various other variables as their DV's (for example, perceived value or service quality).

A summary of the 31 studies and their usage of satisfaction and loyalty as their dependent variables is shown in Table 2.4.

Table 2.4 Key DV's Used in Studies

DV's Used in Study	# Studies (of 31)	Years Represented	Example Authors
Satisfaction	5	1980-1999	Oliver (1980), Bolton et al. (1991), Anderson et al. (1993)
Loyalty	16	1993-2014	Rust et al. (1993), Bolton (1998), Gustafsson et al. (2005), Chen et al. (2014)
Satisfaction + Loyalty	10	2004-2020	Yang et al. (2004), Ha et al. (2013), Gallarza et al. (2016), Dey et al. (2020)

The operationalization of the satisfaction and loyalty dependent variables varies across the research, often based on the specific product or service industry under investigation.

Customer satisfaction is determined through survey questions, either self-administered (online or using paper surveys) or administered by the researcher (telephone or in-person). The satisfaction DV may be measured by a single, generic satisfaction question (for example, "Overall, I am satisfied with my current mobile network provider" as used by Dey et al. (2020)) or may be a multi-item construct combining several questions related to satisfaction (see Ha et al. (2013) and Calvo-Porral et al. (2015) for examples). Most frequently, the responses to the survey questions are measured using a Likert-type scale where the respondent may indicate the degree to which they agree or disagree with the provided statement. Less frequently employed are semantic differential scales, where the purpose is to gather a more subjective assessment of satisfaction (that is, how they may "feel" about their level of satisfaction, rather than a more objective assessment of whether or not they are satisfied). Of the 15 studies summarized in Appendix A, only two reference the

use of semantic scales for measuring satisfaction, the remainder use a Likert or Likert-like scale.

Measures of loyalty as the dependent variable are more varied than that of satisfaction and can be classified into two primary groups: i) those that measure the attitudinal loyalty or the behavioural intention to remain loyal, and ii) those that measure actual loyalty through action.

Of the two groups, measures of attitudinal loyalty and behavioural intention dominate the research literature. This is likely a result of the complexity involved in measuring actual loyalty, either through observation of the desired behaviour (such as customer re-purchasing the company's product or service again in the future) or as a self-reported action by the survey respondent. In both cases, the action that constitutes demonstration of loyalty may not take place in a frequent manner and would require follow-up over some period of time (perhaps years, in the case of infrequent purchases such as with an automobile). For most studies, this time requirement is a substantial practical obstacle, so researchers instead measure the customer's intention to remain loyal.

When measured as an attitude or intention to remain loyal, the loyalty measure is operationalized in a manner similar to that of satisfaction: using a single- or multi-item construct of survey questions, measured with Likert-type scales. Questions to derive attitudinal loyalty often take the form of measuring the likelihood of the customer to recommend the company to others and/or spread positive work of mouth (for example, Andreassen et al. (1998), Aydin et al. (2005), Bodet (2008), Ha et al. (2013)), while questions aimed at determining intention to remain loyal will focus on the customers' likelihood of continued repeat purchasing (see Homburg et al. (2001), Bodet (2008), Flint et al. (2011)).

While rarer, a few studies have used actual or claimed demonstrations of loyalty as their definition of the dependent variable for loyalty. Rust et al. (1993) worked backwards from the act of loyalty by identifying participants in their study that had both stayed with their current bank and those that had recently defected and moved their financial accounts to another bank. From that sample, they determined satisfaction levels for both groups and modelled the impact of that satisfaction on loyalty to remain with the original bank. In one of the few longitudinal studies of satisfaction and loyalty, Bolton (1998) measured the duration of a customer's relationship during a 22-month period as a measure of loyalty.

2.5.2 Key Independent Variables

The studies reviewed in detail for this section utilize a wide variety of independent variables (IV).

Of the 10 studies that included both customer satisfaction and customer loyalty as dependent variables, seven of those studies also evaluated customer satisfaction as both an IV and DV within their study. The models in these studies included other variables as IVs, measuring their influence on satisfaction (as a DV); then measuring the influence of satisfaction (now acting as an IV) on loyalty.

The 31 studies under review utilized over 25 unique IVs (other than customer satisfaction), as might be expected based on the variety of antecedents of customer satisfaction (25 antecedents alone as identified by Thusyanthy et al. (2017), and discussed previously in Section 2.2.5).

The most common IVs used in the 31 studies reviewed include:

- Customer satisfaction (23 studies)
- Perceived value (12 studies)
- Perceived service or product quality (8 studies)
- Corporate image (5 studies).

The remaining 23 IVs were observed in 2 or fewer of the studies examined. The wide variety of IVs utilized by researchers reflects the variety of industries in which the studies take place, the available of data available to the researchers, and other criteria (such as interest areas of the researchers themselves).

2.5.3 Methodological Approaches Employed

The 31 studies examined in this section employed several different methodologies in conducting their research, of which the most common was a cross-sectional study, conducted on a given sample of the study population at a given point in time (the method used by 27 of the 31 studies). These cross-sectional studies employed surveys (mail, email, web or telephone) to gather the data on the IV and DV constructs used in the study.

One study (Churchill Jr et al. 1982) conducted a laboratory-type experiment in a mall, allowing for manipulation of the IVs across the study participants.

Only three studies were of a longitudinal design, where the data on the variables of interest were measured over a period of time:

- Bolton (1998) measured the duration of mobile phone subscribers' relationships with their providers over a 22-month period, with two waves of telephone surveys to gather data on satisfaction.
- Bolton et al. (1999) analyzed panel data for two study groups (mobile phone and interactive TV service users) over periods of 12 and 6 months, respectively.
- A longitudinal study of telecommunications customers (mobile phone, fixed phone and Internet) was conducted over a 9-month period by Gustafsson et al. (2005).

2.5.4 Key Findings

The research studies reviewed for this section found - on the whole - evidence consistent with the relationships put forward in prior sections (specifically on the antecedents of and the positive relationship between Customer Satisfaction and Customer Loyalty).

Many of the antecedents to customer satisfaction that were reviewed in Section 2.2 found support in these 31 studies. In particular, perceived value was cited most frequently as an independent variable with a significant effect on customer satisfaction. Evidence against the strong role of perceived value was occasionally found, such as the greater role of corporate image (Andreassen et al. 1998, Ha et al. 2013).

In some cases, the IV's were found to have a greater effect on loyalty directly, rather than through satisfaction (Tam 2004, Yang et al. 2004).

2.6 Conclusion

The evidence to support the influence of customer satisfaction on customer loyalty is substantial, with multiple studies across a range of consumer product and service industries supporting the linkage of these two customer attributes, along with the influence of other variables as antecedents or moderators of both satisfaction and loyalty.

However, there are several areas where the existing research is lacking or in need of further study. These proposed questions regarding opportunity for further research are summarized below and explored in further detail in the remainder of this section.

- How does cumulative satisfaction impact loyalty over time?
- Can customer satisfaction be shown to influence behavioural loyalty (and not just attitudinal loyalty)?
- Can practitioners causally affect loyalty, and do so profitably?
- Is perceived value as important an IV as claimed?

If customer satisfaction is accepted as a cumulative assessment by the customer (as discussed in Section 2.2.3), the use of a longitudinal, panel study methodology may contribute important insights into how changes in satisfaction over time are reflected in the loyalty of those same customers, as recently called for by Dey et al. (2020). Questions that might be answered could include: Is the loyalty outcome different for customers who have the same end state of satisfaction, but whose cumulative journeys are different? Does the satisfaction:loyalty relationship hold true when accounting for all exogenous and endogenous factors?

As noted in Section 2.5.1, few studies examine the impact of customer satisfaction on *behavioural* loyalty, that is, does the customer actually demonstrate loyalty through continued purchasing of the company's goods or services over time. Mittal et al. (2001) observed that "Virtually all published studies on satisfaction, even those relying on commercial surveys …, have use repurchase-intent as the criterion variable…" (p. 140). This statement appears to still be true 20 years later. Potentially more worrisome is the risk that attitudinal loyalty may not be a predictor of behavioural loyalty (Bodet 2008).

Accordingly, future research would benefit from studies that remove stated intention of a participant to remain loyal (attitudinal loyalty) as the DV of focus and replace it with a measure – ideally not self-reported but measured through actual transactions – of actual behavioural loyalty.

Studies using a longitudinal approach and a focus on behavioural loyalty might also introduce the use of an experiment, to manipulate the antecedents and independent variables that have been shown to effect loyalty but do so in a way that might demonstrate causality between those IVs and the DV of loyalty.

The use of an experimental methodology would be useful in addressing calls to demonstrate the impacts of customer satisfaction and loyalty on firm profitability (Hu et al. 2009, Cotiu 2013). Understanding what treatments could be employed, at what cost, and to what effect on loyalty would permit a cost/benefit analysis to be conducted on the efficacy of those treatments.

Lastly, the frequency of studies indicating that perceived value has a strong influence on loyalty – either directly or working through customer satisfaction – suggests that further examination of this relationship under the controlled environment of an experimental study might yield insights to support and quantify this influence.

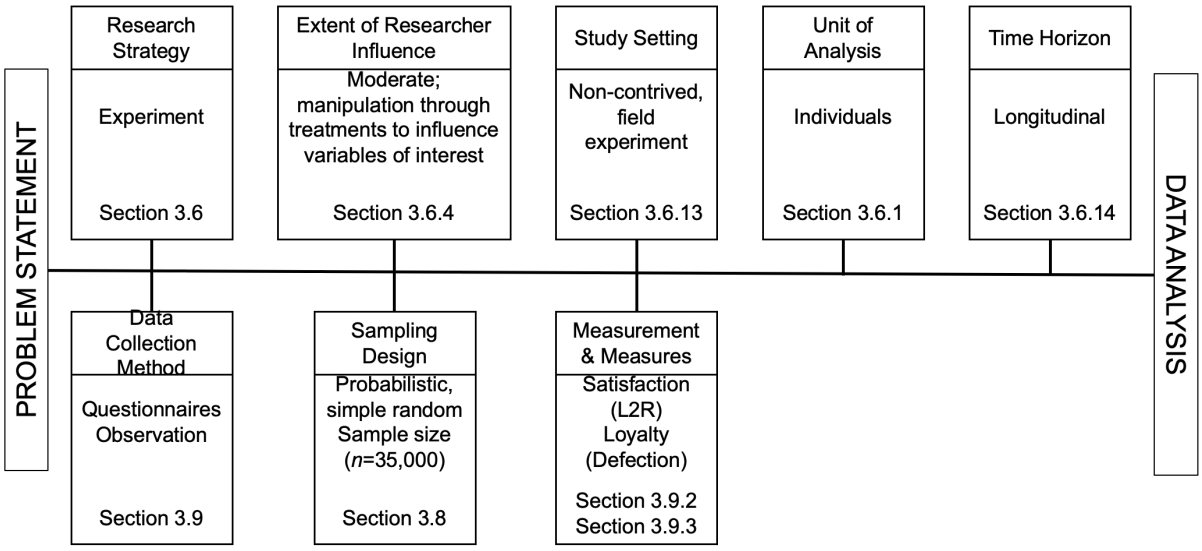
3 Methods

3.1 Introduction

The research problem identified in Section 1.2 made a case for providing practitioners with evidence as to whether their efforts to improve customer satisfaction and customer loyalty were causal. The literature reviewed in Chapter 2 demonstrated that while there is a substantial body of work to support the relationship between efforts to improve satisfaction and increased loyalty, there is little in the way of causal proof that such a relationship exists. Also, there is limited research showing the degree to which the actions of an organization causally influence the antecedents of customer satisfaction and loyalty themselves.

The following sections of this chapter detail the conceptual model and hypotheses to address the research problem, and the research methods selected to test those hypotheses. The various elements addressed in the study design are shown in Figure 7 (Sekaran et al. 2016, p. 96).

Figure 7 Study Design



Source: Adapted from Sekaran et al. (2016)

3.2 Ontological and Epistemological Perspective

The study developed from the ontological position of realism. This realist position takes the view that there are facts that exist and can be measured, and that together these facts describe a single truth (Easterby-Smith et al. 2015).

The positivistic approach makes several philosophical assumptions that align strongly with the intended study. In particular, Easterby-Smith et al. (2015) highlight the following assumptions which are foundational to this study:

- Causality – identification of causal explanations in human behaviour
- Hypothesis and deduction – process of hypothesizing relationships and deducing types of observations can support or falsify those hypotheses
- Operationalization – facts that can be measured quantitatively
- Generalization – random sampling from a whole, which supports inferences about the broader population

The study adopts the epistemological position of positivism. The implications of this position include (Easterby-Smith et al. 2015):

- The observer must be independent
- Human interests should be irrelevant
- Explanations must demonstrate causality
- Research progresses through hypotheses and deductions
- Concepts need to be defined so they can be measured
- Units of analysis should be reduced to the simplest terms
- Generalization through statistical probability
- Sampling requires large numbers selected randomly

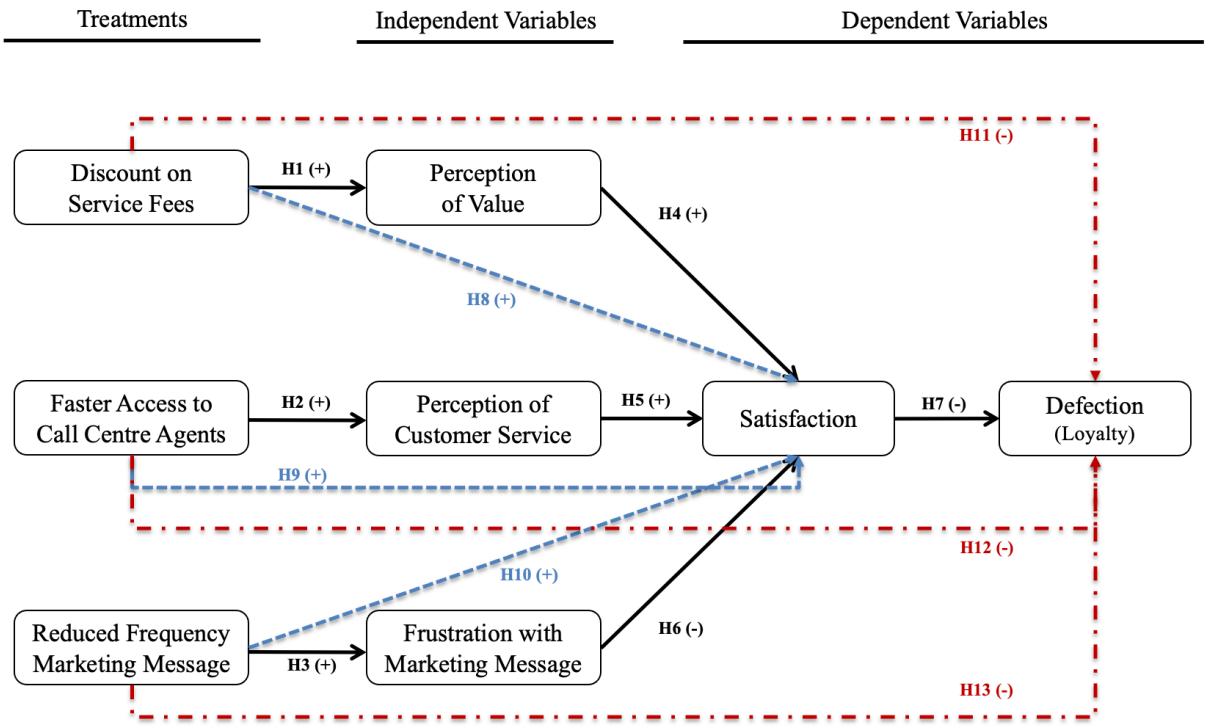
Subsequent sections of this chapter illustrate how the study design conforms to these implications.

3.3 Study Model

The introduction to this study (Section 1.3) set out the research problem to be investigated, specifically to determine i) if a causal linkage between customer satisfaction and customer loyalty (churn) exists, and ii) if there exists an ability to influence customer satisfaction and loyalty through specific, practical and scalable treatments.

Based on the previous studies reviewed in Chapter 2, and the research questions established above, the study utilized the model shown in Figure 8. The model establishes a series of three treatments, each of which intends to influence a respective independent variable. These three independent variables, in turn, are hypothesized to impact customer satisfaction and loyalty.

Figure 8 Study Model



For reference throughout the rest of this study, the following names were defined for the treatments, independent variables and dependent variables shown in Figure 2:

Table 3.1 Treatment and Variable Names in Model

Name	Component of Model
Treatments:	
VALUE	Discount on Service Fees
CSERV	Faster Access to Call Centre Agents
CONTACT	Reduced Frequency of Marketing Message
Independent Variables:	

PERCVALUE	Perception of Value
SERVICE	Perception of Customer Service
MARKETING	Frustration with Marketing Message
Dependent Variables:	
CSAT	Satisfaction
DEFECT	Defection (Loyalty)

3.4 Hypotheses

The model in Figure 8 consists of thirteen hypotheses to be tested in the study, each stated below in their null (H_{0x}) and alternative (H_{ax}) hypothesis formats:

H_{01}: ***There is no relationship between receiving a discount on monthly service fees and a customer's perception of value***

H_{a1}: ***There is a positive relationship between receiving a discount on monthly service fees and a customer's perception of value***

H_{02}: ***There is no relationship between faster access to call centre agents and a customer's perception of customer service quality***

H_{a2}: ***There is a positive relationship between faster access to call centre agents and a customer's perception of customer service quality***

H_{03}: ***There is no relationship between reduced frequency of marketing messages and a customer's frustration with receiving such marketing messaging***

H_{a3}: ***There is a positive relationship between reduced frequency of marketing messages and a customer's frustration with receiving such marketing messaging***

H_{04}: ***There is no relationship between perception of value and customer satisfaction***

H_{a4}: ***There is a relationship between perception of value and customer satisfaction***

H_{05}: There is no relationship between perception of customer service and customer satisfaction

H_{a5}: There is a relationship between perception of customer service and customer satisfaction

H_{06}: There is no relationship between frustration with marketing messaging and customer satisfaction

H_{a6}: There is a relationship between frustration with marketing messaging and customer satisfaction

H_{07}: Higher levels of customer satisfaction have no impact on customer defection (i.e. improved customer loyalty)

H_{a7}: Higher levels of customer satisfaction cause a decrease in customer defection (i.e. improved customer loyalty)

H_{08}: Receiving a discount on monthly service fees has no impact on customer satisfaction

H_{a8}: Receiving a discount on monthly service fees causes an improvement in customer satisfaction

H_{09}: Faster access to call centre agents has no impact on customer satisfaction

H_{a9}: Faster access to call centre agents causes an improvement in customer satisfaction

H_{010}: Reducing the frequency of marketing messages has no impact on customer satisfaction

H_{a10}: Reducing the frequency of marketing messages causes an improvement in customer satisfaction

H_{011}: Receiving a discount on monthly service fees has no impact on customer defection

H_{a11}: Receiving a discount on monthly service fees causes an improvement in customer defection (i.e. lower rates of defection)

H_{012}: Faster access to call centre agents has no impact on customer defection

H_{a12}: Faster access to call centre agents causes an improvement in customer defection

H_{013}: Reducing the frequency of marketing messages has no impact on customer defection

H_{a13}: Reducing the frequency of marketing messages causes an improvement in customer defection

These thirteen hypotheses influence the approach to be taken in the research design. In particular, the causal hypotheses expressed in H_4 through H_{13} indicate an opportunity to leverage a specific research design that can determine if such causal relationships exist and if so, the nature of that relationship.

3.5 Research Design

The research design for this study takes the form of an experimental, longitudinal, randomized controlled trial (RCT). The use of this experimental design – especially with the use of large, randomly assigned participant groups (including a control group) – allows for the identification and measurement of the causal relationships hypothesized in Section 3.4. The design follows that of (Bolton 1998), whose research examined the cumulative impact of customer satisfaction on repurchase behaviour over an extended period of tim.

This section begins with an overview of the critical aspects of experimental research design and the criteria to be met in evaluating causal relationships. The remainder of this section then details the specific aspects of the research design implemented in this study.

3.5.1 Experimental Design Overview

The earliest examples of experimental design originate in the medical (18th century) and agricultural (19th century) fields (Dean et al. 2015). These experiments were critical to the evolution of experimental design and the statistical principles associated with it, and have become useful in the social sciences as well (Salkind 2010).

Experimental research has been described as "…the ideal way of collecting knowledge" (Balnaves et al. 2001), yielding the strongest data to address the research problem proposed by the researcher. Salkind (2010) calls experimental designs "…the most rigorous quantitative research methods", with a primary advantage being it can be used to test a hypothesis and demonstrate a cause-and-effect relationship. The ability to identify causal effects was the motivation (Duflo et al. 2017) for the examination of randomized experiments by researchers early in the 20th century, notably Splawa-Neyman et al. (1923) and Fisher (1925)

According to Shadish et al. (2002, page 23), research experiments that consist of the combination of methods employed in this one study are quite rare, partly due to the requirement for "…a high level of resources and degree of logistical control that is rarely feasible". The research design utilized in this study leverages a unique opportunity to overcome most – if not all – of the difficulties customarily faced by researchers in this area and provide a unique contribution to the study of consumer behaviour.

3.5.2 Definition of Experimentation

The criteria that define experimental research are widely agreed. Shadish et al. (2002, page xvii) define experimental research as "…a systematic study designed to examine the consequences of deliberately varying a potential causal agent.". The same authors proceed to summarize four critical requirements for an experimental study:

1. Variation in the treatment(s)
2. Post-treatment measures of the outcomes
3. A minimum of one unit upon which the observations are made
4. A mechanism for determining what the outcome would have been without the treatment

The research design reviewed in section 3.6 provides the detailed plans to meet these four requirements in this study, as well as other requirements proposed as the basis for robust experimental research design.

3.5.3 Defining a "Good" Experiment

Beyond the minimum criteria set forth in Section 3.5.2 by Shadish et al., other criteria have been proposed to define what constitutes good experimental research design

Oehlert (2010) suggests that a good experimental design must:

- Avoid systematic error (suggesting randomization as a solution)
- Be precise (through robust research design)
- Allow for estimation of error
- Have broad validity

This study satisfies these requirements through the use of random sampling and random assignment (Section 3.8), the use of an experimental RCT design (Section 3.6), support for multiple statistical analysis techniques (Section 4.2) and strong validity controls (Section 3.7).

3.5.4 Criteria for Establishing Causal Relationships

Three concepts are proposed to be associated with causation by Mosteller and Tukey (as cited in Oehlert 2010), with the authors stating that at least two are required to support the existence of a causal relationship:

1. Consistency – all other things being equal, the relationship between two variables is consistent across populations in direction and maybe in amount
2. Responsiveness – one can go into a system, change the causal variable, and watch the response variable change accordingly
3. Mechanism – a step-by-step mechanism leading from cause to effect

3.6 Design of Experiment Utilized in this Study

Taking into account the criteria for the definition of strong experimental research design and the support for claiming a causal relationship (should one be found), this study utilizes an experimental, field-based, longitudinal, randomized controlled trial (RCT) design.

The specific research design implemented in this study can be described as a "Pretest-Posttest Control-Group Factorial Design" (Christensen et al. 2015). Details of the research design are reviewed in the remaining sections of this chapter, but in summary, the design consists of:

- Pretest-Posttest – measurements of the dependent variables are conducted before and after the experimental treatments are applied, allowing for measurement of the treatment effect(s)
- Control Group – a group of study participants, equal in all other ways to the treatment groups, to which no treatment is applied. The control group allows for control of extraneous variables.
- Factorial Design – for experiments with more than one treatment or levels of treatment, the various combinations (or factors) are each measured separately through the experiment

This study exhibits the strongest characteristics of good experimental design (Shadish et al. 2002, Salkind 2010) and minimizes threats to internal validity (see section 3.7 for a detailed review of validity threats in this study).

As outlined by Christensen et al., this design offers several advantages, including:

1. Allows for validation of the randomization process (specifically the random assignment of study participants into control and treatment groups)
2. If there are differences between the control and treatment groups at the time of the pretest, covariance techniques can be used at the conclusion of the study to control for those differences
3. Provides for identification of whether a ceiling or floor effect is likely to occur (i.e. a pretest score so high or low that variance as a result of the treatments might be unlikely)
4. Provides (most importantly for Christensen et al.) the ability to empirically demonstrate that a change in the dependent variable (DV) occurred from the pretest to the posttest.

Christensen et al. also highlight one potential disadvantage to the pretest-posttest design – the possibility that the administration of the pretest itself might affect the research participants somehow. In this study, the pretest will be given to both the control and treatment groups, and the pretest measure itself (a short customer satisfaction survey) is a common measurement tool that many of the participants in the study would have received at some point prior (either from the subject organization or from other companies), and the survey itself does not make it obvious that the participant is the subject of a study.

3.6.1 Experimental Unit

The experimental unit in this study is a consumer mobile service subscriber, selected from the population of such subscribers from a large Top 3 Canadian telecommunications provider (the “subject organization”).

Specifically, the population on which the study is based is defined as those in the Consumer Premium Postpaid subscriber segment (as described in Section 1.1) as follows:

- Consumer – subscribers of the mobile service plans offered to the consumer segment of the market (as opposed to small or medium business)
- Premium – subscribers on the primary, main brand service offered by the subject organization
- Postpaid – subscribers on postpaid rate plans

This Consumer Premium Postpaid segment is, based on the industry information reviewed in Section 1.1, the largest segment of the mobile phone market in Canada.

In addition, the same mobile service subscribers defined here as the experimental unit will also serve as the measurement unit, in that measures of the dependent variables CSAT and DEFECT will be obtained on and from each individual subscriber included in the study.

3.6.2 Dependent Variables

The study utilizes two dependent variables (DV) to measure the outcomes of the experiment, as shown in the study model earlier in this chapter:

- Defection / Loyalty (DEFECT)
- Customer Satisfaction (CSAT)

Defection (the “DEFECT” DV) in the mobile services industry is defined by the event that occurs when the subscriber’s service is cancelled with a given provider. This cancellation may take one of two forms:

1. Voluntary cancellations, where the subscriber initiates the cancellation. This cancellation may take place by the subscriber notifying their service provider of the intent to cease service, or (most frequently) by transferring their service to another provider utilizing the number portability regulations discussed in Section 1.1.1.
2. Involuntary cancellations, whether the service provider initiates the cancellation of the service, usually as a result of non-payment for services on the part of the subscriber.

With either type of cancellation, the subscriber's service is terminated, the account status is changed to "cancelled", and the date and reason for the cancellation are noted. This change in status allows for the observed measurement of the DEFECT DV without a self-reporting requirement from the study participant. Further details of how this DV is implemented in the study are reviewed in Section 3.9.2.

The Customer Satisfaction dependent variable (the "CSAT" DV) is based on self-reported data collected through pretest and posttest surveys deployed to participants in the study. In order to ensure maximum applicability of the research findings to the subject organization, this study utilizes the same measure for customer satisfaction that the organization uses in its primary research and performance management metrics. Within the organization, a given customer's likelihood to recommend (L2R) the company to others is the primary measure of customer satisfaction and has been used consistently for over ten years to guide investments and performance (as part of company and business unit scorecards).

Details of the instrumentation of the CSAT DV are provided in Section 3.9.3.

3.6.3 Independent Variables

The conceptual model proposed includes three independent variables (IVs), one IV for each of the three treatments:

- Perception of Value (PERCVALUE)
- Perception of Customer Service (SERVICE)
- Frustration with Marketing Message (MARKETING)

Each of these three IVs is measured through the same pretest and posttest surveys used for the CSAT DV; the instrumentation details are provided in Section 3.9.4.

3.6.4 Treatments

Core to the experimental study method is the use of one or more treatments (Salkind 2010, Creswell 2014). The identification of appropriate treatments to be used in this study is based on a combination of two factors:

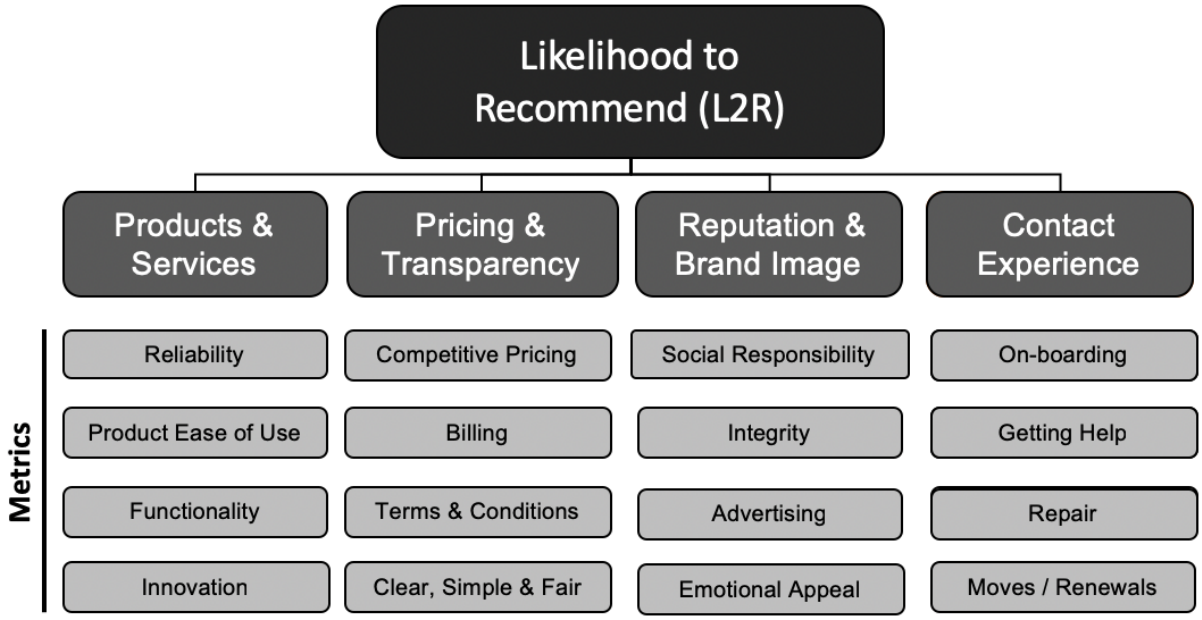
- The treatments suggested by the antecedents of customer satisfaction and customer loyalty in previous research and reviewed previously in Chapter 2
- Discussion with multiple teams within the subject organization

From prior research, multiple possible treatments are suggested for further examination in an experimental study such as this. Section 2.2.5 details antecedents that could form the basis for potential treatments in this study.

In addition to the potential treatments identified in the literature, it was also determined that the treatments should be based on input from the subject organization itself. There are several reasons for including this input in the treatment selection. Firstly, the organization has a great deal of experience and history in managing customer satisfaction and loyalty, and the use of this experience in determining appropriate treatments for the study would allow for greater opportunity for the research findings to be applicable to the real-world business challenges faced by the organization. Secondly, as the experimental study is conducted in the field on actual customers of the organization, the treatments need to be those that can be implemented within the constraints of the operations of the company (including, but not limited to, billing systems, call centre applications, employee training and financial systems). Thirdly, the treatments needed to be scalable, both at the experimental study level (on the sample size determined in section 3.8) and, if the insights of the research show that the treatment could create better customer and financial outcomes, at full scale (which could entail several million customers).

Over the past ten years, the subject organization has undertaken substantial primary research to identify those factors that impact customer satisfaction. This research has resulted in the development of a framework to measure the various attributes that contribute to the organization's primary measure of customer satisfaction, which is the same as one of two DVs in this study: Likelihood to Recommend (L2R). The high-level L2R framework used by the subject organization is shown in Figure 9.

Figure 9 Likelihood to Recommend (L2R) Framework



Each of the four categories shown in the L2R framework used by the organization consists of several metrics which, based on the subject organization's experience, have been seen to be strongly correlated with likelihood to recommend.

The determination of which treatments should be used in the experimental study included discussions with the in-house research team that administers and reports the outcome of the quarterly L2R survey, as well as key executive personnel including the Executive Vice-President Marketing, various Vice-Presidents of Marketing and the Director of Research.

Based on a combination of prior research in the field, and input from the subject organization, the following three treatments were identified to be utilised in the experimental study:

1. Value (VALUE) – a treatment to manipulate the cost of the monthly service paid by each participant in the study
2. Customer Service (CSERV) – a treatment to manipulate the amount of time a participant needs to wait until speaking to a call centre agent
3. Marketing Frequency (CONTACT) – a treatment that manipulates the frequency that a participant in the study receives unsolicited sales and marketing communications (email or text messages)

The specific details of each treatment's design, implementation in the study, and potential impact on the independent variables (IV) defined in the model are explored in the following three sections.

3.6.5 "VALUE" Treatment

The VALUE treatment for this study was derived from findings by the subject organization that customer perceptions of value have an impact on satisfaction. When asked "What is it that makes you hesitate about recommending…" a wireless provider, the response "Pricing/too expensive/costs" is the most cited reason by a factor of over 10:1 versus other Pricing and Transparency responses.[11]

The cost of service for a mobile subscriber is determined by their total monthly service invoice, which in turn is comprised of two types of charges: monthly recurring charges (MRC) and variable usage charges. The MRC is the primary component of the monthly invoice, and is the fee associated with the price plan that the customer has chosen for their service as well as any "add-ons". The monthly cost of the plan varies depending on the attributes of the plan chosen and can range from a low of $15 to $200 or more. Unless the subscriber

[11] Source: Subject organization Likelihood to Recommend research results

changes their rate plan or adds or removes add-ons, the MRC amount on their service invoice remains the same month-over-month. Usage charges can vary for each subscriber, from no charges if the subscriber does not use any services not covered by the MRC of their monthly plan to several hundred (or thousand) dollars if the customer uses a large amount of a service not included in their rate plan (e.g. long-distance calls or additional data).

Examples of rate plans available to customers at the time of the study are shown in Table 3.2.

Table 3.2 Example Monthly Rate Plans

Plan Component	Plan A	Plan B	Plan C
Monthly Fee	$85	$95	$85
Included Attributes	Unlimited nationwide talk Unlimited nationwide text messaging 10GB data	Unlimited Canada & US talk Unlimited Canada & US text messaging 10GB data	Unlimited nationwide talk Unlimited nationwide text messaging 20GB data shareable with connected devices
Add Ons	Voicemail @ $5/month Additional data packages @ $varies Extended Warranty @ $14/month		
Usage Fees	Voice roaming outside Canada @ $1.50/minute Long-distance voice calls to the US @ $0.60/minute Canada to US text messages @ $0.40/message Additional data over included amount @ $10/100MB		

The VALUE treatment implemented in this study takes the form of a discount on the participant's monthly service invoice. There are several ways in which a discount can be provided to a mobile subscriber, with the usual forms being either a fixed dollar discount or a percentage discount on some or all of the service costs. In order to provide each study participant with a discount that has a similar relative value for each participant, a percentage discount in the amount of 10% off the MRC charges was chosen as the treatment. This

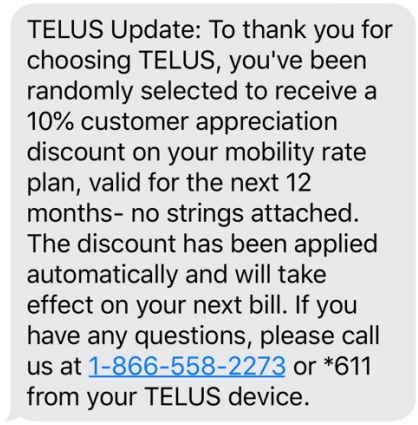
treatment design also allowed for accurate forecasting of the cost of providing the treatment used in the study, which prior to the study execution was estimated at $1.3 million, and was approved and paid for by subject organization management.

The execution of the VALUE treatment consisted of the creation of a discount code within the study organization's billing system. When added to the account of study participants that were designated to receive the VALUE treatment in the study, the participant would have their monthly invoice reduced by the calculated amount of Total MRC x 10%. For example, of the participant' total MRC charges each month were $85, they would receive a discount on that same month's invoice in the amount of $8.50.

Upon commencement of the experiment, participants in the VALUE treatment group were informed that they would be receiving a discount for a twelve-month period. This communication was required as the participant would see the discount on their monthly invoice, and if not informed of their receiving the discount in advance, might be concerned that the discount was applied in error.

Each participant in the VALUE treatment group received a text message at the beginning of the study, as shown in Figure 10.

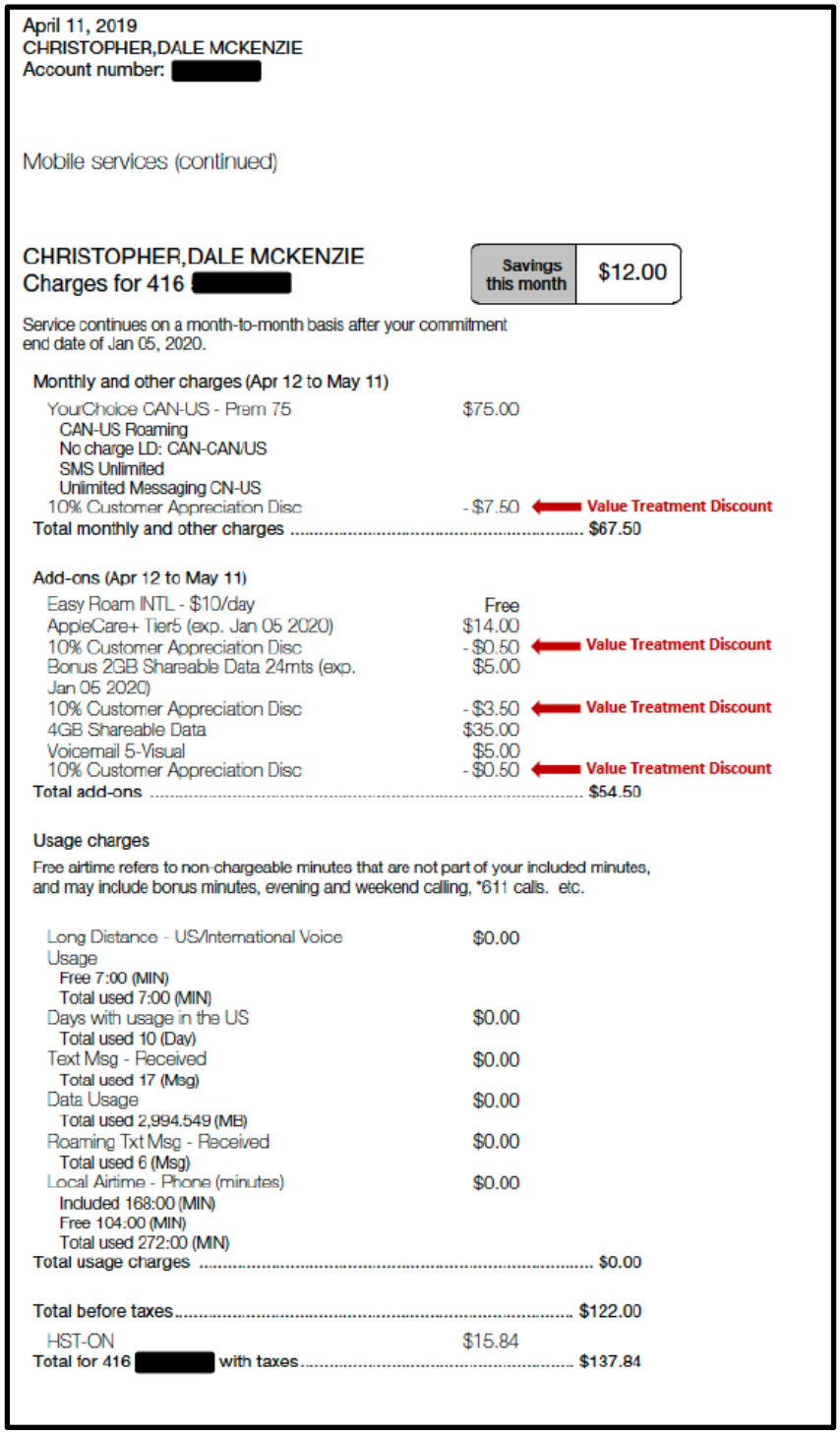
Figure 10 VALUE Treatment Text Message



For the 12 months of the experiment, participants in the VALUE treatment groups would see the 10% discount on their monthly invoice. An example invoice is shown in Figure 11. In this example, the study participant would have received the 10% discount on the MRC of their monthly rate plan (shown as $75.00, with a discount of $7.50) as well as a discount on the MRC of other charges (shown under "Add-ons", with three discounted items receiving a total discount of $4.50). Usage charges, which are not MRC fees, did not receive any discount. For the entire monthly invoice shown, the study participant would have received a $12.00 discount.

Treatment data is collected monthly on each participant in the study, capturing the various charges each participant incurred on their service invoice (both MRC and variable usage), as well as the discount amount if they were in the assigned treatment group. The results for both treatment and non-treatment subjects are reviewed in Section 4.3.1.

Figure 11 VALUE Treatment Monthly Invoice Example



April 11, 2019
CHRISTOPHER,DALE MCKENZIE
Account number: [redacted]

Mobile services (continued)

CHRISTOPHER,DALE MCKENZIE
Charges for 416 [redacted]

Savings this month $12.00

Service continues on a month-to-month basis after your commitment end date of Jan 05, 2020.

Monthly and other charges (Apr 12 to May 11)

YourChoice CAN-US - Prem 75	$75.00	
CAN-US Roaming		
No charge LD: CAN-CAN/US		
SMS Unlimited		
Unlimited Messaging CN-US		
10% Customer Appreciation Disc	-$7.50	← Value Treatment Discount
Total monthly and other charges		**$67.50**

Add-ons (Apr 12 to May 11)

Easy Roam INTL - $10/day	Free	
AppleCare+ Tier5 (exp. Jan 05 2020)	$14.00	
10% Customer Appreciation Disc	-$0.50	← Value Treatment Discount
Bonus 2GB Shareable Data 24mts (exp. Jan 05 2020)	$5.00	
10% Customer Appreciation Disc	-$3.50	← Value Treatment Discount
4GB Shareable Data	$35.00	
Voicemail 5-Visual	$5.00	
10% Customer Appreciation Disc	-$0.50	← Value Treatment Discount
Total add-ons		**$54.50**

Usage charges

Free airtime refers to non-chargeable minutes that are not part of your included minutes, and may include bonus minutes, evening and weekend calling, *611 calls. etc.

Long Distance - US/International Voice	$0.00	
Usage		
Free 7:00 (MIN)		
Total used 7:00 (MIN)		
Days with usage in the US	$0.00	
Total used 10 (Day)		
Text Msg - Received	$0.00	
Total used 17 (Msg)		
Data Usage	$0.00	
Total used 2,994.549 (MB)		
Roaming Txt Msg - Received	$0.00	
Total used 6 (Msg)		
Local Airtime - Phone (minutes)	$0.00	
Included 168:00 (MIN)		
Free 104:00 (MIN)		
Total used 272:00 (MIN)		
Total usage charges		**$0.00**

Total before taxes		**$122.00**
HST-ON	$15.84	
Total for 416 [redacted] with taxes		**$137.84**

3.6.6 "CSERV" Treatment

One of the primary methods that consumer customers of the subject organization can obtain assistance with their mobile service is to call the company's contact centre to speak to a customer service agent. On average, 23% of the customer base of mobile subscribers will call the contact centre in any given month[12]. Reasons for calling may include questions regarding their monthly invoice, requesting changes to their service, technical support or a wide variety of other reasons.

There are multiple factors that could influence a customer's overall level of satisfaction as a result of an interaction with calling the contact centre. Within the subject organization, the primary determinants of customer satisfaction when making a call into the contact centre are believed to be the following:

- Queue time – the amount of time spent in the queue prior to speaking to an agent
- First call resolution – the ability of the agent to resolve the customer's reasons for calling on the first attempt
- Abandonment – the need for customers to abandon their call (that is, hang up) prior to speaking to an agent; usually the result of long queue times
- Length of call – the length of time the customer spent on the call, both in the queue and with the agent

The first three of these four factors (queue time, first call resolution and abandonment) were identified in previous research as antecedents to customer satisfaction (see Section 2.2.5.2).

Results from the subject organization's Likelihood to Recommend survey indicates that "Poor customer service" and "Slow service/long wait periods to speak to reps" were the top two reasons cited for not recommending a given wireless provider[12].

From a treatment perspective, it was identified that specific technology platforms employed by the subject organization could be utilized to support a treatment that combines the "queue time" satisfaction determinant identified by the company, and indirectly, the rate of abandonment identified by prior research. Details of the treatment design and implementation follow.

[12] Based on subject organization results for Q1 2017

In many organizations, the contact centre operations function as follows (simplified view):

1. The customer initiates a call by dialling a toll-free (e.g. 1-800) or specially provided number (such as *611) from their mobile device
2. The call is answered by an interactive voice response (IVR) system. The IVR performs multiple functions:
    a. Identification and authentication of the calling customers. This may happen automatically for customers calling from their mobile device, or by the customer entering an account number in the IVR
    b. Determination of the reason for the call, and the appropriate queue to assign the call. Different agents have the ability to solve different customer problems (such as specially trained technical support agents or agents that can assist with more general inquiries)
    c. Management of the queues to ensure that each customer is held "on hold" until the next available agent is able to take their call. In most systems and for most customers, calls are handled on a first-come, first-served basis
    d. Upon availability of a suitable agent the IVR transfer the call to the agent
3. Contact centre agent receives the call and proceeds to address the reason for the customer's call

Based on a sample week's statistics[13], customers of the subject organization experienced an average queue time of 3.3 minutes. The distribution of actual queue times for individual caller can vary substantially from the average. In the sample obtained, approximately 25% of callers experienced a queue time in excess of 5 minutes, and 8% experienced a queue time of 10 minutes or more.

Many factors can influence the queue time experienced by a customer, including overall volume of calls being received at a given time, the time of day, the number of agents available to take calls, and the average length of each call.

The CSERV treatment involves the manipulation of how the subject organization's IVR handles the queueing of the customer calls. Rather than the traditional first-come, first-served approach, study participants receiving this treatment were prioritized to have their call handled by the next available agent, regardless of the participant's order in the queue. This results in a shorter queue time for each occurrence where the customer calls for support.

[13] Sample statistics from subject organization's contact centre, week of January 6-12, 2017

Implementation of this treatment was conducted by the contact centre technology team of the subject organization. A file containing the subscriber information for all study participants assigned to the CSERV treatment was provided, and the technology team implemented various system rules to affect the treatment.

Data is collected on each participant in the study, identifying the number of calls made by the participant in a given time period and the amount of queue time for those calls. As the treatment is only experienced by those participants who attempt to call the contact centre, not all study participants will necessarily be exposed to the treatment. The length of the program study, combined with data from the research organization that almost a quarter of all customers call the contact centre in a given month, helps to mitigate this risk of low exposure to the treatment.

The actual difference in queue times experienced by those participants in the treatment versus those not in the treatment throughout the experiment is provided in Section 4.3.2.

3.6.7 "CONTACT" Treatment

The third treatment is one that modifies the frequency of unsolicited marketing communications sent by the study organization to participants. As with most commercial organizations, the subject organization sends email and text message communications to its customers, informing them of special offers, new products and services, and other marketing messages designed to maximize the financial relationship.

Dreze et al. (2003) showed that the frequency of marketing messages sent to customers could have an impact on that customer's satisfaction and retention levels, with longer periods between communications positively impacting customer behaviour. The subject organization itself includes a question in its likelihood to recommend survey that asks respondents to rate their frustration level with communications from the company, including relevance or frequency of emails.

The CONTACT treatment reduces the frequency of marketing communications to any study participant in the treatment group relative to the other participants in the study. This treatment was executed within the database marketing (DBM) function of the study organization through the uploading of participant identification information into the DBM systems, with instructions to remove those participants from marketing communications.

Study participants in all groups, treatment or otherwise, continued to receive email and text messages that were required for regulatory reasons, associated with managing their service or were initiated automatically in systems outside of the DBM team control.

Data is collected on each participant in the study, tabulating the number of communications (marketing and regulatory) received by each participant each month. The difference communication frequency experienced by those participants in the treatment versus those not in the treatment throughout the experiment is provided in Section 4.3.3.

3.6.8 Treatment Levels

Each of the three treatments was implemented at two levels:

- Normal – no treatment; participant receives no change to their monthly invoice amount, call centre treatment, or frequency of marketing communications
- Enhanced – treatment is applied; participant receives the treatment(s) as assigned to their study group

The complete set of treatments employed in the experimental study are summarized in Table 3.3.

Table 3.3 Summary of Treatments

Treatment Code	Treatment Name	Treatment Description	Levels in Treatment	Implemented By
a	VALUE	Recurring 10% discount on monthly MRC fees	2	Marketing team
b	CSERV	Lower queue times when calling into the contact centre	2	Call Centre Technology group
c	CONTACT	Reduced frequency of marketing emails and text messages	2	Database Marketing team (DBM)

3.6.9 Factorial Design

Factorial designs are an efficient method of studying the impact of multiple treatments within a single experiment (Salkind 2010, Montgomery 2013), allowing for smaller sample sizes when each participant is able to receive one or more of the treatments at the same time.

This efficiency allows for the ability to pose the following questions in the study:

- What effects do the three treatments each have on the dependent variables?
- Is there any one treatment that would consistently impact customer satisfaction or loyalty, regardless of the other treatments?

Oehlert (2010) suggests two additional advantages to using a factorial design for a research experiment:

1. Factorial experiments allow for the estimation of the interaction between treatment factors, something that cannot be done with one-at-a-time factor experiments. Oehlert suggests that the use of one-at-a-time experiments in the presence of potential interactions between factors can lead to misunderstandings of the causal relationships.
2. In the absence of interactions between factors, the factorial design is more efficient in that the main effects for all treatment factors can be measured simultaneously.

At the same time, Oehlert suggests that a factorial design may be wasteful if:

- There is a belief there will be no higher-order interactions amongst the independent variables (IVs)
- The purpose of the study is to screen a large number of treatments to identify those of interest for a follow-up study
- There are limited resources to support the study (for example, access to the number of study participants to sample or management of the treatment groups)

In this study, there is no evidence to suggest that there might not be higher-order interactions amongst the IVs. In summarizing the advantages of factorial design, Oehlert (2010, p. 170) concludes that “Factorial structure is a win, whether or not we have interaction”. Additionally, the subject organization has made sufficient resources available – including access to the study population – that the issue of limited resources as outlined by Oehlert, does not exist.

Several authors (Shadish et al. 2002, Salkind 2010) identify potential challenges with factorial experiments, specifically the need for “…rigorous control over the combinations of treatments administered to each experimental unit…” while noting that “…maintaining control over multiple factors and multiple levels can pose special challenges” (Salkind 2010, p. 1,545). The three treatments utilized in this study are executed by a combination of enterprise technology systems in the subject organization. Care was taken to ensure that each treatment was implemented according to the assignment determined in the factorial design below, and that non-treatment participants did not receive a treatment which they should not have.

The factor design suggested by the treatments outlined in sections 3.6.4 through 3.6.7 follows that of an important class of factorial designs: the 2^k factor design (Montgomery 2013).

In the case of this study, the factor design would be noted as a 2^3 design; that is, a factorial design consisting of three factors, each at two levels. Using the treatment code designations from Table 3.3 and assigning the treatment levels as 0 (Normal) or 1 (Enhanced), the implementation of the 2^3 design results in eight distinct treatment combinations represented by the notation format below.

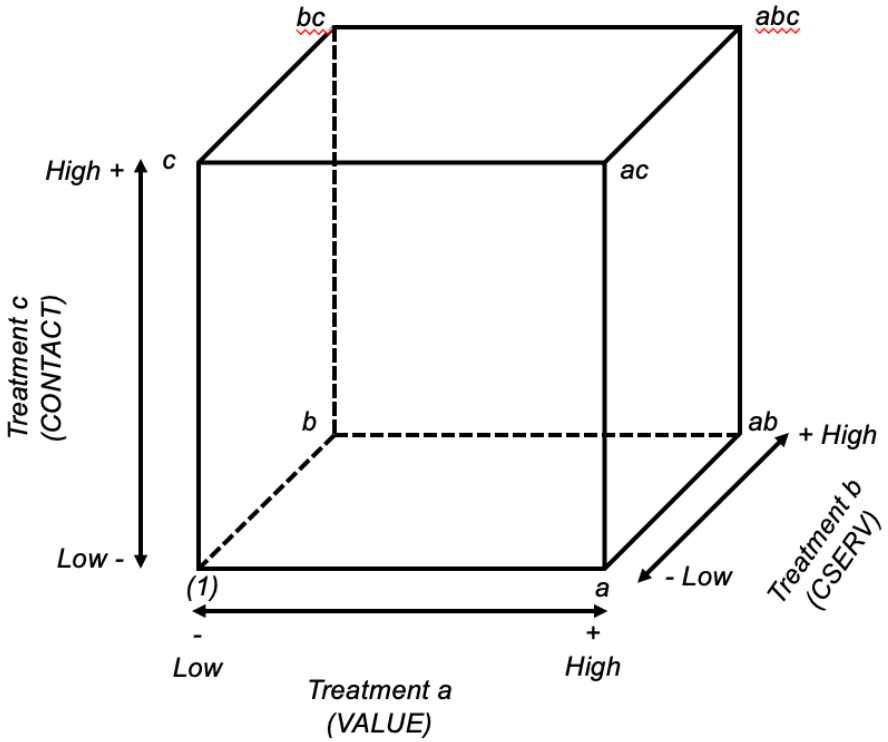
000 = $a_0b_0c_0$ = (1)	a, b, and c are all at first (normal) level; control group
100 = $a_1b_0c_0$ = a	a at second level (enhanced), b and c at first level
010 = $a_0b_1c_0$ = b	b at second level, a and c at first level
110 = $a_1b_1c_0$ = ab	a and b at second level, c at first level
001 = $a_0b_0c_1$ = c	c at second level, a and b at first level
101 = $a_1b_0c_1$ = ac	a and c at second level, b at first level
011 = $a_0b_1c_1$ = bc	b and c at second level, a at first level
111 = $a_1b_1c_1$ = abc	a, b and c at second level

In this factorial arrangement, the treatment combination identified as $a_0b_0c_0$ (that is, with all factors at 0 or normal treatment levels) serves as the control group against which the other effects of the other treatment combinations will be compared.

The design results in seven degrees of freedom amongst the eight treatment combinations, with three degrees of freedom associated with the main effects of *a*, *b* and *c*, and a further four degrees of freedom from the interactions of *ab*, *ac*, *bc*, and *abc* (Montgomery 2013).

The factorial arrangement can also be expressed graphically, as shown in Figure 12.

Figure 12 Graphical Representations of the 23 Factorial Design



Adapted from (Montgomery 2013, p. 240)

3.6.10 Balanced Factorial Design

The experimental study utilizes a balanced factorial design format, whereby the number of measurement units will be the same in each treatment combination. This is an important aspect of the design in order to support the standard analysis of factorial responses at the conclusion of the research experiment (Oehlert 2010).

3.6.11 Between Subjects Design

Within the factorial design of the study, the treatments were implemented in a between-subjects design. In this design, each of the participants in a given treatment group will receive only that one set of treatments (which may consist of no treatment, or one to three treatments, as defined in the factorization design in Section 3.6.9).

Generally, the study was not constrained in the creation of the sample sizes required for each treatment group, thereby eliminating the advantage of the competing design: the within-participants design (Christensen et al. 2015). Additionally, the between-subjects design

provides for a strong degree of control for internal validity (see Section 3.7.3 for a discussion of the internal validity of this study).

3.6.12 Pretest/Posttest Measures of the CSAT Dependent Variable

In addition to the factorial design, the study also incorporates another important aspect of strong research design: the pretest-posttest design. The use of a pretest allows for the identification of differences between the treatment groups prior to the commencement of the experiment, which therefore allows for stronger inferences to be made from the posttest observations after the experiment is concluded (Salkind 2010). As noted by Shadish et al. (2002, p. 260) "…some observers cite the need for a pretest as one of the most important lessons to emerge from the last 20 years of social experiments.".

Several conditions are also proposed by Shadish et al. (2002) for when it might be inappropriate to include a pretest in an experimental study, including i) when the pretest might have a sensitizing effect on the participants, ii) when a pretest is impractical, or iii) when the observations to be gathered in the pretest are known to be constant across the treatment groups. In this study, the pretest used is a common survey metric that is not believed to contribute to any sensitization of the participants and can be easily deployed using existing survey resources of the subject organization. Although effective random sampling and assignment techniques (Section 3.8) should ensure that there are no pre-experiment differences in the treatment groups, the use of the pretest provides data to support this assumption and therefore contributes to stronger analysis and conclusions of the treatment effects.

This design requires that one of the two dependent variables (DVs) in the study be measured for each study group at the start of the experiment, prior to the implementation of any of the treatments. Specifically, the DV for customer satisfaction (CSAT) will be measured at the beginning of the study period and again at the end of the study. Details on the definition and collection of the CSAT DV are provided in Section 3.9.3.

The second DV in the study (DEFECT, measuring defection/loyalty) is measured throughout the study, as it is an observed measure and is recorded in the study organization's data system as defection events occur. By definition, there is no opportunity for a pretest of defection: if the participant had deactivated their mobile service prior to the start of the study, they would not have been able to have been included in the study at all. Section 3.9.2 provides details on the definition and measurement of this DV.

In place of a pretest for the DEFECT DV, attention will be given to the design, execution and validation of the sampling strategy (Section 3.8) to ensure equivalence of the treatment groups prior to the start of the experiment.

3.6.13 Field and Natural Experiment Design

The study takes the form of a field experiment, as opposed to the alternative of laboratory-based study design. This design supports the ability to test the hypotheses from Section 3.4 through the use of the three treatments described previously (Salkind 2010). The selection and implementation of the treatments overcome a major constraint in field experimentation, namely the ability to alter specific variables through experimental treatments (Eden 2017).

Part of the decision to utilize a field experiment method is practical: the design of the study requires the implementation of the treatments and measures over a period of 12 months (see Section 3.10 for a discussion of the study timeframe). It would be impractical (if not impossible) to have implemented the desired study design within the controlled environment of a laboratory for such an extended period.

There are additional advantages to the field approach that were incorporated into the study (Maner 2016). These advantages include:

1. Increased replicability
2. Enhanced impact

The increased replicability provided by the field-based study design derives from five advantages, as outlined in Table 3.4.

A final aspect of field research that has been applied to the study design is that of a three-dimensional naturalistic approach, as espoused by Tunnell (1977). According to Tunnell: “The most convincing studies have incorporated high levels of control with as much naturalness as possible” (1977, p. 434), and proceeds to outline two advantages of designing a field experiment around the three-dimensional naturalistic approach:

1. Possibility of the discovery of new empirical laws, through the operationalization of the DVs, IVs, and background setting in the real world; and
2. Increase in external validity, as the natural research setting may increase the generalizability of the results

The three dimensions proposed by Tunnell and their method of implementation in this study are summarized in Table 3.5.

.

Table 3.4 Replicability Advantages Provided by Field Design

Replicability Issue	Feature of Field Research	Implementation in this Study	Advantages
Participant diversity	Field studies often employ samples that are highly diverse	Fully randomized sample groups from a large frame of the total population (see Section 3.8 for details of the sampling approach used in the study)	Findings derived from diverse samples are more robust and may be more likely to be replicated across other samples than those derived from homogenous samples
Presence of extraneous variables	Field studies usually entail less control over extraneous sources of variance	No attempt made to control extraneous variables; all study groups (control and treatment) are subject to the same extraneous variables on a random basis	Effects demonstrated under uncontrolled circumstances should be more robust to contextual factors than those studies in which such factors are held constant
Measurement of behaviour	Field studies often focus on directly measuring behavioural dependent variables	The study will measure two DVs (customer satisfaction and defection)	Behavioural DVs in field studies are relatively robust to variables that might moderate the intention-behaviour gap, and this may be more replicable (and impactful)
Researcher degrees of freedom	Field studies sometimes focus on fewer DVs; entail less daily control over data collection	No researcher control over the collection of defection data as this occurs automatically when a study participant leaves the subject organization; customer satisfaction data collection is routinized and executed by a third party	Field studies may leave relatively less room for the researcher to exploit inclusion or exclusion of data that do not adhere to the researcher's hypotheses
Publication bias	Field studies tend to be high-investment, less likely to be relegated to the file drawer	Strong interest and desire to incorporate study results into the subject organization's ongoing strategy and programs; study involves resources across various teams of the subject organization	Field studies may be less likely than other types of research to suffer from publication bias

Adapted from Maner (2016)

Table 3.5 Definition and Application of Naturalistic Dimensions in Field Research

Dimension	Defining Criteria	Implementation in this Study
Natural behaviour	Behaviour is part of the participants' current repertoire of natural responses	The study participant's response to factors impacting customer satisfaction, and their decision to leave the company or stay, is one they already make today and would continue to do so absent the experiment
Natural setting	Research participants exist in their natural environment (outside the lab)	The study participants continue to function in their natural environment and are not brought into any artificial setting to administer the treatments or measure impacts to the DVs
Natural treatment	Treatment is a natural, discrete event that might have occurred outside the study	Each of the three treatments utilized in the study exists naturally outside the experiment

Adapted from Tunnell (1977)

3.6.14 Longitudinal Design

The study takes the form of a longitudinal design, with a total in-field duration of 12 months. The importance of the longitudinal approach to the study lies in the nature of the two dependent variables of interest: customer satisfaction (CSAT) and defection (DEFECT).

Prior research suggests that the impact of customer satisfaction antecedents may take place over extended periods of time (Bolton et al. 1991, Kumar et al. 2013, Chen et al. 2014). Experience within the subject organization has shown that changes to customer satisfaction can take many months or quarters (and in some cases, years). With little prior research on the timing of the impact of the treatment effects employed in this study on the CSAT DV, the 12-month period used provides a longer timeframe from which to anticipate a measurable effect.

The DEFECT DV is an outcome of the customer deciding to cease receiving service from the mobile phone provider. Mobile phone services are differentiated from many types of products and services used by consumers in that there is a strong need by most customers to ensure continuity of service, even when they have made a decision to leave a given carrier. Failure to maintain this continuity results in customers losing their assigned mobile phone number, or the ability to contact or be contacted via voice or text messaging or utilize any Internet-enabled applications.

Few mobile subscribers exit the market (that is, decide they no longer require any mobile service from any provider). Instead, customers will determine that they want to move their service to another provider and will seek to transition their service to that new provider as seamlessly as possible. Such transitions are enabled by government-mandated support for number portability (described previously in Section 1.1.1). In most cases where a customer transfers service to another competitor, the company losing the customer only finds out at the time the new service is being activated at the new carrier, and the old carrier must facilitate the transfer of service in a mandated period of time (by regulation, this is 2.5 hours or less[14]; in practice the number transfer is almost instantaneous).

Although the ability to maintain continuity of service through this transfer process exists, the decision to make this move is a result of a fair amount of work on behalf of the customer. The customer must decide which of several carriers and brands they want to move to, whether they want to take their existing mobile device with them or purchase a new one, and which rate plan and features they wish to subscribe to.

Customers may also face barriers to leaving their current provider, such as the existence of a balance owed from the financing of a mobile phone device.

These barriers imply that the action to leave a given carrier may not be coterminous with a change in customer satisfaction level, even after an event that might substantially impair their level of satisfaction. As a result, the longitudinal approach to this study is appropriate to provide sufficient time for the treatments to have an effect (or not) on the DVs.

One final consideration in the choice of the 12-month longitudinal study design is the cyclical nature of the mobile phone market. Throughout the course of a year there are several cyclical events that may act as extraneous influences on customer behaviour throughout the study. In particular, competitive activities of the carriers in the market can vary quite widely by time of year. Examples of this include the launch of new major mobile phone devices (such as the Apple iPhone, which launches in September of each year) or key sales periods such Back-to-School (August and September), Black Friday (November) and Christmas/Boxing Day (December).

The study was deployed in the field for a full 12 months to increase the validity of the study and the ability to generalize the results across the full calendar year of business operations. The 12-month timeframe allowed for a full cycle of the external events to take place that might impact the DVs.

---

[14] CWTA. https://www.cwta.ca/for-consumers/wireless-number-portability/

Consideration was given to the potential for attrition of participants over the 12 month study period in determining the sample size required, as described in Section 3.8 (Salkind 2010).

3.7 Research Validity and Control Techniques

An objective in the design of this study is the achievement of high degrees of validity for the inferences that may be drawn from the outcomes, implying a greater approximation of the truth (Shadish et al. 2002) or correctness (Maxwell et al. 2017) of those inferences. In this manner, it is hoped that the claims of causal relationships as a result of the study may be claimed as "approximately true" (Shadish et al. 2002).

To achieve this result, the study design includes specific strategies for achieving valid results (Christensen et al. 2015), and as recommended by Shadish et al. (2002, p. 40) as a way to "…minimize the number and plausibility of those threats that remain by the end of the study.". Christensen et al. suggest that there are four types of validity required for a quantitative study to achieve accurate inferences, including:

- Statistical Conclusion Validity
- Construct Validity
- Internal Validity
- External Validity

The following sub-sections outline the strategies that were employed that are appropriate for the experimental research design.

3.7.1 Statistical Conclusion Validity

This form of validity is determined by the ability to show a statistically significant relationship between the independent (IV) and dependent (DV) variables.

This form of validity relates to two types of statistical inferences that are expected to be evaluated in the analysis phase of the study (Shadish et al. 2002):

a. Whether the presumed cause and effect covary – a Type I error if it is erroneously concluded that the cause and effect covary when they do not, or a Type II error if it is concluded that they do not covary when they do
b. How strongly cause and effect covary – an over- or under-estimation of the magnitude of covariation, as well as the degree of confidence

The research design collects the data necessary to determine if such a relationship exists. No additional explicit strategies are employed over and above the study design itself to address this validity.

3.7.2 Construct Validity

Threats to construct validity (i.e. the degree to which the measures in the study adequately represent the construct, the example of which in this study is the use of the likelihood to recommend (L2R) measure to represent customer satisfaction) that are relevant to this study's design are summarized in Table 3.6, along with the corresponding strategy to mitigate that threat in the study.

Table 3.6 Potential Threats to Construct Validity

Threat	Mitigation Strategy
Participant reaction to the experimental situation	Study participants are not informed they are participating in the study
Compensatory Equalization	The three treatments employed in the study cannot be assigned to other study participants. The VALUE, CSERV and CONTACT treatments are controlled through systems that the study participants cannot influence and cannot accidentally be applied to participants not receiving the assigned treatment via controls inside the study organization. Changes in the environment outside the study, such as changes in the pricing of services, queue times in the contact centre or the ability of a participant to opt-out of marketing messages are equally available to all participants regardless of their treatment status.
Compensatory Rivalry	Participants are not aware they are participating in the study and are not aware that there are treatments other than the one(s) they are receiving. Of the three treatments employed, only the VALUE treatment is overt and known to the study participant (as detailed in Section 3.6.4). The sampling strategy employed (Section 3.8) ensures that no two mobile subscribers are selected for the study from the same household, reducing the likelihood that any one participant may be aware of the VALUE treatment received by another participant.

Treatment Diffusion	All treatments were implemented through the programming of various systems within the subject organization, based on the assigned treatment groups. Once participants were assigned to their treatments, the system changes were locked to prevent any changes of participants into or out of a given treatment.
Experimenter Effects	There is no direct interaction between the researcher and the participants; data for the measurement of the DVs are recorded by third-party sources (as in the case of the CSAT DV) or through automated reporting systems (such as with the DEFECT DV).

Adapted from Christensen et al. (2015)

3.7.3 Internal Validity

Internal validity is perhaps the most important aspect of validity to be addressed in this research design, as it refers to the validity of the inference(s) that the IVs and DVs in the study are causally related (Christensen et al. 2015). The ability to support statements of causal relationships is a key objective of this study.

While Duflo et al. (2017) state that "…well-executed randomized experiments by definition have internal validity, and that the problem of internal validity is one that plagues only observational studies or compromised random experiments", the study design nonetheless strives to ensure that any changes observed in the DVs were in fact caused by changes in the IVs, and not as a result of any extraneous variables (known or unknown).

The research design aspect most closely associated with ensuring internal validity is randomization in both the selection of the sample and the assignment of participants within the sample to the treatment groups (Christensen et al. 2015).

The use of randomization in this manner provides internal validity by the fact the existence of any such extraneous variables is approximately equally distributed amongst all participants, thereby ensuring that the influence of these extraneous variables is held constant and there is no differential influence on the DVs (Maxwell et al. 2017).

Christensen et al. also identify several threats to internal validity that are mitigated by the random sampling strategy of the study. These threats include:

- History – events that occur after the start of the study prior to the posttest measurement
- Maturation – changes in the state of the study participants as a function of time
- Testing – changes in measures of a DV (such as CSAT in this study) as a result of the previous measurement

- Attrition – loss of participants through the study timeframe that produces differences in the groups other than as a result of the treatments
- Selection – where participants are not randomly assigned to treatment groups

The rigour of the randomized sampling strategy employed in this study, and detailed in Section 3.8, is explicitly aimed at achieving the highest degree of internal validity possible.

3.7.4 External Validity

Of importance to the subject organization, and for the potential contribution of this study to the overall body of knowledge, is the degree of external validity that can be achieved in this study. External validity concerns the ability to extend causal inferences from the study to other persons, settings and part or future situations (Shadish et al. 2002, Creswell 2014).

The ability to generalize the findings across the target population of the subject organization is of critical importance for operationalizing the findings and obtaining financial value from the research. The ability to conduct random sampling for study participants, while "…rarely feasible in experiments" simplifies the external validity inferences (Shadish et al. 2002). The study similarly strives for relevance with the broader research community.

Christensen et al. (2015) proposes three main types of external validity, which the design of this study seeks to maximize. Each type is reviewed in Table 3.7, with details of the strategy employed in the study to maximize each.

Table 3.7 External Validity Types Addressed in Research Design

External Validity Type	Study Maximization Strategy
Population Validity	The sampling strategy of the study (Section 3.8) defines the sampling frame as broadly as possible in order to represent the target population. The target population of 3.6 million represents a large percentage of total consumers in Canada, and data can be obtained from other sources to evaluate the equivalency to the mobile service customers of other carriers. The total number of mobile subscribers is 33.2 million[15], which represents 89% of the population of Canada of 37.2 million[16]. Demographic and psychographic data can be obtained for each study participant to utilize in generalizing results across or to broader populations.
Temporal Validity	Study design uses a 12-month experimental period to account for seasonal variations and other cyclical variations such as competitive activity.
Ecological Validity	The naturalistic field experiment design of the study provides for the ability to generalize the results as it is not constrained by a laboratory environment. Randomized selection ensures a variety of environments (such as rural or urban) are represented in the study sample.

Adapted from Christensen et al. (2015)

3.7.5 Validity Summary

Christensen et al. (2015) suggests that the type of validity that is most important (internal or external) is a function of the primary purpose of a given study. Kazdin (as cited by Christensen et al. 2015, p. 200) suggests:

> "…it would seem logical to design experiments using a diverse sample of research participants, treatment variations, outcome measures, and settings across several different time periods in order to increase external validity. The problem with this strategy is that there tends to be an inverse relationship between internal and external validity. When external validity is increased, internal validity tends to be sacrificed; when internal validity is increased, external validity tends to suffer."

[15] CRTC. Communications Monitoring Report (2019)
[16] Statistics Canada. Table 17-10-0009-01 Population estimates, quarterly

The research design employed in this study attempts to strike a balance between the suggested trade-offs of internal versus external validity. It does this by including a randomly selected sample group from a large target population, situated in a natural field setting and studied across a lengthy timeframe (all contributing to external validity), while simultaneously enhancing internal validity through the same randomization strategy.

3.8 Sampling Strategy

One of the most critical aspects of a randomized controlled trial experiment is the sampling strategy for study participants, with the random assignment of participants to treatment groups being "…the backbone of experimental inference" (Kenny 1979, p. 1). Done properly, the researcher may make strong claims on the causal relationship(s) between the variables under study. Done poorly, and the study may suffer from unintended (and unknown) extraneous influences and be limited in the ability to generalize from its findings.

Randomization, as employed in the sampling strategy defined here, is "…a blind balancing act that controls infinite potential confounders – including those unknown to the experimenter…" and "…imbues the experiment with supreme internal validity and makes it the gold standard." When combined with a field experiment design, this allows for control over both internal and external validity, making the field experiment "sterling gold" (Eden 2017, p. 96).

Kothari (2004) suggests five characteristics of a good sample design, which the sampling design for this study attempts to emulate:

1. Sample design must result in a truly representative sample
2. Sample design must be such which results in a small sampling error
3. Sample design must be viable in the context of funds available for the study
4. Sample design must be such so that systemic bias can be controlled
5. The sample should be such that the results of the sample study can be applied, in general, for the universe with a reasonable level of confidence

This section details the sampling strategy employed in the study, made possible by the access of the researcher to the mobile subscriber base of a major Canadian wireless carrier ("subject organization"). The target population for the study is defined, and the methods and procedures used to create the sample population and randomly assign participants to the appropriate treatment and control groups are reviewed. Finally, an analysis of the target population, sample frame, sample group and treatment group assignment is completed to validate the implementation of the randomization strategy.

3.8.1 Target Population

Within the mobile wireless industry, there are several different types of subscribers, defined by a number of attributes of the service they subscribe to (detailed in Chapter 1). For the purposes of this study, the largest single group of subscribers within the subject organization has been chosen. This group is defined by the attributes in Table 3.8.

Table 3.8 Target Population Definition

Attribute	Target Population Defined By	Description
Brand	Primary Brand	Most carriers in Canada operate multiple brands. Primary brands include Bell, Rogers and TELUS; flanker brands include Fido, Koodo and Virgin Mobile
Customer Segment	Consumer	Two main segments of customers include Consumer and Business
Billing Method	Postpaid	Two types of billing methods exist: Postpaid and Prepaid. The majority of Consumer, and all Business customers, utilize a postpaid service.

At the time the sampling activity was completed in February 2019, the target population (as defined in Table 3.8) for the subject organization was 3,574,893 mobile service subscribers. This represented 42% of the total mobile subscriber base of the subject organization (for all brands, customer segments and billing methods), and approximately 11% of the total mobile subscriber base in Canada.

3.8.2 Sampling Frame

From the target population, the sampling frame has been defined to take into account several characteristics of the population that would make some mobile subscribers less relevant for the study or to address issues related to ensuring internal validity (as described in Section 3.7.3).

The design attempts to ensure that the sampling frame is as representative of the target population as possible, and only removes a potential study participant where there is a justifiable reason for exclusion. Each of the criteria for defining an excluded participant for the purpose of this study is described in detail below and consists of two types of exclusions based on 1) contract status and 2) multiple subscriber households.

As outlined in Section 3.6.14, the study will take place over a twelve-month period. The DEFECT DV is based on participants choosing to leave the subject organization (or not)

during the study period, requiring consideration to be given to the inclusion of mobile subscribers whose contract status might otherwise deter their defection, regardless of their levels of customer satisfaction or other factors.

Each customer electing a subscribe to postpaid mobile phone service in Canada must make a decision at the time of initiating service whether they wish to purchase a new mobile phone handset, and if so, whether they will choose to pay a lower price upfront by having the carrier subsidize or finance some of the cost of this device. The decision to take a subsidy or financing offer creates a device balance (equal to the value of subsidy or financing received) that depreciates in a straight line over (usually) 24 months. In essence, this creates a form of contract between the customer and the carrier, as a customer who wishes to deactivate their service prior to the device balance reaching $0 must pay off the balance immediately and in full. This contract relationship creates a situation where subscribers may be temporarily restrained from deactivating their service, in spite of low levels of customer satisfaction or attractive competitive offers.

Subscribers who remain with a carrier at the conclusion of their contract (i.e. they have a $0 subsidy or financing balance) are deemed to become month-to-month (MTM) subscribers. This same status also applies to new subscribers who sign up for new service with the carrier but bring their own mobile devices or pay the full up-front cost of the device (also referred to as BYOD subscribers). In these situations, subscribers are free to move to a competitor at any time without a financial cost.

Including subscribers in the sample frame whose contract spans the study timeframe would not create a flaw in the experiment design, but it would have the potential to artificially reduce the treatment effects on the two DVs, and in particular, the DEFECT DV.

As a result, the sampling frame employed in this study excludes any customers who have greater than 12 months remaining on their contract as of the start of the study. The sampling frame will, therefore, only include customers who are either:

- On contract at the time of study commencement, but whose contract will end within the study timeframe (i.e. in 12 months or less)
- MTM and BYOD subscribers at the time of study commencement, but who may become re-contracted during the study timeframe (as a result of a service renewal transaction)

The percentage of the target population excluded from the sample frame as a result of these criteria is estimated at 38%[17].

In addition to excluding customers on contracts, the population frame also needs to take into account households where there is more than one subscriber on a given account for that household. For example, there might be several mobile phone services for two parents as well as one or more children of the household.

The challenge presented by the multiple subscriber household is two-fold:

1. Which of the subscribers on the account might be considered the "owner" of the account, and therefore might be expected to be the primary decision maker? The study would only want to target this primary subscriber for the treatments and DV measures, as this is where the effect of the treatments, if any, would manifest itself. This ownership within a multiple subscriber household is not cleanly defined in the subject organization's subscriber database.
2. The existence of multiple subscribers on a single account in a household creates a situation where all of those subscribers in the household would be included in the sample frame, and therefore might i) be included in the sample population, and ii) be randomly assigned to different treatment groups. This would create a challenge to construct validity (specifically compensatory rivalry as reviewed in Section 3.7.2), where one study participant might receive the VALUE treatment and another participant in the same household might not.

As a result of the two challenges outlined above, the sampling frame for this study only includes single subscriber accounts/households. Including this criterion in the definition of the sampling frame ensures that the study participants fulfil the definition of "customer", as defined in Section 2.1.1. The percentage of the target population excluded from the sample frame as a result of these criteria is estimated at 24%[18].

As a result of the exclusions from the target population described above, the sampling frame for the study therefore consists of a subset of the target population. The two exclusions are not mutually exclusive, as an account can both be under a contract greater than 12 months and have multiple subscribers on the account at the same time. The two criteria were applied to the target population of 3,574,893, resulting in a sampling frame for the study of 930,930

[17] 2016 average percentage of the target population on contract with >12 months remaining, based on subject organization internal reporting
[18] Based on analysis of data retrieved for a pilot study in March 2017

subscribers. This sampling frame represents 26% of the target population, 11% of total subscribers of the subject organization and 2.9% of all mobile subscribers in Canada.

3.8.3 Study Sampling Technique

In support of the overall design of the study (which requires a random sample of participants), the study sample will be drawn from the sample frame using an Equal Probability Selection Method (EPSM). This method ensures that the sample is representative of the population under study (Christensen et al. 2015).

Within EPSM, Christensen outlines four methods of sampling. The "…definitive case of an (EPSM)…" (p. 165) is the simple random sampling method. This is the method employed in this study, both to create the overall sample for the study from the sampling frame, and to assign participants to the various treatment and control groups of the experiment.

Simple random sampling is an unrestricted sampling method, meaning that it ensures a completely random and unpredictable allocation of participants, and is, therefore "…superior to the other methods in terms of bias prevention." (Schulz and Grimes, as cited in Arifin 2012, p. 131).

Alternative consideration was given to the use of stratified random sampling; however, the research design did not identify any stratification variables that would dictate the need to consider this sampling approach. In addition, the efficiency advantage attributed to the stratified sampling method (i.e. that of requiring fewer participants) was not a constraint in this study.

3.8.4 Sample Size Determination

Given the lack of constraints on determining the sample size for this study (with a population frame of almost one million), the study design considered the creation of the sample population as large as is feasible given other constraints (for example, the cost of providing the discount associated with the VALUE treatment). With a large enough sample population, the stratification of results post-study completion will be possible.

As stated previously, the study benefited from a general lack of constraints on the sample size to support the experimental study. This allows for a determination of the sample size to be focused primarily on the minimum size required to support the research analysis at the desired confidence levels. The upper limit to sample size was set at the point where further increases in size would not contribute to any additional measurement value.

Three factors were considered in the determination of the sample size for the study:

1. Response rates to the survey measurement of the CSAT DV
2. Attrition of participants during the study timeframe due to defection
3. Number of participants expected to exhibit the defect behaviour

The first factor to be considered was the ability to obtain survey measures for the CSAT DV. With a response rate expected to be less than 100% (as with all surveys), the sample size must be such that sufficient valid responses are received to meet the criteria for statistical validity.

The study participants that do not respond to the customer satisfaction survey will still be measured for the DEFECT DV (which does not require any response by the study participants), and therefore the ability to claim a causal relationship between the treatments and the DEFECT DV would remain intact. However, without a valid measure of the CSAT DV for a given treatment group, it would be unknown if the intermediating CSAT DV was involved in the causal relationship.

Previous customers satisfaction research by the subject organization on the same group of mobile subscribers as the target population for this study has historically experienced a 10% response rate to surveys.

The second factor to be considered in determining the appropriate sample size is attrition. This factor is directly related to the DEFECT DV, so a given level of participants leaving the study as a result of defection is expected. An expected rate of attrition is included in the sample size design to ensure that the remaining participant sample does not drop below the level required to satisfy the study design.

The target population identified in Section 3.8.1 exhibits a rate of defection of slightly less than 1% per month[19]. At this rate, the study sample can be expected to shrink by approximately 11% over the 12-month study period.

The third factor to be considered is that of the desired sample size of study participants expected to exhibit the DEFECT DV measure. The ability to determine statistically valid analyses on the study participants who defect (and therefore are measured by the DEFECT DV) is a critical outcome of the study. The sample size required for the study must, therefore be large enough that the estimated number of participants to defect over the 12-month study will be equal to or greater than the statistically required sample size.

[19] Based on subject organization results, average monthly defections for the target population in 2016 was 0.83%

With these three factors in mind, the next step is an examination of several common guidelines used in sample size determination. Christensen et al. (2015) identify the following relevant guidelines:

1. Obtain relatively large sample sizes when feasible
2. Examine other research studies
3. Use a sample size table
4. Use a sample size calculator, such as one based on the Power method

Constraints on sample size for this study are few, with the exception of consideration for the cost of the VALUE treatment described in Section 3.6.5. There are few studies utilizing a similar research model (as discussed in the literature review section); therefore, the determination of the appropriate sample size looks to other approaches.

A commonly used method of determining sample size is to utilize a pre-calculated table, such as that shown in Table 3.9.

Table 3.9 Suggested Sample Sizes for Populations up to 50 Million

N	n	N	n	N	n	N	n	N	n
10	10	130	97	250	152	950	274	10,000	370
20	19	140	103	260	155	1,000	278	20,000	377
30	28	150	108	270	156	1,100	285	30,000	379
40	36	160	113	280	162	1,200	291	40,000	380
50	44	170	118	290	165	1,300	297	50,000	381
60	52	180	123	300	169	1,400	302	75,000	382
70	59	190	127	400	196	1,500	306	100,000	384
80	66	200	132	500	217	2,000	322	250,000	384
90	73	210	136	600	234	3,000	341	500,000	384
100	80	220	140	700	248	3,500	346	**1,000,000**	**384**
110	86	230	144	800	260	4,000	351	10,000,000	384
120	92	240	148	900	269	8,000	357	50,000,000	384

N represents the size of the population, *n* stands for the recommended sample size. Sample sizes based on 95% confidence level.

From Christensen et al. (2015)

Using the sampling frame population of 930,930 determined in Section 3.8.2, the recommended sample size based on a 95% confidence level is 384. This is the sample size that would be required for each of the treatment groups, as well as for the control group. The sample size chart shows that there is no further increase in sample size required as the population increases beyond 100,000, suggesting 384 as the upper limit for valid sample sizes.

The final determination of the sample size of participants required for the study is now a combination of the sample size table recommendation and the accommodation of the three factors identified at the beginning of this section. The sample size requirement is considered for each of the three factors individually, with the largest required sample size identified being the one to be used for the study.

In considering the response rates to the customer satisfaction survey (that is, the expected 10% response rate), the sample size for each participant group in the study would need to be at least 3,840 (allowing for the estimated 384 survey responses to be received).

From an attrition perspective, a monthly defection rate of 0.83% would require an initial starting sample size of 424 to have at least 384 participants remaining after 12 months, as shown in Table 3.10.

Table 3.10 Sample Size Requirement - Attrition Consideration

	Month											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Starting Sample Size	**424**	420	417	414	410	407	403	400	397	393	390	387
Monthly Deactivations	4	3	3	3	3	3	3	3	3	3	3	3
Ending Sample Size	420	417	414	410	407	403	400	397	393	390	387	**384**

Note: Monthly deactivations calculated based on 0.83% monthly churn rate

With a requirement for at least 384 defection observations from each treatment group, the original sample size for each treatment group is 4,100 as shown in Table 3.11 (a starting sample of 4,100 participants is expected to yield approximately 390 defections over the course of the study).

Table 3.11 Sample Size Requirement - Deactivation Volume Consideration

	Month												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Total**
Starting Sample Size	**4100**	4066	4032	3999	3966	3933	3900	3868	3836	3804	3772	3741	
Monthly Deactivations	34	34	33	33	33	33	32	32	32	32	31	31	**390**
Ending Sample Size	4066	4032	3999	3966	3933	3900	3868	3836	3804	3772	3741	3710	

Note: Monthly deactivations calculated based on 0.83% monthly churn rate

The ending sample size of 3,710 falls slightly below the sample size required at the end of the study to meet the requirements for the customer satisfaction survey (identified as 3,840). A starting sample size of 4,244 for each treatment group meets the needs of all three factors to be considered (Table 3.12)

Table 3.12 Sample Size Requirement - All Factors Considered

	Month												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Total**
Starting Sample Size	**4244**	4209	4174	4139	4105	4071	4037	4003	3970	3937	3905	3872	
Monthly Deactivations	35	35	35	34	34	34	34	33	33	33	32	32	**404**
Ending Sample Size	4209	4174	4139	4105	4071	4037	4003	3970	3937	3905	3872	**3840**	

Note: Monthly deactivations calculated based on 0.83% monthly churn rate

The factorial design described in Section 3.6.9 identifies a total of eight treatment groups. A starting sample size of 4,244 for each group yields a starting sample group for the overall study of 33,952. This was rounded up to 35,000 in the execution of the sampling strategy and detailed in Section 3.8.5.

3.8.5 Sampling Execution and Validation

The experimental nature of this study is highly dependent on the diligent execution of the various activities involved in the research design. In particular, the randomized controlled trial aspect of the experiment demands a high degree of attention to the execution of the sampling strategy. Errors in sampling may result in unintended (and possibly unknown) biases to be introduced to the study, and in a worst-case scenario, may render the outcome of the study as unusable. This would be an unwelcome outcome, especially for a study such as this that takes place over more than a year and involves substantial financial costs for the subject organization.

As a result, prior to the commencement of the study, the sampling strategy was carefully executed and an analysis was completed to ensure that the goals of the sampling strategy were achieved. The outcome of this analysis proved satisfactory, and the experimental study proceeded.

3.8.5.1 Sampling Execution

Data files were obtained from the subject organization's subscriber billing system for the target population and the sampling frame, as defined in Sections 3.8.1 and 3.8.2, respectively. Descriptive data fields were obtained to allow for analysis to confirm the randomness of the sampling techniques. These data fields can also be used for post-study stratification analysis.

The data elements shown in Table 3.13 were collected for both the target population and the sampling frame.

Table 3.13 Sampling Validation Data Elements

Data Element	Description
Tenure	Number of days the customer has had service with the provider
Language	English (EN) or French (FR)
Billing Cycle Date	One of 17 calendar days in the month that customer's bill is created
Credit Class Code	One of 12 credit classes assigned at time of service activation
Contract Status	Indicates if the customer is under contract or not
Days Left on Contract	If customer on contract, how many days left

Billing Province	One of 10 Canadian provinces or three territories (or other)
Rural or Urban	Indicates if customer's postal code is designated as Rural or Urban
Revenue Band	One of 9 bands that represent the revenue level of the customer
Do Not Survey Status	Indicates if the customer does not wish to receive surveys
Do Not Contact Status	Indicates if the customer does not wish to receive marketing contacts

3.8.5.2 Validation of Sampling Execution – Target Population and Sampling Frame

Validation of the creation of the sample frame from the target population was completed by the analysis of the data elements summarized in Section 3.8.5.1. Differences in the results across the various elements are expected as a result of the execution of the sampling frame criteria defined in Section 3.8.2. Results of the analysis for each data element are shown in Table 3.14.

Application of the sampling frame criteria is reflected in three of the data elements. These three elements were reviewed to ensure the results match expectations of differences between the target population and the sampling frame based on the latter's criteria.

The sampling frame criteria included exclusions from the target population based on 1) contract status and 2) multiple subscriber households.

The contract status is defined as those subscribers in the target population on contract with 12 or fewer months remaining on that contract, or subscribers that are currently not on contract. The data element Days Left on Contract reflects the difference between the target population (where the range is 1 to 732 days, consistent with the maximum contract length for 99.7% of the target population of 24 months) and the sampling frame (where the range is 1 to 365 days, reflecting the criteria to exclude any subscribers with more than 12 months remaining on their contract).

As a result of excluding subscribers with more than 12 months remaining on their contract, the values for Contract Status (that is, the percentage of subscribers on contract versus not on contract) is substantially different between the target population (61.9% on contract) and the sampling frame (37.5% on contract). This is consistent with the application of the sampling frame criteria.

The second sampling frame criteria consists of excluding any households with more than one subscriber on their account. Of the target population of 3,574,893, 60.8% of those subscribers are on an account where they are the only subscriber. The remaining 39.2% of the subscribers are on accounts where there is at least one other subscriber as well. The sampling frame shows that the accounts with multiple subscribers have been excluded, as 100% of the subscribers in this sample are on accounts with only one subscriber.

Examination of the values across the remaining data elements shows that the distributions within the target population and sampling frame are consistent, with only small variations as would be expected following the application of the sampling frame criteria.

Table 3.14 Data Element Comparison - Target Population versus Sampling Frame

Element	Value	Target Population	Sampling Frame
Number Records		3,574,893	930,930
Unique Accounts		60.8%	100%
Tenure (days)	Minimum	13	13
	Maximum	17,575	17,575
	Mean	3,180	2,818
	St. Dev.	2,329	2,210
Language	English	88.8%	85.3%
	French	11.2%	14.7%
Billing Cycle Date	1	6.0%	6.2%
	2	5.7%	5.8%
	6	5.3%	5.4%
	7	6.1%	5.7%
	8	5.4%	5.8%
	10	6.0%	5.8%
	11	5.7%	5.7%
	12	5.4%	5.3%
	14	5.9%	5.7%
	15	5.7%	5.9%
	17	6.4%	6.2%
	18	5.1%	5.3%
	22	7.1%	6.9%
	25	6.4%	6.5%
	26	5.6%	6.0%

Element	Value	Target Population	Sampling Frame
	28	6.2%	6.0%
	29	5.6%	5.6%
Credit Class Code	B	85.3%	86.0%
	C	3.0%	3.6%
	D	10.7%	8.7%
	L	0.6%	0.5%
	V	0.2%	0.3%
	X	0.2%	0.9%
Contract Status	On	61.9%	37.5%
	Off	38.1%	62.5%
Contract Duration	24 month	99.7%	99.0%
	Other	0.3%	1.0%
Days Left on Contract	Minimum	1	1
	Maximum	732	365
	Mean	395	206
	St. Dev.	203	99
Billing Province	AB	31.8%	27.3%
	BC	23.5%	24.0%
	MB	4.2%	4.7%
	NB	1.2%	1.2%
	NL	1.6%	1.2%
	NS	2.2%	2.0%
	NT	0.1%	0.1%
	NU	<0.0%	<0.0%
	ON	19.9%	20.6%
	PE	0.3%	0.3%
	PQ	12.4%	15.8%
	SK	2.7%	2.3%
	YT	0.1%	0.1%
Rural or Urban	Rural	22.9%	21.1%
	Urban	77.1%	78.9%
Revenue Band	A	5.5%	4.4%
	B	28.5%	29.5%
	C	25.2%	28.2%
	D	17.3%	17.9%
	E	6.5%	9.3%
	F	1.6%	2.0%

Element	Value	Target Population	Sampling Frame
	N	3.7%	1.7%
	S	10.8%	6.5%
	Other	0.9%	0.5%
Do Not Survey	Yes	2.0%	1.8%
	No	98.0%	98.2%
Do Not Call	Yes	0.6%	0.5%
	No	99.4%	99.5%

Note: Values for data elements may not total 100% due to rounding

3.8.5.3 Validation of Sampling Execution – Sample Group

Upon validation that the sampling frame was representative of the target population (based on the definition of the target frame), the sample group was then obtained from the sample frame, using the simple randomization technique. The purpose was to obtain a sample group for the study, based on the sample size determined in Section 3.8.4 (that is, 35,000), that would be representative of the sample frame.

The sample group was then subjected to the same analysis of the data elements to ensure the sample group was representative (Table 3.15). In this case, it is important that the values for each data element be consistent between the sampling frame and the sample group – there are no expected material differences between any of the data elements to be examined.

Table 3.15 Data Element Comparison - Sampling Frame versus Sample Group

Element	Value	Sampling Frame	Sample Group
Number Records		930,930	35,000
Unique Accounts		100%	100%
Tenure	Minimum	13	13
(days)	Maximum	17,575[20]	11,998

[20] Target population and sampling frame datasets include two subscribers with unexpectedly high values for tenure; maximum tenure should be approximately 13,000 days based on launch of the subject organization in 1984. There is no known explanation for the anomalous data, which is not expected to impact the sampling strategy or the study overall.

Element	Value	Sampling Frame	Sample Group
	Mean	2,818	2,805
	St. Dev.	2,210	2,200
Language	English	85.3%	85.2%
	French	14.7%	14.7%
Billing Cycle Date	1	6.2%	6.4%
	2	5.8%	5.9%
	6	5.4%	5.6%
	7	5.7%	5.8%
	8	5.8%	5.6%
	10	5.8%	5.6%
	11	5.7%	5.8%
	12	5.3%	5.2%
	14	5.7%	5.7%
	15	5.9%	5.8%
	17	6.2%	6.1%
	18	5.3%	5.2%
	22	6.9%	7.2%
	25	6.5%	6.6%
	26	6.0%	5.9%
	28	6.0%	5.9%
	28	5.6%	5.6%
Credit Class Code	B	86.0%	86.2%
	C	3.6%	3.6%
	D	8.7%	8.6%
	L	0.5%	0.5%
	V	0.3%	0.3%
	X	0.9%	0.9%
Contract Status	On	37.5%	37.5%
	Off	62.5%	62.5%
Contract Duration	24 month	99.0%	98.9%
	Other	1.0%	1.1%
Days Left on Contract	Minimum	1	1
	Maximum	365	365
	Mean	206.05	204.40
	St. Dev.	99.058	99.797
Billing Province	AB	27.3%	27.4%
	BC	24.0%	23.9%

Element	Value	Sampling Frame	Sample Group
	MB	4.7%	4.8%
	NB	1.2%	1.2%
	NL	1.2%	1.2%
	NS	2.0%	2.0%
	NT	0.1%	0.1%
	NU	<0.0%	<0.0%
	ON	20.6%	20.7%
	PE	0.3%	0.3%
	PQ	15.8%	15.%
	SK	2.3%	2.4%
	YT	0.1%	0.1%
Rural or Urban	Rural	21.1%	21.4%
	Urban	78.9%	78.6%
Revenue Band	A	4.4%	4.4%
	B	29.5%	29.5%
	C	28.2%	28.6%
	D	17.9%	17.5%
	E	9.3%	9.1%
	F	2.0%	2.1%
	N	1.7%	1.6%
	S	6.5%	6.7%
	Other	0.5%	0.5%
Do Not Survey	Yes	1.8%	1.9%
	No	98.2%	98.1%
Do Not Call	Yes	0.5%	0.5%
	No	99.5%	99.5%

3.8.5.4 Validation of Sampling Execution – Treatment Assignment

As the sample group was determined to be a fair representation of the sampling frame, the sample participants were then subsequently randomly assigned one of the eight treatments as determined by the factorization requirements of Section 3.6.9. Once again, comparisons were made across the descriptive elements between each of the resulting eight treatment groups, as well as the original sample group (Table 3.16).

Analysis of the descriptive values for the various data elements across all eight treatment groups revealed no anomalies, and the randomized assignment of the sample group across the treatments was deemed successful.

The final outcome of the sampling execution was a set of eight treatment groups, each containing between 4,284 and 4,404 subjects (mean of 4,375 per treatment group). The smallest treatment group consists of 4,284 subjects, meeting the minimum size requirement of 4,244 defined in Section 3.8.4.

Table 3.16 Data Element Comparison - Sample Group versus Treatment Groups

			Treatment Group Assignment							
Element	**Value**	**Sample Group**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Number Records		35,000	4,400	4,284	4,397	4,335	4,412	4.384	4,404	4,384
Unique Accounts		100%	100%	100%	100%	100%	100%	100%	100%	100%
Tenure (days)	Min	13	13	13	13	13	13	13	14	14
	Max	11,998	10,809	11,125	10,961	11,340	11,224	11,998	10,577	11,671
	Mean	2,805	2,797	2,850	2,793	2,800	2,768	2,878	2,788	2,767
	St. Dev.	2,200	2,197	2,218	2,189	2,221	2,173	2,259	2,171	2,171
Language	English	85.2%	84.4	85.4	85.2	84.8	85.5	85.7	86.0	84.9
	French	14.7%	15.5	14.6	14.7	15.2	14.5	14.3	13.9	15.0
Billing Cycle Date	1	6.4%	6.1	6.1	6.1	6.4	6.6	7.3	6.6	6.0
	2	5.9%	6.2	6.0	5.7	6.0	6.2	5.7	5.8	5.9
	6	5.6%	5.3	5.0	5.7	6.0	5.9	6.0	5.9	5.4
	7	5.8%	5.3	6.0	5.8	6.0	5.5	5.6	6.0	6.0
	8	5.6%	5.9	5.2	5.8	5.6	5.3	5.5	4.9	6.2
	10	5.6%	5.6	5.6	5.2	5.8	5.4	5.7	6.0	5.7
	11	5.8%	5.7	5.9	6.5	5.7	5.6	5.7	6.1	5.6
	12	5.2%	5.0	2.2	5.1	5.3	5.0	6.0	5.0	5.0
	14	5.7%	6.0	5.4	5.8	6.1	5.5	5.8	6.0	5.4
	15	5.8%	6.0	5.7	6.2	5.2	5.5	529	5.6	6.2
	17	6.1%	6.0	6.3	6.3	6.0	6.3	5.5	6.2	5.8
	18	5.2%	5.0	5.3	5.1	5.1	5.7	5.0	5.0	5.3
	22	7.2%	7.3	8.1	7.0	7.7	6.9	7.0	7.0	6.7
	25	6.6%	7.1	6.3	6.5	6.9	6.6	5.9	6.4	7.0
	26	5.9%	6.2	6.3	6.0	5.2	5.5	5.6	6.1	6.3
	28	5.9%	5.6	6.1	6.3	5.5	6.5	5.5	6.4	5.7
	29	5.6%	5.6	5.6	4.8	5.4	6.1	6.1	5.1	5.8

Element	Value	Sample Group	Treatment Group Assignment							
			1	2	3	4	5	6	7	8
Credit Class Code	B	86.2%	86.8	86.3	86.3	85.9	86.1	86.4	85.6	86.4
	C	3.6%	3.7	3.8	3.5	3.4	3.8	3.4	3.5	3.4
	D	8.6%	8.0	8.3	8.6	8.7	8.4	8.6	9.3	8.6
	L	0.5%	0.6	0.5	0.5	0.7	0.4	0.4	0.5	0.4
	V	0.3%	0.3	0.3	0.3	0.3	0.2	0.2	0.2	0.3
	X	0.9%	0.6	0.7	0.8	0.9	1.1	1.0	0.7	1.0
Contract Status	On	37.5%	38.0	37.1	37.6	36.4	38.5	37.0	37.5	37.7
	Off	62.5%	62.0	62.9	62.7	63.6	61.5	63.0	62.5	62.3
Contract Duration	24 month	98.9%	99.3	98.7	98.8	98.8	98.9	98.3	99.0	99.2
	Other	1.1%	0.7	1.3	1.2	1.2	1.1	1.7	1.0	0.8
Days Left on Contract	Minimum	1	1	1	1	1	1	2	2	1
	Maximum	365	365	365	365	365	365	365	365	365
	Mean	204.40	203.36	204.72	204.84	206.89	212.26	205.48	202.50	203.63
	St. Dev.	99.797	101.411	99.633	99.359	100.359	97.868	99.057	98.969	98.840
Billing Province	AB	27.4%	27.0	27.1	28.6	27.2	27.4	27.4	27.7	26.4
	BC	23.9%	24.2	23.9	23.7	24.2	23.4	23.3	24.4	23.7
	MB	4.8%	4.5	4.6	5.0	4.8	5.1	4.8	4.2	5.2
	NB	1.2%	1.1	1.1	1.2	1.3	1.1	1.2	1.4	1.5
	NL	1.2%	1.2	1.3	1.3	1.2	1.1.	1.0	1.1	1.1
	NS	2.0%	2.2	2.5	1.7	1.7	2.2	1.9	2.2	1.6
	NT	0.1%	0.1	0.1	0.1	0.0	0.1	0.2	0.1	0.0
	NU	<0.0%	<0.0	<0.0	<0.0	<0.0	<0.0	<0.0	<0.0	<0.0
	ON	20.7%	20.2	21.0	19.5	20.5	20.5	21.6	20.9	21.1
	PE	0.3%	0.3	0.2	0.4	0.2	0.4	0.3	0.4	0.4
	PQ	15.%	16.3	15.8	15.9	16.4	15.8	15.5	14.9	16.1
	SK	2.4%	2.5	2.1	2.4	2.1	2.6	2.5	2.7	2.6
	YT	0.1%	0.2	0.1	0.1	0.1	0.1	0.2	0.0	0.0
Rural or Urban	Rural	21.4%	22.8	21.3	21.2	21.6	20.6	21.2	21.2	20.9
	Urban	78.6%	77.2	78.7	78.8	78.4	79.4	78.8	78.8	79.1
Revenue Band	A	4.4%	4.6	4.6	3.8	4.5	4.3	4.4	4.4	4.7
	B	29.5%	29.0	29.5	30.6	29.6	30.2	28.2	31.0	27.9
	C	28.6%	28.6	27.9	28.2	27.7	28.1	29.5	28.4	30.0
	D	17.5%	17.5	17.7	17.9	17.8	17.7	16.9	17.3	17.4
	E	9.1%	9.8	9.7	8.6	9.0	8.4	9.6	8.8	9.2
	F	2.1%	2.3	1.9	2.1	1.9	2.1	2.3	1.8	2.6
	N	1.6%	1.4	1.4	1.5	1.7	1.9	1.8	1.6	1.5
	S	6.7%	6.6	6.7	6.8	7.1	6.8	6.7	6.3	6.1
	Other	0.5%	0.2	0.6	0.5	0.7	0.5	0.6	0.4	0.6

Element	Value	Sample Group	Treatment Group Assignment							
			1	2	3	4	5	6	7	8
Do Not Survey	Yes	1.9%	98.1	97.1	98.0	98.1	98.3	98.3	98.2	97.7
	No	98.1%	1.9	1.9	2.0	1.9	1.7	1.7	1.8	2.3
Do Not Call	Yes	0.5%	0.5	0.4	0.4	0.5	0.5	0.4	0.6	0.6
	No	99.5%	99.5	99.6	99.6	99.5	99.5	99.6	99.4	99.4

3.8.5.5 Mapping Treatments and Treatment Groups

The final step in completing the randomized sampling and assignment was the allocation of the eight treatment combinations determined by the factorial design (Section 3.6.9) to the eight treatment groups. This allocation determines which study participants receive which treatment(s), shown in Table 3.17.

Table 3.17 Treatment and Treatment Group Mapping

Treatment Group	Factorial Notation	Receives Treatment:		
		VALUE	CSERV	CONTACT
1	$a_0b_0c_0$	N	N	N
2	$a_1b_0c_0$	**Y**	N	N
3	$a_0b_1c_0$	N	**Y**	N
4	$a_1b_1c_0$	**Y**	**Y**	N
5	$a_0b_0c_1$	N	N	**Y**
6	$a_1b_0c_1$	**Y**	N	**Y**
7	$a_0b_1c_1$	N	**Y**	**Y**
8	$a_1b_1c_1$	**Y**	**Y**	**Y**

3.9 Data Collection

Data on the two dependent variables identified in the theoretical model of Section 3.3 - defection (DEFECT) and customer satisfaction (CSAT) - will be gathered through both observational means and questionnaires, respectively.

3.9.1 Use of Observational Collection Methods

The study utilizes a unique aspect to data collection available as a result of the integration of the study into the subject organization. With this integration, data that reflects certain behavioural activities of the study participants (including defection) can be obtained through observational methods without requiring action by the participant.

The use of observational data collection supports the naturalistic approach taken with this study design, and as discussed previously in Section 3.6.12 (Christensen et al. 2015).

Christensen et al. suggest several advantages to the use of observational data that are relevant to this study; the most relevant include:

- Ability to measure what study participants do, without having to rely on what they say they do (or say they would do)
- Provides information that participants might be otherwise unwilling to talk about
- Provides an objective measurement of behaviour, unbiased by the participant or the researcher
- Behaviour is recorded in its natural setting and provides a greater degree of realism

Christensen et al. also describe potential disadvantages to the use of observational data collection; however, many of these potential disadvantages are ameliorated by the nature of how the observational data in this study is collected. The behaviour of the participant who defects during the course of the study is recorded in the subject organization's customer billing system at the time service is discontinued. As a result, the following potential disadvantages to observational data collection are less of a concern in this study:

- Reactive effects as a result of the participant's knowledge they are being observed
- Researcher effects
- Limitations on the sampling and measurement of large and dispersed populations
- Time-consuming data analysis, and associated expense

Attitudinal data, such as that for the measure of the CSAT DV, are measured through email / online surveys.

3.9.2 Data Collection for the DEFECT DV

The dependent variable DEFECT is measured in this study in the same manner as the subject organization's measurement of subscriber churn, a key performance indicator (KPI) of the telecommunications industry.

Deactivations are measured for each of the eight treatment groups, each month, for the duration of the study. At the end of each calendar month throughout the study timeframe, data for each of the 35,000 participants is obtained from the subject organization's customer billing system. The data obtained consists of the following information:

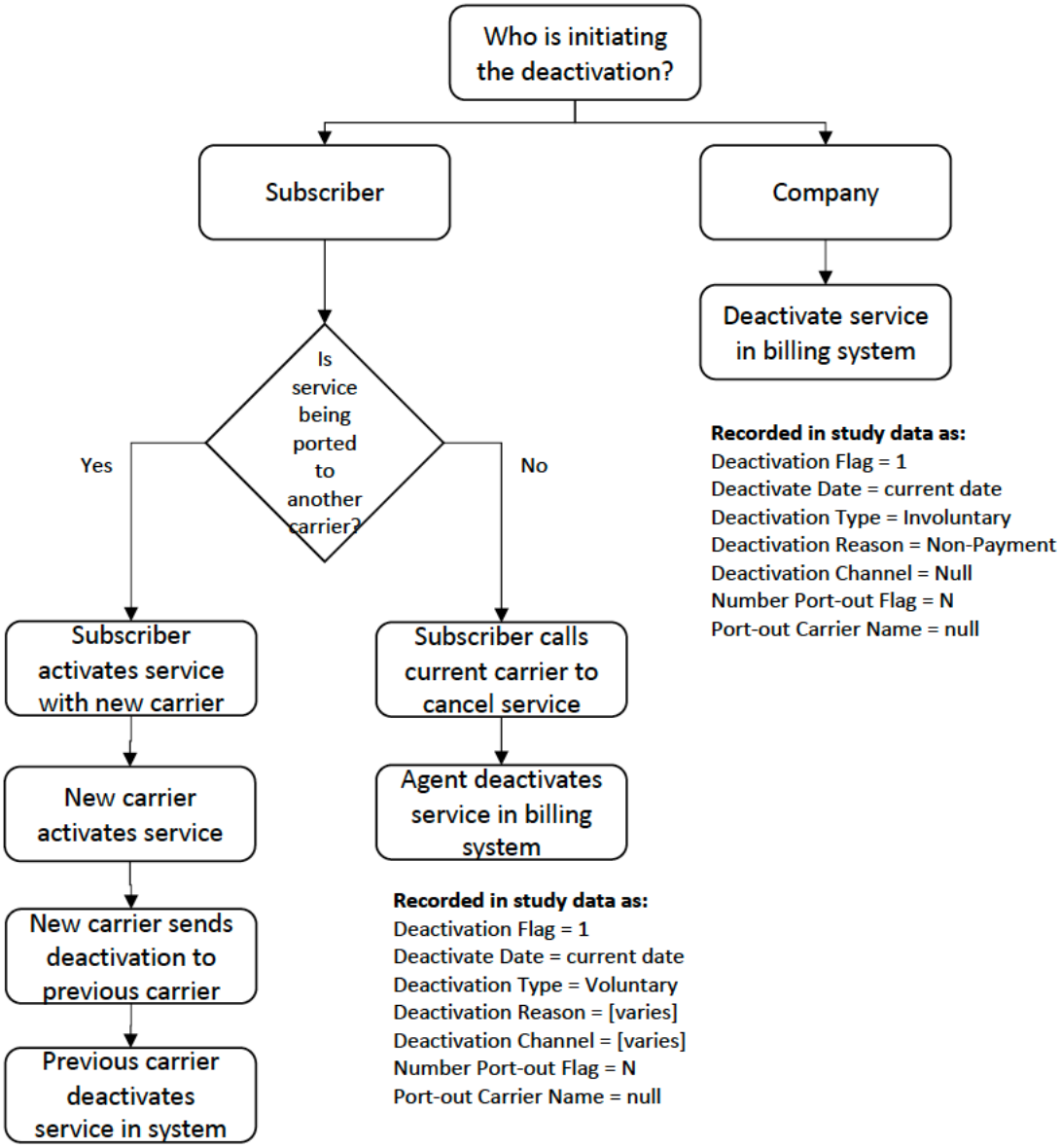
- Deactivation flag
- Deactivation date
- Deactivation type
- Deactivation reason
- Deactivation channel
- Number port-out flag
- Port-out carrier name

Appendix B provides an explanation of each of the data fields listed above and example values for each field.

The overall flow for deactivation behaviour by a study participant and the recording of the related deactivation data elements is shown in Figure 13.

For each participant in the study, if there is a deactivation of service for that participant (whether the participant or the subject organization initiates the deactivation), a record is created in the billing system. Figure 13 shows a high-level view of the process, with three types of deactivation behaviours detailed. The figure also shows the values recorded for each field in the deactivation record.

Figure 13 Deactivation Process Flow and Outcome Measurement



3.9.3 Data Collection for the CSAT DV

Section 3.6.2 proposed that the dependent variable represent satisfaction levels of the participants in the study replicate that used by the subject organization in its quarterly customer satisfaction research. The use of the same definition for the customer satisfaction dependent variable (CSAT DV) provides familiarity with the management team of the subject organization, useful for interpretation and application to actual business practices.

The specific measure employed in this study as the CSAT DV is a likelihood to recommend (L2R) score, obtained from each study participant through an online survey. The question posed to survey participants is as follows:

> "If a colleague, friend or family member were looking for a Wireless service provider, what is the likelihood that you would recommend your current service provider to them?"

Responses are captured on a 6-point Likert-like scale, consisting of the following:

- Definitely
- Probably
- Maybe
- Probably Not
- Definitely Not
- Don't Know

The overall customer satisfaction survey deployed in this study consists of a total of eleven questions, replicating the quarterly satisfaction survey of the subject organization. The full satisfaction survey is shown in Appendix C and was deployed in both English and French as required for each study participant. Appendix D outlines the various response scales used throughout the survey for each question.

The inclusion of the full set of satisfaction questions used by the subject organization allows for observation on the impact, if any, of the experimental treatments on other aspects of customer satisfaction measured in the survey.

The satisfaction survey was deployed to all study participants (with the exception of those participants that had indicated that they did not wish to receive marketing contacts or surveys, following the permission marketing guidelines of the subject organization).

The survey was communicated via email (an example is shown in Appendix E), which included a URL link to the online survey.

Execution of the survey was conducted by a third-party research company, which also conducts other surveys and data collection activities for the subject organization.

Following the pretest-posttest design of the experimental study described in Section 3.6.12, the satisfaction survey was deployed immediately prior to the commencement of the treatments (pretest), and again at the end of the 12-month study timeframe (posttest).

3.9.4 Data Collection for the Independent Variables

The conceptual model proposed in section 3.3 identifies three independent variables (IVs) that are targeted for manipulation by the treatments in the study, one IV for each treatment. Each of the IVs is measured in the same satisfaction survey as the dependent variable for overall satisfaction (L2R); the corresponding survey question for each IV is shown in Table 3.18.

Table 3.18 Data Collection for Independent Variables

Independent Variable (IV)	Survey Question	Measurement Scale (Appendix D)
PERCVALUE	"Considering the overall quality and the overall price you pay, how would you rate (subject organization) for overall value in terms of being worth what you pay for?" (Question 6 on the survey)	Excellent Very Good Good Fair Poor Unsure
SERVICE	"Thinking of your most recent (contact with customer service over the phone), how would you rate your experience with (the time you waited to speak to a representative)?" (Question 9)	Excellent Very Good Good Fair Poor
MARKETING	"How often are you frustrated with (communication quality, e.g. relevance or frequency of emails, other marketing)?" (Question 7)	Very often Often Sometimes Rarely Not at all Unsure / Not applicable

3.9.5 Data Collection for the Treatments

The precise nature of how each treatment manifests itself for each participant in the study will vary, depending on specific attributes of the participant or their behaviour outside the study. Data were collected on a monthly basis for how each treatment unfolded for each participant as an individual, and for each of the eight treatment groups as a whole. Details on the data collected for each treatment are shown in Table 3.19. Analysis of the treatment outcomes is explored in Chapter 4.

Table 3.19 Data Collection for Treatments

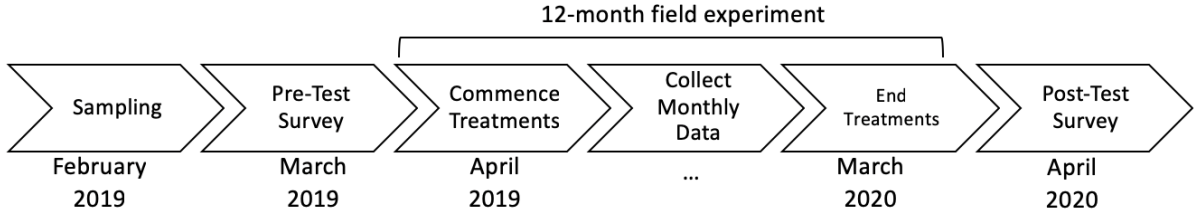
Treatment	Monthly Data Collected for Each Participant
VALUE	Amount of discount received on invoice
	Total recurring fees on the invoice
	Total usage fees on the invoice
	Total amount of invoice before discounts (recurring fees + usage fees)
	NET amount of invoice after discounts (recurring fees + usage fees – discount)
CSERV	Number of calls into the call centre
	Average time each call waited in a queue before reaching an agent
CONTACT	Number of marketing messages sent
	Number of mandatory regulatory messages sent

3.10 Experiment Timeline

The previous sections of this chapter identify a number of activities that were executed in a specific sequence. The experimental nature of the study required close attention to the scheduling and execution of each step of the experiment, spanning a total of 15 months.

Figure 14 provides a high-level view of the timing of the key components of the study.

Figure 14 Experiment Timeline



3.11 Summary

This chapter of the study established the conceptual model and hypotheses to respond to the research problem identified in Section 1.2 and detailed the research methods used to test those hypotheses.

The experimental, longitudinal, randomized controlled trial (RCT) design of the study required detailed planning and execution of many important activities before the experiment could begin, and data could be collected. In particular, care and attention were paid to the definition and measures of the independent and dependent variables, and the treatments used to manipulate the antecedents to the DVs under investigation.

The randomized sampling and randomized assignment of study participants was defined, and analysis completed to confirm the achievement of the desired randomization outcomes, which are critical to supporting the internal and external validity of the study.

The robust design and execution of the methods used in the study create confidence for the results to be reviewed in the following chapter.

4 Outcomes and Results

4.1 Introduction

This chapter tests the hypotheses outlined in Section 3.4. Overall interpretation of the results and a discussion of the implications for practice will be addressed in Chapter 5 (Discussion). Results of a practical, but not statistical, significance are reported in Appendix J but are not discussed in this or the following chapter.

This chapter begins with a summary (Section 4.2) of the statistical methods to be used throughout the analysis of the experiment, followed by four main areas of analysis:

- analysis of the three treatments employed in the study, examining how they manifested themselves to the study participants (Section 4.3);
- a review of how those treatments affected the independent variables (IVs) described in the study model (Section 4.5);
- examination of the relationships between the IVs and the two dependent variables (Sections 4.6 and 4.7)
- examination of the direct impact of the treatments on the two dependent variables (Sections 4.8 and 4.9).

4.1.1 Note Regarding the Impact of COVID-19 Pandemic

The 12-month experimental randomized controlled trial (RCT) commenced April 1, 2019. As the experiment approached its conclusion in March 2020, the COVID-19 pandemic was advancing around the world. In Canada - where the experiment was conducted - the greatest impact took effect mid-March as the border with the United States was closed, Parliament was suspended, businesses closed their offices and most public venues such as restaurants and gyms were closed[21].

The experiment was completed on March 31, 2020 as planned, and there was no impact on the gathering of data following the completion of the experiment, with the exception of the posttest customer satisfaction survey which was delayed slightly to mid-April 2020. Details of the impact of the pandemic on the execution of the posttest survey are reviewed in Section 4.4. For the remainder of the components of the experiment, the relatively brief overlap of the pandemic shutdown and the end of the experiment is not expected to materially impact the

[21] COVID-19 in Canada: A Two-year Update on Social and Economic Impacts
https://www150.statcan.gc.ca/n1/pub/11-631-x/11-631-x2022001-eng.htm

results of the study. In addition, the randomized design of the study ensures that any pandemic effects would be equally represented in the various treatment groups.

4.1.2 Presentation of Treatment Groups in Analysis

To aid the understanding of the analysis for each of the eight treatment groups utilized in the study, the naming structure shown in Table 4.1 is employed throughout this chapter. The treatment group name is constructed of the one or more treatments experienced by that group, as defined in Section 3.6.9.

Table 4.1 Treatment Group Naming Convention

Treatment Group Number	Treatment Group Name
1	CONTROL
2	VALUE
3	CSERV
4	VALUE + CSERV
5	CONTACT
6	VALUE + CONTACT
7	CSERV + CONTACT
8	VALUE + CSERV + CONTACT

4.2 Statistical Analysis Methods Used

The determination of the appropriate methods to be employed in analyzing the data from the experiment follows from the objectives of the Discussion (Chapter 5) and Conclusion (Chapter 6) chapters of this study. Mohajeri et al. (2020) propose that - in addition to reporting results based on statistical significance - quantitative researchers should supplement their reporting with an analysis that supports an understanding of the practical significance of the study, stating that “Practical significance is as equally important a measure as statistical significance when it comes to rigorous quantitative research.”. The focus of this chapter, however, will be the analysis of results based on statistical significance; references to findings of practical significance – when relevant to the interpretation of the experimental results – are confined to Appendix J.

The statistical analysis of randomized controlled trial studies may employ one or more techniques, depending on the specific design of the study. Consideration is given to the research problem proposed in Section 1.2, together with how the study findings might be utilized in practice to determine the appropriate analytical methods to employ.

Analytical approaches to determining statistical significance are reviewed first, followed by the review of analytical tools to examine the practical significance of the results.

4.2.1 Methods for Evaluating Statistical Significance

The research problem concluded with the question “Are these decisions (efforts to improve customer loyalty through improved customer satisfaction) supported by actual customer behaviour?”. The researcher seeks answers to that question through an examination of the impacts (if any) of the treatments employed in the study, and in particular:

1. Are there differences in customer satisfaction or loyalty between the groups receiving the different (or no) treatments?
2. Are there differences in customer satisfaction or loyalty outcomes between groups, and if so, which treatments are associated with those differences?
3. And if there are specific treatments associated with changes in customer satisfaction or loyalty, how much of an impact do those treatments have on the outcomes?

With these questions in mind, a review of statistical techniques was undertaken to determine the appropriate set of analyses to be completed.

The study design yields a continuous outcome variable (deactivation rates between 0.00 and 1.00, represented by the dependent variable DEFECT) across eight unrelated treatment groups. Based on this study design, Kim et al. (2017) recommend the use of Analysis of Variances (ANOVA) to respond to the question posed in #1 above, as shown in Table 4.2. The use of ANOVA to analyze a factorial design study is also supported by Salkind (2010).

Kim et al. also identify statistical techniques to be used in answering question #2, examining the association between an outcome (dependent) variable and one or more explanatory variables. Regression modelling is proposed as the appropriate approach. Table 4.3 summarizes the specific regression model appropriate to a given study design, specifically the nature of the outcome variable. The outcome variable in the study presented here is continuous, indicating that a linear regression model is an appropriate technique to be applied. Although the outcome variable is designed to yield continuous values, the underlying value of the variable for each individual participant is not continuous but is dichotomous with a value of 0 or 1. In order the ensure that the use of a linear regression in analysing the dependent variable DEFECT is appropriate, logistical regression and probit regression

analyses (appropriate for use where the outcome variable is dichotomous) are also employed and the results compared to the linear regression analysis.

The selection of both the ANOVA and linear regression analysis techniques for randomized controlled trials is supported by Solomon et al. (2009, p. 133) who state “Testing of hypotheses is the final step in operationalizing the RCT. Moving beyond ANOVA (analysis of variance), the classic model for testing the effectiveness of an intervention is a regression model…”.

Consideration was also given to an analytical method often used in econometric studies, namely the difference-in-differences (DID) method. At first glance, RCT and DID studies appear quite similar as both have a well-defined study population and a given set of treatment conditions to be studied, and both also share attributes such as the ability distinguish between a treatment group and a control group, as well as between pre-treatment and post-treatment time periods (Wing et al. 2018).

It is noted, however, that DID is “a quasi-experimental research design that researchers often use to study causal relationships…where randomized controlled trials (RCTs) are infeasible” (Wing et al. 2018, p. 453) and that it is a “…popular way to estimate causal relationships” (Bertrand et al. 2004, p. 2).

This study is a full RCT experiment that provides proof of the existence (or not, as the case may be) of casual relationships between the variables of interest. However, the difference-in-differences method will be used where applicable (that is, where there is a measurement of a DV both pre- and post-treatment is available) in order to provide additional confirmation of the results obtained in the ANOVA and regression analyses.

A significance level of α = 5% will be employed to identify results deemed statistically significant.

Table 4.2 Statistical Tests to Determine if Outcome Variable is Different Across Two or More Explanatory Groups

Outcome Variable	Explanatory Variable			
	Dichotomous (Unrelated)	**Dichotomous (Related)**	**Three or More Subgroups (Unrelated)**	**Three or More Subgroups (Related)**
Continuous (normally distributed)	Two-sample *t-test*	Paired *t-test*	**Analysis of Variance (ANOVA)**	Mixed-effects model for repeated measures
Ordinal, Continuous (not normally distributed)	Wilcoxon rank-sum test	Wilcoxon signed-rank test	Kruskal-Wallis test	Friedman test, Skillings-Mack test
Categorical	Chi-square test, Fisher exact test	McNemar test	Chi-square test, Fisher exact test	Cochran *Q* test

Adapted from Kim et al. (2017)

Table 4.3 Regression Models to Examine Association of Outcome Variable with One or More Explanatory Variables

Outcome Variable	Regression Model
Continuous variable	**Linear regression**
Ordinal variable	Ordinal logistic regression
Count variable	Poisson regression
Dichotomous variable (various types)	Conditional/unconditional logistic regression, Log binomial regression, Multinomial logistic regression

Adapted from Kim et al. (2017)

4.2.2 Methods to Determine Practical Significance

Practical significance refers to "…evaluations of significance or importance made based on the data analysis for a study, but with a focus separate from the obtained *p* value" (Spurlock 2019, p. 624). Spurlock also suggests that while experimental research is "…amenable to evaluation of the practical significance of the statistical findings" it is rarely seen in published

research. Given the desire to demonstrate practical applications for the results of the study, an effort will be made to explore the practical significance of the findings.

Communication of the practical significance of the study results will include effect magnitudes (Peeters 2016, Mohajeri et al. 2020) and confidence intervals (Mohajeri et al. 2020). Given the large sample size of this study (n = 35,000), it is possible that the relationships between the treatments and the DVs may be quite small, but simultaneously be statistically significant (Howitt et al. 2017). As noted by the last two authors, the inclusion of effect size analysis will contribute to the understanding of "…whether or not the researcher's findings are of any real substance.".

Multiple measures of effect size exist, with the appropriate measure to be used dependent on the study design. Where regression analysis is conducted, adjusted R^2 will be used to measure the proportion of variance in the dependent variable that is attributable to the independent variable(s) (Peeters 2016). The use of adjusted R^2 rather than simply R^2 ensures that the impact of multiple independent variables is accounted for. Further, for the analysis of practical significance, the partial Eta squared (η^2) effect measure will be determined for each of the treatment factors in the study. The η^2 effect measure provides a standardized report of the proportion of variance explained by the factor (or combination of factors).

The categorization of the effect sizes observed using the partial Eta squared approach follows the recommendations by Howitt and Cramer (2017) that an effect size of 0.01 or less be considered Small, around 0.06 be considered Medium, and that of 0.13 or more be considered Large.

The results obtained through analyses of practical significance are presented in Appendix J.

4.3 Treatment Application and Participant Experience

Section 3.6 defined the three treatments utilized in the experiment:

- VALUE: 10% discount on monthly recurring service fees
- CSERV: Prioritized access to customer service representatives
- CONTACT: Reduced frequency of marketing communications

The individual treatment experience of each study participant varies depending on certain characteristics and the behaviours of the study participants themselves. The analysis in this section of the study details how each of these three treatments was experienced by each participant through an examination of the treatment results across each of the eight study groups in the experiment.

4.3.1 VALUE Treatment Analysis

The VALUE treatment consisted of a 10% discount on the monthly recurring charges (MRC) incurred by a participant assigned the treatment. As defined in Section 3.6.5, the actual monetary value of the treatment will vary for each participant as the MRC also varies for each. The MRC amount may also change over time, as the study participant makes changes to his or her mobile service plan. MRC amounts presented throughout this section exclude any discount associated with the VALUE treatment; the treatment discount is calculated separately and subtracted from the participant's total invoice.

Analysis of the MRC for each of the study groups yields the results shown in Table 4.4. At the beginning of the experiment, the mean MRC for the entire study group was $73.78, which declined to $58.11 (-21%) by the end of 12 months. This decline is consistent with the experience of the subject organization: over time, mobile phone subscribers may change their mobile rate plan, either as a stand-alone action or in conjunction with buying a new phone and recommitting their contract (renewing). Competitive actions in the market generally result in lower prices for the same service, or more services for the same price, over time.

Decreases in MRC across the treatment groups ranged from an 18% decrease in group 6 to a 23% decrease for group 5. Collectively, the four groups receiving the VALUE treatment exhibit a decrease in MRC at the end of the experiment of -22.5%, compared to the non-treatment groups which showed a decrease of -20.0%. A one-way ANOVA analysis was conducted to determine if there was any statistical difference between the mean MRC values for each group, both at the beginning of the experiment (April 2019) and at the end (March

2020). The ANOVA resulted in p = .065 at the beginning of the experiment, and p = .137 at the conclusion; both of which are not statistically significant at the 5 percent level.

Table 4.4 Mean Monthly Recurring Charge by Treatment Group

Treatment Group	Mean Monthly Recurring Charge ($)											
	Apr 2019	May 2019	Jun 2019	Jul 2019	Aug 2019	Sep 2019	Oct 2019	Nov 2019	Dec 2019	Jan 2020	Feb 2020	Mar 2020
1	73.68	72.07	71.25	69.95	68.17	66.24	64.28	62.71	61.06	59.10	58.36	57.30
2	73.57	72.15	71.22	70.17	68.20	66.07	64.61	63.26	62.01	60.24	59.38	58.27
3	74.58	73.35	72.18	70.63	68.63	66.94	64.62	63.42	61.86	59.99	59.00	57.85
4	73.72	73.01	71.99	70.72	69.06	67.40	65.20	63.71	62.02	60.43	59.69	58.51
5	73.90	71.84	71.58	70.01	68.01	65.89	63.86	62.15	60.56	58.83	57.91	56.88
6	71.39	71.23	71.13	69.99	68.18	66.86	65.08	64.02	62.40	60.85	59.70	58.62
7	75.07	73.33	72.68	71.21	69.36	67.43	65.52	63.75	61.98	60.31	59.25	58.40
8	74.30	73.38	72.10	71.58	69.48	68.04	65.92	64.65	63.47	60.48	60.22	59.08
ALL	73.78	72.55	71.77	70.53	68.64	66.86	64.88	63.46	61.92	60.03	59.19	58.11

Note: Groups 2, 4, 6 & 8 received the VALUE treatment

Analysis of the mean monthly discount received by each participant in each treatment group is shown in Table 4.5. As expected, only treatment groups 2, 4, 6 and 8 received discounts, consistent with the treatment assignment. This confirms the absence of any treatment diffusion, thereby contributing to construct validity (Section 3.7.2).

Across the total study period, the mean monthly discount amount received by each participant was $6.27, with a range of $6.22 (groups 2 and 6) to $6.29 (groups 4 and 8). This discount amount represented 8.1% of the MRC; the difference between the actual discount rate of 8.1% and the treatment design of 10% is explained by the existence of certain MRC services that were not eligible for the monthly discount, as determined by policies of the subject organization (for example, services such as device insurance and roaming plans are not eligible for the treatment discount).

The total cost of the VALUE treatment over the study period was $1,307,171 or $75.18 per participant receiving the treatment, consistent with the amount estimated at the time of the treatment design as described in Section 3.6.5.

Table 4.5 Mean Monthly Discount Received by Treatment Group

Treatment Group	Average Mean Discount ($)											
	Apr 2019	May 2019	Jun 2019	Jul 2019	Aug 2019	Sep 2019	Oct 2019	Nov 2019	Dec 2019	Jan 2020	Feb 2020	Mar 2020[a]
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	6.80	7.14	6.97	6.88	6.72	6.51	6.36	6.22	6.07	5.91	5.49	3.58
3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	6.81	7.25	7.06	6.94	6.81	6.63	6.44	6.29	6.11	5.97	5.55	3.66
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	6.60	7.04	6.93	6.85	6.68	6.56	6.41	6.30	6.14	5.97	5.51	3.66
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	6.83	7.20	7.03	6.98	6.78	6.65	6.43	6.32	6.17	5.98	5.51	3.64
Mean of Groups 2,4,6,8	6.76	7.16	7.00	6.91	6.75	6.59	6.41	6.28	6.12	5.96	5.51	3.64

[a] The discount received by all study participants in March 2020 (the last month of the study) is pro-rated based on the participant's bill cycle date. For example, a participant with a bill cycle date of 18 would have received the treatment discount for the portion of their monthly invoice that occurred between March 1st and 18th, and no discount for the remaining 13 days of the month.

4.3.2 CSERV Treatment Analysis

The CSERV treatment enabled participants assigned this treatment to access customer service agents more quickly (when calling into the call centre) than participants not receiving the treatment. As reviewed in the treatment design in Section 3.6.6, not all study participants will call the call centre in any given month or may not call at all through the 12-month duration of the experiment.

This section of the study analyzes the calling behaviour of all participants in the experiment and examines the experience those participants received in terms of how long it took to reach an agent. Note: Due to a reporting system issue at the subject organization, data on the calling behaviour of participants was only made available for 11 of the 12 months of the study (May 2019 through March 2020, missing April 2019).

4.3.2.1 Frequency of Calling Activity

Throughout the timeframe of the study, participants could have chosen to call the subject organization's call centre for some aspect of support. An analysis of the calling behaviour of the participants in each treatment group is shown in Table 4.6.

The percentage of participants who made one or more calls to the call centre over the study was 43.6%, with the percentage across treatment groups ranging from 42.5% to 44.2%. The mean number of calls made by all participants in the study was 1.48, with a range among treatment groups of 1.41 to 1.62. Table 4.7 shows the mean number of calls for those participants who made one or more calls over the study period: the overall mean for all of these participants was 3.39, with the range across treatment groups of 3.20 to 3.65.

A one-way ANOVA analysis was conducted to determine if there was any statistical difference between the Total Number of Calls over the course of the study by the participants of each treatment group. The ANOVA results (p = .213) indicate no statistically significant difference in the number of calls made between one or more of the treatment groups.

Further ANOVA analysis was conducted for each of the 11 months of the study data, shown in Table 4.8. This analysis shows that there was no statistical difference in the number of calls made by each treatment group in 10 of the 11 months reported. Only in one month (March 2020) was there a statistically significant difference between one or more groups (p = .028, < .05). Analysis of the call volumes between the treatment groups identified as being statistically different (Tukey post-hoc analysis indicates treatment group 4 to be different from groups 5 and 7), showed the reason for the statistical significance to be a result of a slightly higher percentage of the participants in treatment group 4 that made one or more calls in the month of March 2020. In treatment group 4, 10.2% of participants made one or more calls; in treatment groups 5 and 7, only 8.2% of participants made one or more calls. This difference was sufficient for the ANOVA analysis to indicate a statistically significant difference between the groups, but in practical terms, the differences are non-consequential.

The total number of calls made by the treatment groups (3, 4, 7 & 8) compared to the non-treatment groups (1, 2, 5 & 6) are shown in Figure 15. The increased volume of calls made by all participants starting in November is consistent with patterns at the subject organization, where call volumes increase in the heavy promotional months of November and December, with carry-over into the New Year. The increase in call volumes in March 2020 is not consistent with historical results of the organization, but reflect the onset of the COVID-19 pandemic in Canada at that time, and a corresponding increase of calls from customers regarding their telecommunications services and billing questions.

Table 4.6 Participant Calling Behaviour (May 2019 through March 2020)

Treatment Group	Treatment Group Size	Participants with >0 Calls	% Participants with >0 Calls	Total Number of Calls	Mean Calls per Participant	Minimum Calls by a Participant	Maximum Calls by a Participant
1	4,400	1,941	44.1%	6,520	1.48	0	47
2	4,284	1,892	44.2%	6,119	1.43	0	71
3	4,397	1,938	44.1%	6,685	1.52	0	159
4	4,335	1,921	44.3%	7,018	1.62	0	342
5	4,412	1,873	42.5%	6,618	1.50	0	150
6	4,384	1,865	42.5%	6,472	1.48	0	75
7	4,404	1,908	43.3%	6,207	1.41	0	40
8	4,384	1,933	44.1%	6,183	1.41	0	60
Total	35,000	15,271	43.6%	51,822	1.48	0	342

Note: Calling analysis based on the 11-month period of May 2019 to March 2020; Groups 3, 4, 7 & 8 received the CSERV treatment

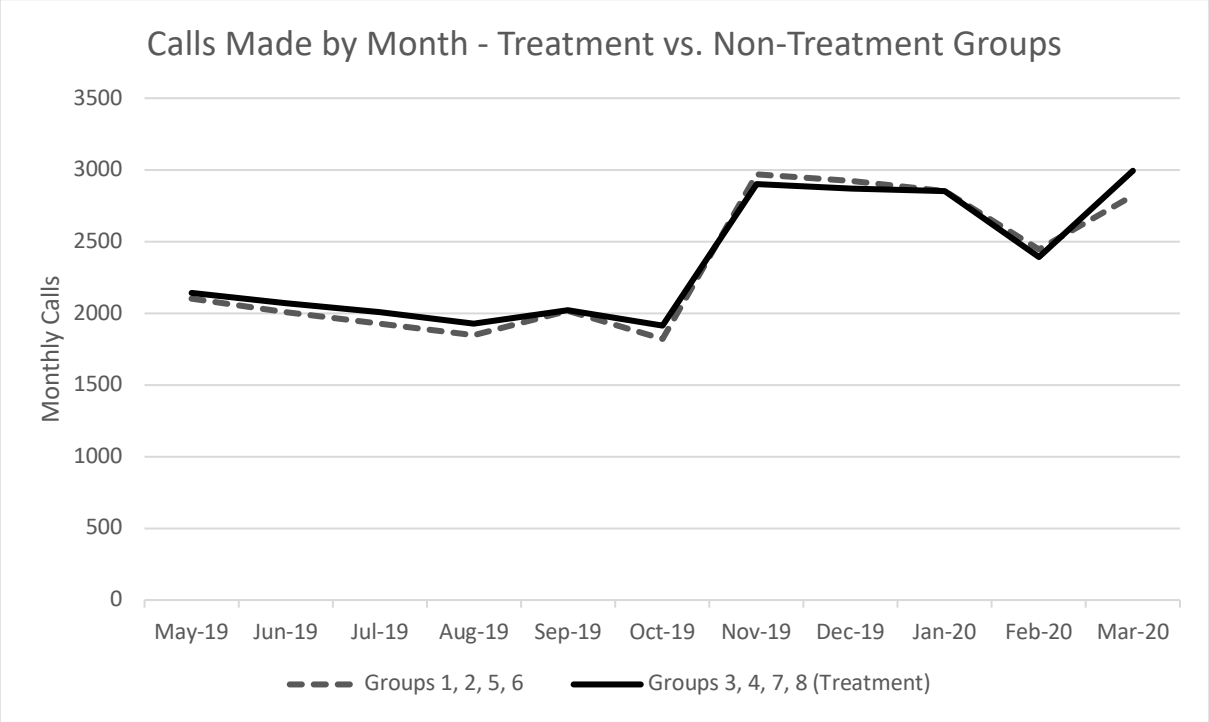
Table 4.7 Call Behaviour for Participants with >0 Calls

Treatment Group	Treatment Group Size	Participants with >0 Calls	Mean Calls per Participant
1	4,400	1,941	3.36
2	4,284	1,892	3.23
3	4,397	1,938	3.45
4	4,335	1,921	3.65
5	4,412	1,873	3.53
6	4,384	1,865	3.47
7	4,404	1,908	3.25
8	4,384	1,933	3.20
Total	35,000	15,271	3.39

Table 4.8 One-Way ANOVA Analysis of Number of Calls by Group

	Month of Study										
	May '19	**Jun '19**	**Jul '19**	**Aug '19**	**Sep '19**	**Oct '19**	**Nov '19**	**Dec '19**	**Jan '20**	**Feb '20**	**Mar '20**
Significance	.919	.972	.124	.126	.158	.517	.918	.729	.299	.740	**.028**

Figure 15 Monthly Call Volumes - Treatment and non-Treatment Groups



4.3.2.2 Wait Time Experience

Whenever a study participant called the call centre, they would experience a wait time in the Interactive Voice Response (IVR) system as described in Section 3.6.6. The CSERV treatment aimed to provide those participants receiving the treatment with faster access to an agent than those participants not receiving the treatment. The difference in wait times for those receiving the treatment versus those not receiving the treatment will vary, and in some cases, may not be much different at all. The actual experience of the treatment depends on a number of factors, including the day of the week the call was made, the time of day, and the overall busy-ness of the contact centre at the time of the call.

The mean wait time for all calls made by each treatment group is shown in Table 4.9.

A summary of the wait times experienced by the CSERV treatment participants versus those in the non-treatment groups is illustrated in Figure 16. The groups receiving the CSERV treatment show lower mean wait times from May 2019 through August 2019; from September 2019 through February 2020 the mean wait times appear similar between treatment and non-treatment groups, with a difference re-appearing in the final month of the study (March 2020) when overall call volumes increased due to COVID-19. At the time of the study design, overall wait times for all customers of the subject organization were substantially longer than was the case during the in-field timeframe of the study (for example, averaging 198 seconds in January 2017 during the initial study design, versus 5.1 seconds in January 2020 during the experiment). Between the time of the study design and the commencement of the experiment, the subject organization made substantial financial investments to improve (lower) the wait times experienced by its customers calling the call centre. As a result, the anticipated impact of the CSERV treatment may be muted as the treatment itself provides less of an expected benefit.

A one-way ANOVA of the mean wait times by month and group reveals a statistically significant difference amongst one or more groups in the months of May, June, July and August (*p*-values of .020, .001, .002 and .003. respectively, each $p < .05$). As overall wait times declined from September onwards, there was no significant difference in the mean wait times among groups ($p > .05$ in all cases). For the total 11-month period of the data, there was a significant difference in the means between one or more groups ($p = .000$, $p < .05$).

Table 4.10 summarizes the results of the One-Way ANOVA across each month. For the four months where a significant difference was observed, further post-hoc analysis was completed to determine between which treatment groups a significant difference in mean wait times existed. Tukey post-hoc analysis indicates that the significant differences are observed between all four of the treatment groups (groups 3, 4, 7 and 8) versus three of the non-treatment groups (groups 1, 2 and 5). The former groups had statistically significant lower mean wait times than the latter groups, as expected given the treatment design.

Table 4.9 Mean Wait Times (in seconds) by Group and Month

Treatment Group	Month of Study										
	May '19	Jun '19	Jul '19	Aug '19	Sep '19	Oct '19	Nov '19	Dec '19	Jan '20	Feb '20	Mar '20
1	16.2	39.9	18.3	18.1	14.2	13.3	8.0	4.4	5.8	3.1	16.9
2	17.2	36.5	22.9	21.9	17.4	11.3	8.5	6.0	4.4	3.2	12.0

Treatment Group	Month of Study										
	May '19	Jun '19	Jul '19	Aug '19	Sep '19	Oct '19	Nov '19	Dec '19	Jan '20	Feb '20	Mar '20
3	13.3	20.4	13.5	13.0	14.7	12.2	7.3	5.1	4.5	1.9	11.4
4	11.7	28.9	13.8	13.3	13.5	10.7	5.9	4.4	4.4	3.2	13.0
5	21.2	31.6	22.6	22.0	15.3	11.6	5.1	4.8	5.6	3.3	13.7
6	15.9	30.4	19.0	18.3	13.4	10.8	5.9	6.1	4.8	3.9	12.7
7	16.7	27.0	13.1	12.4	11.3	10.7	9.4	4.1	4.0	3.3	12.4
8	17.4	22.8	13.8	13.5	17.2	11.9	5.3	5.0	7.0	3.6	8.4
All Groups	16.2	29.7	17.1	16.5	14.6	11.6	6.9	5.0	5.1	3.2	12.6

Note: Groups 3, 4, 7 & 8 received the CSERV treatment

Figure 16 Mean Wait Times by Month - Treatment and Non-Treatment Groups

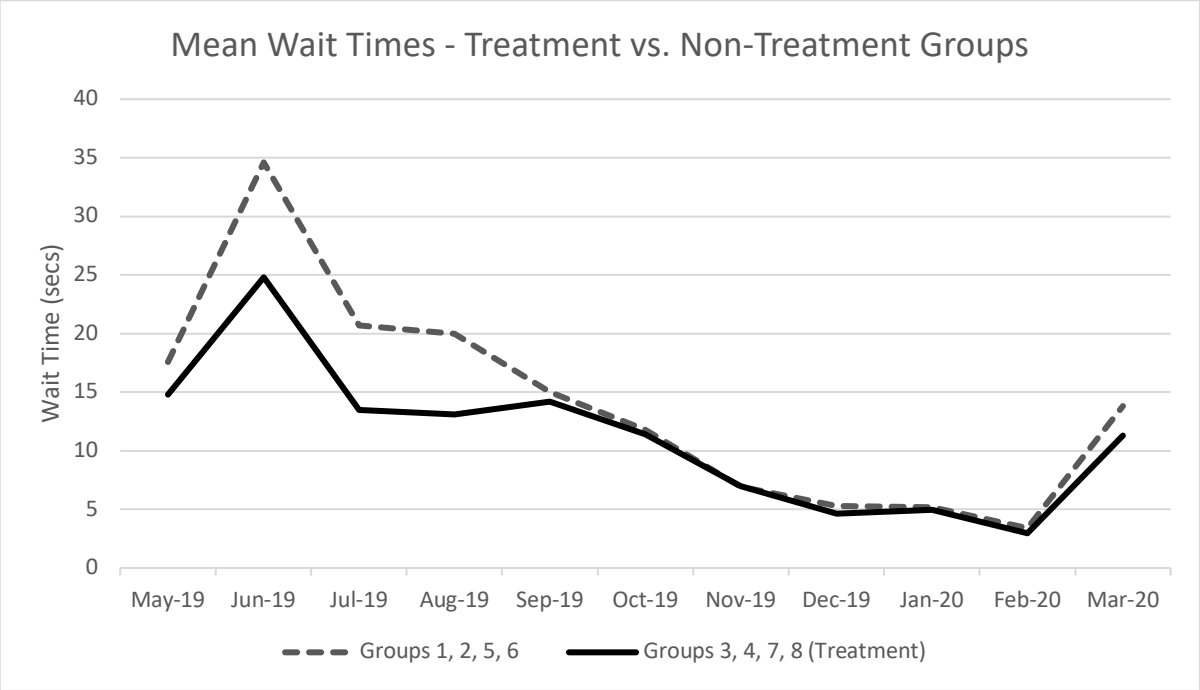


Table 4.10 One-Way ANOVA Analysis of Wait Times by Group

	Month of Study											
	May '19	**Jun '19**	**Jul '19**	**Aug '19**	**Sep '19**	**Oct '19**	**Nov '19**	**Dec '19**	**Jan '20**	**Feb '20**	**Mar '20**	**Overall 11 mths**
Significance	**.020**	**.001**	**.002**	**.003**	.222	.944	.327	.833	.614	.624	.286	**.000**

4.3.3 CONTACT Treatment Analysis

The CONTACT treatment reduced the frequency of discretionary marketing communications sent to study participants that were assigned to that treatment. Due to operational constraints within the subject organization at the time of the study launch, the implementation of this treatment was delayed for the first two-and-one-half months of the study (April 2019 through mid-June 2019) but was fully executed from mid-June 2019 onwards.

Table 4.11 shows the average monthly communications received by each treatment group over the 12 months of the study. Once the CONTACT treatment was fully implemented, marketing communications received by the treatment participants dropped to near zero, as participants only received limited automated messaging that could not be suspended for the purposes of this study.

Over the course of the study, participants that were not assigned the CONTACT treatment received a mean of 25.46 marketing messages; the treatment group (when counting all 12 months of the study) received 5.79 messages. Analysis of the nine months where the treatment was fully implemented (July 2019 through March 2020) shows that non-treatment participants received a mean of 18.96 messages over those nine months, whereas the treatment participants received 0.23 messages. Figure 17 shows the mean number of messages received by participants in the treatment versus non-treatment groups.

A one-way ANOVA of the mean number of marketing messages by month and group reveals a statistically significant difference amongst one or more groups in all months with the exception of May 2019 (p-value of .031 for April 2019, and .000 for the months of June 2019 through March 2020; each $p < .05$). Table 4.12 summarizes the results of the ANOVA analysis across each month.

Table 4.11 Mean Marketing Communications by Treatment Group

Treatment Group	Month of Study											
	Apr '19	May '19	Jun '19	Jul '19	Aug '19	Sep '19	Oct '19	Nov '19	Dec '19	Jan '20	Feb '20	Mar '20
1	2.12	2.29	2.06	1.86	2.41	2.18	2.60	2.02	2.60	1.42	2.38	1.42
2	2.07	2.32	2.05	1.86	2.38	2.11	2.59	2.04	2.60	1.42	2.35	1.40
3	2.12	2.31	2.06	1.88	2.39	2.16	2.62	2.11	2.61	1.44	2.41	1.42
4	2.15	2.32	2.11	1.90	2.42	2.15	2.65	2.08	2.66	1.46	2.43	1.42
5	2.09	2.29	1.09	0.10	0.00	0.01	0.03	0.00	0.00	0.01	0.03	0.03
6	2.13	2.34	1.09	0.09	0.00	0.00	0.04	0.00	0.00	0.01	0.03	0.03
7	2.13	2.29	1.12	0.11	0.01	0.01	0.03	0.01	0.01	0.02	0.04	0.04
8	2.20	2.34	1.13	0.10	0.01	0.00	0.04	0.00	0.00	0.01	0.03	0.03
All Groups	2.13	2.31	1.58	0.98	1.20	1.07	1.32	1.03	1.30	0.72	1.21	0.72

Note: Groups 5, 6, 7 & 8 received the CONTACT treatment; implementation of treatment did not occur until mid-June 2019

Figure 17 Mean Monthly Marketing Messages by Treatment Group

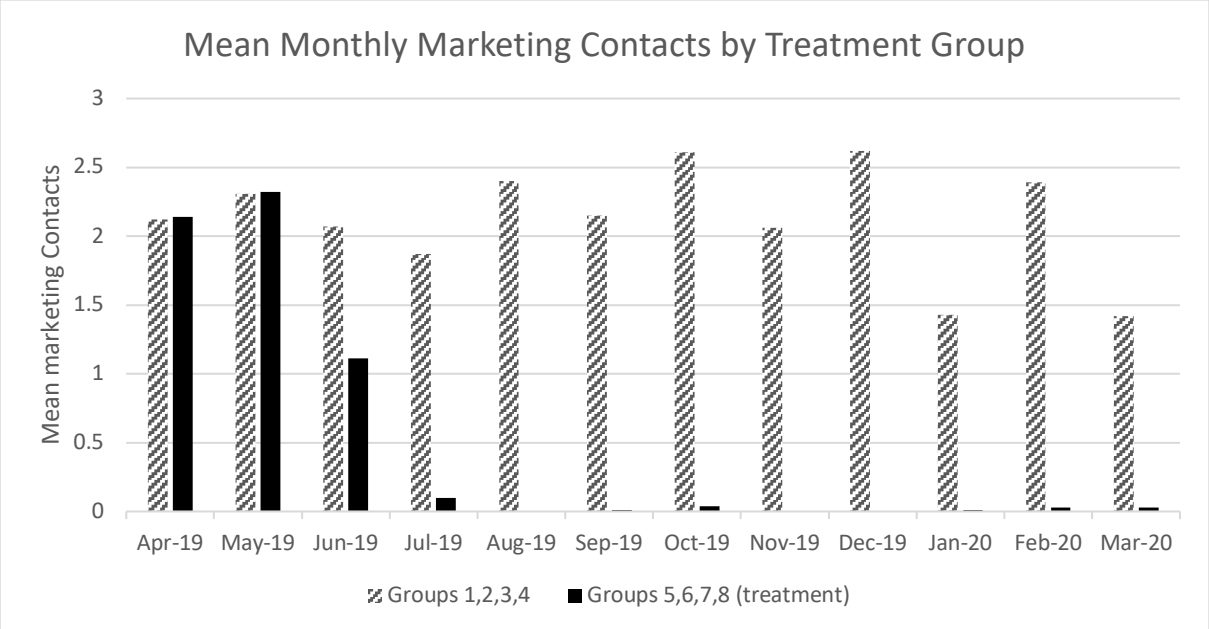


Table 4.12 One-Way ANOVA Analysis of Marketing Messages by Group

	Month of Study												
	Apr '19	**May '19**	**Jun '19**	**Jul '19**	**Aug '19**	**Sep '19**	**Oct '19**	**Nov '19**	**Dec '19**	**Jan '20**	**Feb '20**	**Mar '20**	**Overall**
Significance	**.031**	.817	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**	**.000**

4.3.4 Treatment Analysis Summary

The treatment experience of participants in the study provides important context when next examining the effect of those treatments on the independent variables, and ultimately, the dependent variables.

Table 4.13 summarizes the treatment experiences across the eight groups included in the study. Results shown in brackets indicate the differences between that treatment group's experience of the treatment versus the control group (group 1). For example, the VALUE treatment experience of group 2 resulted in that group receiving a mean monthly discount between $5.49 and $7.14 over the study period versus the control group, and the CSERV treatment experience of group 3 resulted in that group receiving a mean monthly wait time experience of 1.2 seconds through 19.5 seconds lower than the control group.

Table 4.13 Treatment Experience Summary

Group	VALUE (mean monthly discount received in $)	CSERV (mean monthly wait time in secs)	CONTACT (mean monthly marketing messages)
1 (control)	$0.00	3.1~39.9	1.42~2.60
2 (VALUE)	$5.49~7.14 **($5.49 ~ 7.14)**	3.2~36.5	1.40~2.60
3 (CSERV)	$0.00	1.9~20.4 **(-1.2 ~ -19.5)**	1.42~2.62
4 (VALUE + CSERV)	$5.55~7.25 **($5.55 ~ 7.25)**	3.2~28.9 **(+0.1 ~ -11.0)**	1.42~2.65
5 (CONTACT)	$0.00	3.3~31.6	0.00~0.03 **(-1.42 ~ -2.57)**

6 (VALUE + CONTACT)	$5.51~7.04 **($5.51 ~ 7.04)**	3.9~30.4	0.00~0.04 **(-1.42 ~ -2.56)**
7 (CSERV + CONTACT)	$0.00	3.3~27.0 **(+0.2 ~ -12.9)**	0.01~0.04 **(-1.41 ~ -2.56)**
8 (VALUE + CSERV + CONTACT)	$5.51~7.20 **($5.51 ~ 7.20)**	3.6~22.8 **(+0.5 ~ -17.1)**	0.00~0.04 **(-1.42 ~ -2.56)**

Note: Results in () indicate variance of the treatment experience versus control group

4.4 Pretest and Posttest Survey Summaries

Results from specific pretest and posttest survey questions as they relate to the IVs and DVs are covered in the following sections of this chapter. Prior to the analysis of the results of the experiment on the IVs and DVs, this section summarizes the overall characteristics of the two deployments of the survey: pretest and posttest.

4.4.1 Pretest Survey Summary

Prior to the start of the experiment, the email survey described in Section 3.9 was deployed to the study participants in all treatment groups. Of the 35,000 participants in the study, the pretest survey was sent to 30,200 individuals (Table 4.14). The survey excluded those participants who had indicated that they did not wish to be contacted by the subject organization for the purposes of marketing messages or surveys (DNC/DNS), representing 13.7% of the total study population.

Following deployment of the survey, a total of 2,614 complete responses were received, with the number of responses for each treatment group ranging from 315 to 342. Table 4.14 shows the complete response rates by treatment group.

The determination of the sample size for this study was based, in part, on a target of 384 responses per treatment group, or a total of 3,072 overall (see Section 3.8.4). The actual number of complete responses to the survey falls slightly short of this target, both in the aggregate and for each treatment group. The impact of the slightly smaller sample sizes is an increase in the margin of error for any results obtained from the survey. Using the smallest response rate obtained (n = 315, for group 7), the margin of error is increased from the target of 5.0% (as used in the sampling strategy in Section 3.8.4) to approximately 5.5%. This higher margin of error will be taken into account in the analysis of the data and is not anticipated to materially impact the results obtained from the study.

Table 4.14 Pretest Survey Deployment Results

Treatment Group	# Participants in Group	DNC/DNS Count	DNC/DNS Rate	Surveys Sent	Complete Responses	Complete Response Rate
1	4,400	721	16.4%	3,679	326	8.9%
2	4,284	576	13.4%	3,708	324	8.7%
3	4,397	602	13.7%	3,795	324	8.5%
4	4,335	607	14.0%	3,728	316	8.5%
5	4,412	565	12.8%	3,847	342	8.9%
6	4,384	602	13.7%	3,782	330	8.7%
7	4,404	601	13.6%	3,803	315	8.3%
8	4,384	616	14.1%	3,768	337	8.9%
Total	35,000	4,800	13.7%	30,200	2,614	8.7%

4.4.2 Posttest Survey Summary

At the conclusion of the experiment, the same email survey utilized in the pretest survey was re-deployed to the study participants in all treatment groups. Of the original 35,000 participants in the study 31,344 participants remained by the conclusion twelve months later. The posttest survey was sent to 25,840 participants (Table 4.15), with a DNC/DNS rate of 17.6% overall.

The complete response sample size for the posttest survey was n = 1,379 with the number of responses for each treatment group ranging from 136 to 195. The posttest survey was sent to study participants in early April 2020, following the conclusion of the experiment, but at the height of the COVID-19 pandemic shutdown in Canada. It is expected that the impact of the pandemic suppressed survey responses at this time of upheaval in the lives of the participants. The survey was deployed regardless, as the experimental treatments were ending as planned on March 31, 2020 and could not be extended.

Of the 1,379 respondents to the post-test survey, 457 also responded to the pre-test survey. The post-test statistical analysis presented here is conducted on the full set of 1,379 responses as these are representative of each treatment group, regardless of whether a

study participant answered both (or any) surveys[22].

Similar to the evaluation of the smaller than planned sample sizes for the pretest survey, the impact of the smaller sample sizes on the posttest survey data can be determined based on the impact of the margin of error that now accompanies the data. Based on the smallest sample size of 136 (for group 3), the margin of error has increased from the planned 5.0% to 8.4%.

Table 4.15 Posttest Survey Deployment Results

Treatment Group	# Participants in Group	DNC/DNS Count	DNC/DNS Rate	Surveys Sent	Complete Responses	Complete Response Rate
1	3,925	680	17.3%	3.245	156	4.8%
2	3,871	709	18.3%	3.162	164	5.2%
3	3,915	719	18.4%	3,196	136	4.3%
4	3,901	697	17.9%	3,204	189	5.9%
5	3,895	681	17.5%	3,214	172	5.4%
6	3,962	688	17.4%	3,274	192	5.9%
7	3,933	652	16.6%	3,281	175	5.3%
8	3,942	678	17.2%	3,264	195	6.0%
Total	31,344	5,504	17.6%	25,840	1,379	5.3%

4.4.3 Attrition Analysis

Attrition of participants in the study group was expected, and indeed, is the focus of the study through measurement of defection. If there were no attrition, there would be no participant behaviour to study.

Section 1.1 described the context of the mobile market in Canada, indicating that the competitive nature of the market was a key driver of customer defections from one carrier to another. Upon deactivation of service with the subject organization, a reason for the deactivation is recorded. The largest percentage of study participants (61%) deactivated in

[22] Analysis of posttest survey data for study participants that responded to both the pretest and posttest surveys yielded similar statistical results as those that only answered the posttest survey.

order to move their service to another mobile provider, as shown in Table 4.16. The second highest reason for deactivation was termination of service by the subject organization for non-payment by the participant (referred to as involuntary churn in Section 1.1.2).

Table 4.16 Reason for Participant Deactivation

Deactivation Reason	# Participants Who Deactivated	% of Participants Who Deactivated
Ported (moved) service to another mobile carrier	2,360	60.7%
Non-payment	524	13.5%
No longer need or use	485	12.5%
Deceased	137	3.5%
Left network area	66	1.7%
Service now provided by employer	58	1.5%
Financial / cost issues	52	1.3%
Other reason given	204	5.2%
Total	3,886	100%

Note: Deactivation Reason data was not available for study participants that deactivated in March 2020 (total 168 participants), accounting for the difference between the total shown above and total deactivations reported during the study period of 4,054.

4.4.4 Attrition Bias in Survey Results

Although attrition was planned for (and expected) throughout the experiment timeframe, it is possible that the post-test survey results explored in the following sections may be biased by the fact that certain groups of participants chose to stay, whilst others chose to leave, in part based on the impact of the experiment treatments. In addition, participants that remained at the end of the study may have been influenced to respond to the post-test survey (or not) as a result of their experience with the subject organization (including the treatments).

The interpretation of the post-test survey results must therefore be viewed with an understanding of these potential underlying biases in the results.

It is important to note, however, that the issue of attrition bias does not impact the analysis of the most important dependent variable in this study, that of the deactivation rate for each treatment group. The decision of a study participant to deactivate their service at some point

in the study (whether by moving to another mobile carrier or otherwise) is precisely the behaviour to be examined in answering the research problem.

4.5 Analysis of Independent Variables

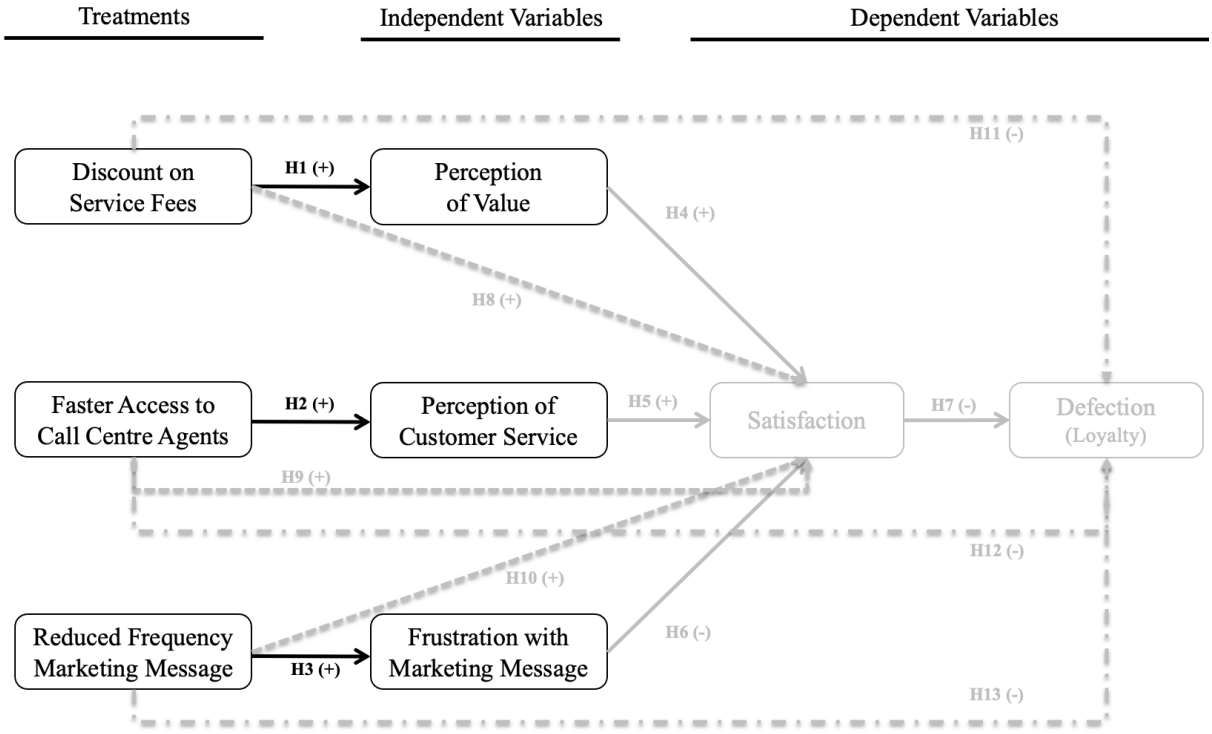
Three independent variables (IVs) were identified in the model proposed in Section 3.3. These three variables included:

PERCVALUE	Perception of Value
SERVICE	Perception of Customer Service
MARKETING	Frustration with Marketing Message

Section 3.9.4 defined the data collection procedure for each of these IVs, namely the use of a pretest and posttest survey of study participants. This section of the study presents the findings of the pretest levels for each IV first, followed by the posttest levels as measured after implementation of the treatments over the 12-month study period.

Of the statistical analysis methods identified in Section 4.2, one-way ANOVA is used to validate that the pre-test response is statistically the same across all treatment groups prior to the start of the experiment and is repeated post-test to measure the statistical significance of any impact of the treatments on the IVs. Regression analyses are also conducted on each of the three IVs, and the results of the hypothesis tests for H1, H2 and H3 (as shown in Figure 18) are provided.

Figure 18 Hypotheses for Relationships Between Treatments and IVs



4.5.1 Pre-test ANOVA Analysis

The purpose of the pre-test ANOVA analysis is to establish whether the values of three IVs are statistically the same for each of the eight treatment groups prior to the commencement of the experiment. The successful execution of randomized assignment of participants to each treatment group is expected to yield similar IV values for each group (prior to the commencement of the experiment).

4.5.1.1 Pre-test Results: PERCVALUE Independent Variable

The independent variable PERCVALUE was measured in the pretest and posttest surveys with the following question:

> "Considering the overall quality and the overall price you pay, how would you rate (subject organization) for overall value in terms of being worth what you pay for?"

Overall responses to the question at the time of the pretest survey are shown in Table 4.17. The mean PERCVALUE score across the 2,562 respondents was 3.19, with means varying from 3.09 to 3.28 across treatment groups (Table 4.18).

A one-way ANOVA analysis was completed on the survey results to determine if the mean values for the PERCVALUE variable are statistically the same prior to the commencement of the treatments. The one-way ANOVA results show a significance level of $p = .441$ (with $p >$

.05), indicating that the mean responses for PERCVALUE across the eight treatment groups are statistically the same prior to the start of the experiment (Table 4.19).

Table 4.17 Overall Pretest Responses for PERCVALUE

Response Value	Frequency	Percent
1 – Excellent	238	9.3%
2 – Very Good	543	21.2%
3 - Good	757	29.5%
4 – Fair	552	21.5%
5 – Poor	472	18.4%
Total	2,562	100.0%

Note: A total of 52 responses of “Unsure” were excluded, and do not contribute to the quantitative interpretation of the responses

Table 4.18 Mean Pretest PERCVALUE Results by Treatment Group

					95% Confidence Interval for Mean			
Treatment Group	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
1	320	3.16	1.217	.068	3.02	3.29	1	5
2	318	3.22	1.236	.069	3.08	3.36	1	5
3	316	3.27	1.195	.067	3.13	3.40	1	5
4	309	3.17	1.282	.073	3.02	3.31	1	5
5	334	3.09	1.201	.066	2.95	3.22	1	5
6	321	3.12	1.244	.069	2.99	3.26	1	5
7	310	3.19	1.228	.070	3.05	3.33	1	5
8	334	3.28	1.203	.066	3.15	3.41	1	5
Total	2,562	3.19	1.225	.024	3.14	3.23	1	5

Table 4.19 One-Way ANOVA Analysis - Pretest PERCVALUE

	Sum of Squares	dF	Mean Square	F	Sig.
Between Groups	10.345	7	1.478	.984	.441
Within Groups	3,835.846	2,554	1.502		
Total	3,846.191	2,561			

4.5.1.2 Pre-test Results: SERVICE Independent Variable

The pretest and posttest surveys included the following question in order to measure the independent variable SERVICE:

> "Thinking of your most recent contact with customer service over the phone, how would you rate your experience with the time you waited to speak to a representative?"

Overall responses to the question at the time of the pretest survey are shown in Table 4.20. Prior to being asked this question, survey respondents had to pre-qualify by indicating that they had contacted customer service within the past six months. Of the 2,614 pretest respondents, a total of 1,229 indicated they had done so. The mean SERVICE score across the 1,229 respondents was 3.42, with means varying from 3.32 to 3.49 across treatment groups (Table 4.21).

A one-way ANOVA analysis was conducted to determine if the mean values for the SERVICE variable are statistically the same prior to the commencement of the treatments. The one-way ANOVA results show a significance level of $p = .835$ (with $p > .05$), indicating that the mean responses for SERVICE across the eight treatment groups are not statistically different prior to the start of the experiment (Table 4.22).

Table 4.20 Overall Pretest Responses for SERVICE

Response Value	Frequency	Percent
1 – Excellent	83	6.8%
2 – Very Good	215	17.5%
3 - Good	325	26.4%
4 – Fair	315	25.6%

5 – Poor	291	23.7%
Total	1229	100.0%

Table 4.21 Mean Pretest SERVICE Results by Treatment Group

Treatment Group	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	152	3.34	1.282	.104	3.14	3.55	1	5
2	160	3.44	1.142	.090	3.26	3.62	1	5
3	158	3.48	1.204	.096	3.29	3.67	1	5
4	160	3.43	1.277	.101	3.23	3.63	1	5
5	166	3.49	1.184	.092	3.31	3.68	1	5
6	147	3.32	1.266	.104	3.11	3.53	1	5
7	146	3.36	1.213	.100	3.16	3.55	1	5
8	140	3.49	1.147	.097	3.29	3.68	1	5
Total	1,229	3.42	1.214	.035	3.35	3.49	1	5

Note: Groups 3, 4, 7 & 8 received the CSERV treatment

Table 4.22 One-Way ANOVA Analysis - Pretest SERVICE

	Sum of Squares	dF	Mean Square	F	Sig.
Between Groups	5.166	7	.738	.499	.835
Within Groups	1804.190	1221	1.478		
Total	1809.356	1228			

4.5.1.3 Pre-test Results: MARKETING Independent Variable

The third independent variable MARKETING was measured in the pretest and posttest surveys with the following question:

"How often are you frustrated with communication quality, e.g. relevance or frequency of emails, other marketing?"

Overall responses to the question at the time of the pretest survey are shown in Table 4.23. The mean MARKETING score across the 2,400 respondents was 3.88, with means varying from 3.86 to 3.92 across treatment groups (Table 4.24).

The one-way ANOVA results show a significance level of $p = .669$ (with $p > .05$), indicating that the mean responses for SERVICE across the eight treatment groups are not statistically different at the start of the experiment (Table 4.25).

Table 4.23 Overall Pretest Responses for MARKETING

Response Value	Frequency	Percent
1 – Very Often	88	3.7%
2 – Often	155	6.5%
3 – Sometimes	501	20.9%
4 – Rarely	869	36.2%
5 – Not at all	787	32.8%
Total	2,400	100.0%

Note: A total of 214 responses of "Unsure / Not Applicable" were excluded, and do not contribute to the quantitative interpretation of the responses

Table 4.24 Mean Pretest MARKETING Results by Treatment Group

					95% Confidence Interval for Mean			
Treatment Group	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
1	302	3.90	0.970	0.056	3.79	4.01	1	5
2	299	3.87	1.101	0.064	3.75	4.00	1	5
3	297	3.86	1.111	0.064	3.74	3.99	1	5
4	277	3.79	1.114	0.067	3.66	3.92	1	5
5	320	3.86	1.034	0.058	3.74	3.97	1	5

6	308	3.96	1.031	0.059	3.85	4.08	1	5
7	286	3.92	1.058	0.063	3.80	4.04	1	5
8	311	3.87	1.014	0.058	3.76	3.98	1	5
Total	2,400	3.88	1.054	0.022	3.84	3.92	1	5

Note: Groups 5, 6, 7 & 8 received the CONTACT treatment

Table 4.25 One-Way ANOVA Analysis - Pretest MARKETING

	Sum of Squares	dF	Mean Square	F	Sig.
Between Groups	5.476	7	0.782	0.073	.669
Within Groups	2659.964	2392	1.112		
Total	2665.440	2399			

4.5.1.4 Pre-test Results - Summary

The results of the pre-test ANOVA analysis for all three independent variables indicate that there were no responses - across any of the treatment groups - that were statistically different from the other treatment groups prior to the commencement of the experiment (Table 4.26). These results validate that the random selection and random assignment of participants to the various treatment groups was successful, and that any differences observed following the completion of the experiment may be attributed to the treatments. The post-test measure of the three IVs is reviewed in the following section.

Table 4.26 Summary of Pre-test Results

Independent Variable	Mean Score	One-way ANOVA Significance (*p*)	Statistical Difference Among Groups?
PERCVALUE	3.19	.441	No
SERVICE	3.42	.835	No
MARKETING	3.88	.669	No

4.5.2 Post-test ANOVA & Regression Analyses

Following the completion of the experiment, the three independent variables were measured using the same survey questions deployed for the pre-test, and analyzed with the same pre-test statistical method (One-way ANOVA) as well as with a regression analysis for each. The results for each IV are reviewed in turn in the following sub-sections.

4.5.2.1 Posttest ANOVA Results - PERCVALUE

Overall responses to the question at the time of the posttest survey are shown in Table 4.27. The mean PERCVALUE score across the 1,358 respondents was 3.00, with means varying from 2.89 to 3.12 across treatment groups (Table 4.28).

As with the pre-test, a one-way ANOVA analysis was conducted to determine if the mean values for the PERCVALUE variable are statistically the same for each treatment group at the conclusion of the treatments. The one-way ANOVA results show a significance level of $p = .559$ (with $p>.05$), indicating that the mean responses for PERCVALUE across the eight treatment groups are not statistically different at the conclusion of the experiment (Table 4.29).

Table 4.27 Overall Posttest Responses for PERCVALUE

Response Value	Frequency	Percent
1 – Excellent	143	10.5%
2 – Very Good	337	24.8%
3 - Good	421	31.0%
4 – Fair	285	21.0%
5 – Poor	172	12.7%
Total	1,358	100.0%

Note: A total of 21 responses of “Unsure” were excluded, and do not contribute to the quantitative interpretation of the responses

Table 4.28 Mean Posttest PERCVALUE Results by Treatment Group

Treatment Group	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	153	3.10	1.213	.098	2.90	3.29	1	5
2	163	3.05	1.191	.093	2.86	3.23	1	5
3	133	2.95	1.167	.101	2.75	3.16	1	5
4	185	2.89	1.248	.092	2.71	3.07	1	5
5	171	3.12	1.137	.087	2.95	3.29	1	5
6	188	2.94	1.143	.083	2.78	3.11	1	5
7	173	3.05	1.120	.085	2.88	3.21	1	5
8	192	2.96	1.199	.087	2.79	3.13	1	5
Total	1358	3.00	1.178	.032	2.94	3.07	1	5

Note: Groups 2, 4, 6 & 8 received the VALUE treatment

Table 4.29 One-Way ANOVA Analysis - Posttest PERCVALUE

	Sum of Squares	dF	Mean Square	F	Sig.
Between Groups	8.099	7	1.157	.834	.559
Within Groups	1873.874	1350	1.388		
Total	1881.973	1357			

4.5.2.2 Posttest Regression Analysis - PERCVALUE

Linear regression analysis was conducted to model the relationship between the treatments utilized in the experiment and the independent variables, and to show the degree (and strength) to which post-test variations in the IVs can be attributed to the treatments. Recall that analysis of the pre-test results indicated that there was no statistically significant difference in the levels of the IVs prior to the commencement of the study (Section 4.4.1).

Therefore, all other things being equal, any statistically significant different results at the end of the study may be ascribed to the treatments.

Dummy variables were created to indicate whether a particular participant received one or more of the three treatments in the study. These dummy variables represent each of the seven treatment combinations that could be received by the study participants (plus the control group, which received no treatments), as shown in Table 4.30. The dummy treatment variables are used in multiple regression analyses throughout the rest of this chapter.

Table 4.30 Dummy Variables for Treatment Combinations

Treatment Combination	Dummy Variable and Assigned Value
CONTROL	dum_CONTROL = 1
VALUE	dum_VALUE = 1
CSERV	dum_CSERV = 1
VALUE + CSERV	dum_VALUE_CSERV = 1
CONTACT	dum_CONTACT = 1
VALUE + CONTACT	dum_VALUE_CONTACT = 1
CSERV + CONTACT	dum_CSERV_CONTACT = 1
VALUE + CSERV + CONTACT	dum_VALUE_CSERV_CONTACT = 1

The model that is examined in the regression analysis for the PERCVALUE IV is shown in (4.1).

$$PERCVALUE_i = b_0 + b_1 VALUE + b_2 CSERV + b_3 VALUE\_CSERV + b_4 CONTACT + b_5 VALUE\_CONTACT + b_6 CSERV\_CONTACT + b_7 VALUE\_CSERV\_CONTACT \quad (4.1)$$

This model defines the outcome of the prediction as $PERCVALUE_i$ with regression coefficients for each of the predictors $VALUE$, $CSERV$, $VALUE\_CSERV$, $CONTACT$, $VALUE\_CONTACT$, $CSERV\_CONTACT$ and $VALUE\_CSERV\_CONTACT$. The predictors correspond to the seven treatment combinations utilized in the study. The parameters b_1, b_2, b_3, b_4, b_5, b_6 and b_7 are associated with each of the seven treatments, respectively. The parameter b_0 is the constant in the model and represents the expected level of PERCVALUE for the control group, without

any treatment effects. The model would predict the level of perceived value ($PERCVALUE_i$) for a population as an outcome of whether they were included in one of the seven treatment combinations (that is, whether b_1, b_2 b_3, b_4, b_5, b_6 and b_7 are 1 or not).

The descriptive statistics resulting from the linear regression are shown in Table 4.31, and show the expected values for the PERCVALUE IV overall (mean = 3.00, N = 1358), consistent with the ANOVA analysis in Section 4.5.2.1. The mean values for each of the dummy variables is consistent with the representation of each treatment combination assigned to study participants (that is, approximately one out of eight participants for each treatment, including the control group, or 12.5% each).

Table 4.31 Descriptive Statistics - Linear Regression Analysis of PERCVALUE

	Mean	Std. Deviation	N
PERCVALUE	3.00	1.178	1358
VALUE	0.12	.325	1358
CSERV	0.10	.297	1358
VALUE_CSERV	0.14	.343	1358
CONTACT	0.13	.332	1358
VALUE_CONTACT	0.14	.345	1358
CSERV_CONTACT	0.13	.334	1358
VALUE_CSERV_CONTACT	0.14	.349	1358

The summary results of the multiple linear regression model are shown in Table 4.32, and the regression model coefficients are shown in Table 4.33. A non-significant regression equation was found ($F(7, 1350) = .834$, $p > .05$), with an R2 of .004 and an adjusted R2 of -.001.

While none of the treatment variables were determined to be a statistically significant predictor of PERCVALUE on their own, an F-test was conducted to determine if any of the treatments are significant in combination. The null hypothesis established for the F-test was that none of the treatments in combination are statistically significant in predicting PERCVALUE. This null hypothesis would be rejected if the F-statistic is greater than the Critical F-Value in the equation. In this case, the analysis yielded an F-statistic of 0.72, which is not greater than the associated Critical F-Value of 2.01635. We fail to reject the null

hypothesis and conclude that the treatments do not jointly predict the PERCVALUE variable in a statistically significant manner.

Table 4.32 Regression Model Summary – PERCVALUE IV

R	R Square	Adj. R Square	Std. Error of Est.	Change Statistics				
				R Square Change	F Change	df1	df2	Sig. F Change
.066[a]	.004	-.001	1.178	.004	.834	7	1350	.559

a. Dependent Variable: PERCVALUE; Independent Variables: VALUE, CSERV, VALUE_CSERV, CONTACT, VALUE_CONTACT, CSERV_CONTACT and VALUE_CSERV_CONTACT

Table 4.33 Regression Model Coefficients – PERCVALUE IV

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations		
	B	Std Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
(Constant)	3.098	.095		35.526	<.001	2.911	3.285			
dum_VALUE	-.049	.133	-.014	-.369	.712	-.309	.211	.014	-.010	-.010
dum_CSERV	-.143	.140	-.036	-1.025	.306	-.417	.131	-.014	-.028	-.028
dum_VALUE_CSERV	-.212	.129	-.062	-1.643	.101	-.464	.041	-.040	-.045	-.045
dum_CONTACT	.019	.131	.005	.144	.885	-.238	.276	.036	.004	.004
dum_VALUE_CONTACT	-.157	.128	-.046	-1.220	.223	-.408	.095	-.021	-.033	-.033
dum_CSERV_CONTACT	-.052	.131	-.015	-.396	.692	-.308	.205	.014	-.011	-.011
dum_VALUE_CSERV_CONTACT	-.134	.128	-.040	-1.053	.292	-.385	.116	-.014	-.029	-.029

4.5.2.3 Posttest ANOVA Results - SERVICE

Overall responses to the SERVICE question at the time of the posttest survey are shown in Table 4.34. The mean SERVICE score across the 612 respondents was 3.09, with means varying from 3.03 to 3.17 across treatment groups (Table 4.35).

The posttest one-way ANOVA results show a significance level of $p = .997$ (with $p > .05$), indicating that the mean responses for SERVICE across the eight treatment groups are not statistically different at the conclusion of the experiment (Table 4.36).

Table 4.34 Overall Posttest Responses for SERVICE

Response Value	Frequency	Percent
1 – Excellent	72	11.8%
2 – Very Good	126	20.6%
3 - Good	172	28.1%
4 – Fair	156	25.5%
5 – Poor	86	14.1%
Total	612	100.0%

Table 4.35 Mean Posttest SERVICE Results by Treatment Group

					95% Confidence Interval for Mean			
Treatment Group	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
1	79	3.10	1.194	.134	2.83	3.37	1	5
2	78	3.08	1.336	.151	2.78	3.38	1	5
3	68	3.03	1.065	.129	2.77	3.29	1	5
4	85	3.15	1.220	.132	2.89	3.42	1	5
5	65	3.17	1.282	.159	2.85	3.49	1	5
6	85	3.11	1.102	.120	2.87	3.34	1	5
7	70	3.09	1.316	.157	2.77	3.40	1	5
8	82	3.04	1.271	.140	2.76	3.32	1	5
Total	612	3.09	1.219	.049	3.00	3.19	1	5

Note: Groups 3, 4, 7 & 8 received the CSERV treatment

Table 4.36 One-Way ANOVA Analysis - Posttest SERVICE

	Sum of Squares	dF	Mean Square	F	Sig.
Between Groups	1.261	7	.180	.120	.997
Within Groups	907.243	604	1.502		
Total	908.503	611			

4.5.2.4 Posttest Regression Analysis - SERVICE

The model that is examined in the regression analysis for the SERVICE IV is shown in (4.2).

$$SERVICE_i = b_0 + b_1 VALUE + b_2 CSERV + b_3 VALUE\_CSERV + b_4 CONTACT + b_5 VALUE\_CONTACT + b_6 CSERV\_CONTACT + b_7 VALUE\_CSERV\_CONTACT \quad (4.2)$$

The descriptive statistics resulting from the linear regression are shown in Table 4.37, and show the expected values for the SERVICE IV overall (mean = 3.09, N = 612), consistent with the ANOVA analysis in Section 4.5.2.3.

Table 4.37 Descriptive Statistics - Linear Regression Analysis of SERVICE

	Mean	Std. Deviation	N
PERCVALUE	3.09	1.219	612
VALUE	.13	.334	612
CSERV	.11	.315	612
VALUE_CSERV	.14	.346	612
CONTACT	.11	.308	612
VALUE_CONTACT	.14	.346	612
CSERV_CONTACT	.11	.319	612
VALUE_CSERV_CONTACT	.13	.341	612

The summary results of the multiple linear regression model are shown in Table 4.38, with the regression model coefficients shown in Table 4.39. A non-significant regression equation was found (F(7, 604) = .120, $p > .05$), with an R^2 of .001 and an adjusted R^2 of -.010.

F-test results yielded an F-statistic of 0.095, which was not greater than the Critical F-Value of 2.02472. As a result, it is also concluded that the treatments do not jointly predict SERVICE in a statistically significant manner.

Similar to the outcome of the regression analysis of the PERCVALUE IV, none of the treatment variables were a significant predictor of SERVICE and therefore subsequent analysis of the results is not conducted.

Table 4.38 Regression Model Summary – SERVICE IV

				Change Statistics				
R	R Square	Adj. R Square	Std. Error of Est.	R Square Change	F Change	df1	df2	Sig. F Change
.037[a]	.001	-.010	1.226	.001	.120	7	604	.997

a. Dependent Variable: SERVICE; Independent Variables: VALUE, CSERV, VALUE_CSERV, CONTACT, VALUE_CONTACT, CSERV_CONTACT and VALUE_CSERV_CONTACT

Table 4.39 Regression Model Coefficients – SERVICE IV

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations		
	B	Std Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
(Constant)	3.101	.138		22.491	<.001	2.830	3.372			
dum_VALUE	-.024	.196	-.007	-.124	.901	-.409	.360	-.006	-.005	-.005
dum_CSERV	-.072	.203	-.019	-.354	.723	-.470	.326	-.019	-.014	-.014
dum_VALUE_CSERV	.052	.192	.015	.270	.787	-.324	.428	.019	.011	.011
dum_CONTACT	.068	.205	.017	.331	.741	-.335	.471	.021	.013	.013
dum_VALUE_CONTACT	.005	.192	.001	.024	.981	-.372	.381	.004	.001	.001
dum_CSERV_CONTACT	-.016	.201	-.004	-.077	.938	-.411	.380	-.003	-.003	-.003
dum_VALUE_CSERV_CONTACT	-.065	.193	-.018	-.335	.738	-.444	.315	-.019	-.014	-.014

4.5.2.5 Posttest ANOVA Results - MARKETING

Overall responses to the question at the time of the posttest survey are shown in Table 4.40. The mean MARKETING score across the 1,285 respondents was 3.98, with means varying from 3.76 to 4.12 across treatment groups (Table 4.41).

The one-way ANOVA results show a significance level of $p = .003$ (with $p < .05$), indicating that the mean responses for MARKETING are statistically different across at least one or more of the eight treatment groups at the conclusion of the experiment (Table 4.42).

Additional post hoc analysis is required to determine which treatment groups have a significant difference in their mean MARKETING values.

Table 4.40 Overall Posttest Responses for MARKETING

Response Value	Frequency	Percent
1 – Very Often	41	3.2%
2 – Often	72	5.6%
3 – Sometimes	245	19.1%
4 – Rarely	444	34.6%
5 – Not at all	483	37.6%
Total	1,285	100.0%

Note: A total of 94 responses of “Unsure / Not applicable” were excluded, and do not contribute to the quantitative interpretation of the responses

Table 4.41 Mean Posttest MARKETING Results by Treatment Group

Treatment Group	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	149	3.89	1.112	.091	3.71	4.07	1	5
2	153	3.86	1.124	.091	3.68	4.04	1	5
3	127	4.05	.925	.082	3.88	4.21	1	5

4	179	3.76	1.098	.082	3.60	3.92	1	5
5	161	3.91	1.057	.083	3.75	4.08	1	5
6	179	4.11	.923	.069	3.98	4.25	1	5
7	158	4.12	.986	.078	3.97	4.28	1	5
8	179	4.12	1.007	.075	3.97	4.27	1	5
Total	1,285	3.98	1.038	.029	3.92	4.03	1	5

Note: Groups 5, 6, 7 & 8 received the CONTACT treatment

Table 4.42 One-Way ANOVA Analysis - Posttest MARKETING

	Sum of Squares	dF	Mean Square	F	Sig.
Between Groups	22.981	7	3.283	3.080	.003
Within Groups	1361.364	1277	1.066		
Total	1384.346	1284			

A post hoc analysis using Tukey was completed to determine which treatment groups exhibit a significant difference in means for the MARKETING IV following the completion of the experiment. This analysis determined the subsets of treatment groups that exhibit a significant difference in reported means.

Table 4.43 shows the results of the Tukey post hoc analysis (with additional details in Appendix F), showing a statistical difference in the MARKETING posttest results between treatment group 4 and groups 6, 7 and 8. There was a statistically significant difference between groups as determined by one-way ANOVA ($F(7,1277) = 3.080$, $p = .003$).

The Tukey posthoc test revealed that the mean MARKETING response was statistically higher for group 6: VALUE+CONTACT (4.11 ±0.923, $p = .028$), group 7: CSERV+CONTACT (4.12 ±0.986, $p = .031$) and group 8: VALUE+CSERV+CONTACT (4.12 ±1.007, $p = .024$) compared to group 4: VALUE+CSERV (3.76 ±1.098). There was no statistically significant difference between the other treatment groups (p values equal to .242 to 1.000).

Table 4.43 Tukey Post Hoc Analysis – MARKETING

		Subset for alpha = 0.05	
Treatment Group	**N**	**1**	**2**
4: VALUE + CSERV	179	3.76	
2: VALUE	153	3.86	3.86
1: CONTROL	149	3.89	3.89
5: CONTACT	161	3.91	3.91
3: CSERV	127	4.05	4.05
6: VALUE + CONTACT	179		4.11
8: VALUE + CSERV + CONTACT	179		4.12
7: CSERV + CONTACT	158		4.12
Sig.		.205	.339

Note: Groups 5, 6, 7 & 8 received the CONTACT treatment

4.5.2.6 Posttest Regression Analysis - MARKETING

The model that is examined in the regression analysis for the MARKETING IV is shown in (4.3).

$$MARKETING_i = b_0 + b_1 VALUE + b_2 CSERV + b_3 VALUE\_CSERV + b_4 CONTACT + b_5 VALUE\_CONTACT + b_6 CSERV\_CONTACT + b_7 VALUE\_CSERV\_CONTACT \quad (4.3)$$

The descriptive statistics resulting from the linear regression are shown in Table 4.44 and show the expected values for the MARKETING IV overall (mean = 3.98, N = 1285), consistent with the ANOVA analysis in Section 4.5.2.5.

Table 4.44 Descriptive Statistics - Linear Regression Analysis of MARKETING

	Mean	Std. Deviation	N
PERCVALUE	3.98	1.038	1285
VALUE	0.12	.324	1285
CSERV	0.10	.299	1285
VALUE_CSERV	0.14	.346	1285
CONTACT	0.13	.331	1285
VALUE_CONTACT	0.14	.346	1285
CSERV_CONTACT	0.12	.329	1285
VALUE_CSERV_CONTACT	0.14	.346	1285

The summary results of the multiple linear regression model are shown in Table 4.45, and the regression model coefficients are shown in Table 4.46.

A significant regression equation was found (F(7, 1277) = 3.080, $p < .05$), with an R^2 of .017 and an adjusted R^2 of .011. As the results were significant, additional interpretation is warranted.

In addition, the F-test results indicate that the treatments jointly are statistically significant in predicting the variable MARKETING. The F-statistic is 2.2544, which is greater than the Critical F-Value of 2.01629. This indicates that the null hypothesis posed in the F-test (that the treatments are not jointly statistically significant in predicting MARKETING) is rejected.

Table 4.45 Regression Model Summary – MARKETING IV

				Change Statistics				
R	R Square	Adj. R Square	Std. Error of Est.	R Square Change	F Change	df1	df2	Sig. F Change
.129[a]	.017	.011	1.033	.017	3.080	7	1277	.003

a. Dependent Variable: MARKETING; Independent Variables: VALUE, CSERV, VALUE_CSERV, CONTACT, VALUE_CONTACT, CSERV_CONTACT and VALUE_CSERV_CONTACT

Table 4.46 Regression Model Coefficients – MARKETING IV

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations		
	B	Std Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
(Constant)	3.886	.085		45.940	<.001	3.720	4.052			
dum_VALUE	-.023	.119	-.007	-.195	.846	-.256	.210	-.041	-.005	-.005
dum_CSERV	.161	.125	.046	1.294	.196	-.083	.406	.022	.036	.036
dum_VALUE_CSERV	-.126	.115	-.042	-1.102	.271	-.351	.099	-.084	-.031	-.031
dum_CONTACT	.027	.117	.009	.231	.817	-.203	.257	-.023	.006	.006
dum_VALUE_CONTACT	.226	.115	.075	1.972	.049	.001	.450	.052	.055	.055
dum_CSERV_CONTACT	.234	.118	.074	1.988	.047	.003	.466	.052	.056	.055
dum_VALUE_CSERV_CONTACT	.231	.115	.077	2.021	.043	.007	.456	.054	.056	.056

Replacing the *b*-values in the model (4.3) with the values shown in Table 4.46, the model estimating the relationship between the predictors and the outcome of SERVICE then becomes Equation (4.4):

$$\begin{aligned} MARKETING_i = {} & 3.886 - (0.023\ x\ VALUE) + (0.161\ x\ CSERV) \\ & - (0.126\ x\ VALUE\_CSERV) + (0.027\ x\ CONTACT) \\ & + (0.226\ x\ VALUE\_CONTACT) \\ & + (0.234\ x\ CSERV\_CONTACT) \\ & + (0.231\ x\ VALUE\_CSERV\_CONTACT) \end{aligned} \tag{4.4}$$

The degree to which each predictor affects the independent variable (when the effects of all other independent variables are held constant) can be summarized as:

- **VALUE (*b* = -0.023, *p* = .846 > .05):** This predictor is not statistically significant
- **CSERV (*b* = 0.161, *p* = .196 > .05):** This predictor is not statistically significant
- **VALUE_CSERV (*b* = -0.126, *p* = .271 > .05):** This predictor is not statistically significant
- **CONTACT (*b* = 0.027, *p* = .817 > .05):** This predictor is not statistically significant

- **VALUE_CONTACT (*b* = 0.226, *p* = .049 < .05):** This predictor is statistically significant
- **CSERV_CONTACT (*b* = 0.234, *p* = .047 < .05):** This predictor is statistically significant
- **VALUE_CSERV_CONTACT (*b* = 0.231, *p* = .043 < .05):** This predictor is statistically significant

Participants' predicted level of frustration with marketing messages (MARKETING) is 3.886 + .226 (VALUE_CONTACT) + .234 (CSERV_CONTACT) + .231 (VALUE_CSERV_CONTACT). Participant's level of frustration deceases (i.e. a higher rating on the survey question, where 1 = "very often" frustrated and 5 = "not at all" frustrated) by .226 if they received both the VALUE and CONTACT treatments, by .234 if they received the CSERV and CONTACT treatments, and by .231 if they received all three treatments.

4.5.3 Summary of Analysis of the Independent Variables

The analysis completed on the three IVs used in the study (PERCVALUE, SERVICE and MARKETING) showed that statistically significant post-experiment differences among treatment groups was found only in the case of the MARKETING IV.

The outcome of the hypothesis tests for these three IVs is summarized in Table 4.47.

Table 4.47 Summary of IV Hypothesis Outcomes

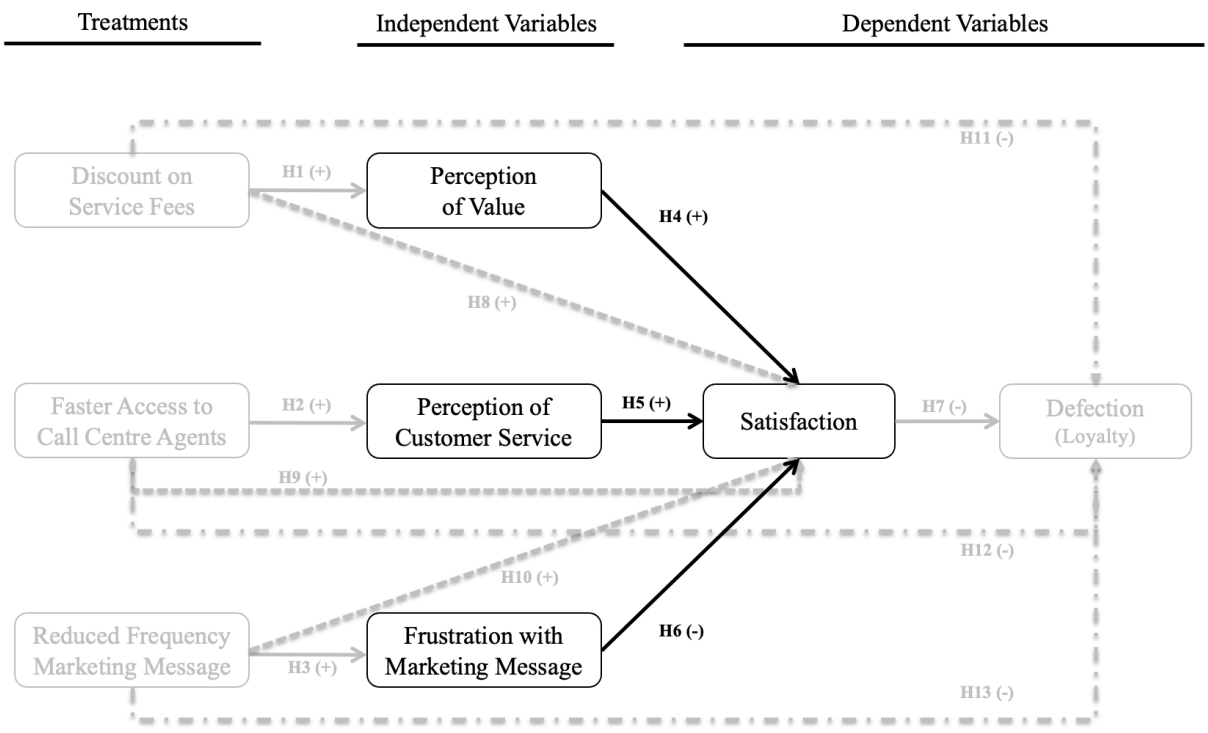
Hypothesis Number	Null Hypothesis Statement	Outcome
H_{01}	***There is no relationship between receiving a discount on monthly service fees and a customer's perception of value.***	Cannot be rejected
H_{02}	***There is no relationship between faster access to call centre agents and a customer's perception of customer service quality.***	Cannot be rejected
H_{03}	***There is no relationship between reduced frequency of marketing messages and a customer's frustration with receiving such marketing messaging.***	Rejected

4.6 Analysis of CSAT Dependent Variable

Following the analysis methods set out in Section 4.2, this section presents the results of the experiment on the CSAT dependent variable (DV). As with the analysis of the independent variables in the preceding sections, the analysis of the CSAT DV begins with an examination of the DV levels prior to the commencement of the experiment, followed by the same analysis post-experiment. Both ANOVA and linear regression methods are used to test the impact, if any, of the treatments and changes in the IVs on the CSAT DV. A difference-in-difference analysis is also conducted, and the outcomes compared to the results from the ANOVA and linear regression analyses.

The section concludes with a review of the hypothesis statements proposed in the study model and highlighted in Figure 19.

Figure 19 Hypotheses for Relationships Between IVs and CSAT DV



The dependent variable CSAT was measured in the pretest and posttest surveys with the following question:

> "If a colleague, friend or family member were looking for a new wireless provider, what is the likelihood that you would **recommend** (subject organization) to them?"

The pretest results are presented first, to establish whether or not the values for this DV were statistically the same prior to the start of the treatments; posttest results are then shown to determine if there were any statistically significant changes between the study groups during the experiment.

4.6.1 Pretest ANOVA Analysis

Overall responses to the survey question measuring the CSAT DV are shown in

Table 4.48. The mean CSAT value across the 2,614 respondents was 2.06, with means varying from 2.01 to 2.10 across treatment groups (Table 4.49).

One-way ANOVA results are presented in Table 4.50. The results of this analysis show a significance level of $p = .939$ ($p > .05$), indicating the mean responses for CSAT across the eight treatment groups are not statistically different prior to the start of the treatments.

Table 4.48 Overall Pretest Responses for CSAT

Response Value	Frequency	Percent
1 – Definitely	946	36.6%
2 – Probably	904	34.9%
3 – Maybe	473	18.3%
4 – Probably Not	170	6.6%
5 – Definitely Not	92	3.6%
Total	2,585	100.0%

Note: A total of 26 responses of “Don’t Know” were excluded, and do not contribute to the quantitative interpretation of the responses

Table 4.49 Mean Pretest CSAT Results by Treatment Group

					95% Confidence Interval for Mean			
Treatment Group	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
1	320	2.05	1.024	.057	1.93	2.16	1	5
2	321	2.10	1.098	.061	1.98	2.22	1	5

3	320	2.08	1.106	.062	1.96	2.20	1	5
4	314	2.08	1.111	.063	1.95	2.20	1	5
5	337	2.01	1.039	.057	1.90	2.12	1	5
6	328	2.02	1.043	.058	1.90	2.13	1	5
7	311	2.03	1.076	.061	1.91	2.15	1	5
8	334	2.09	1.024	.055	1.98	2.20	1	5
Total	2,585	2.06	1.063	.021	2.01	2.10	1	5

Table 4.50 One-Way ANOVA Analysis - Pretest CSAT

	Sum of Squares	dF	Mean Square	F	Sig.
Between Groups	2.648	7	.378	.334	.939
Within Groups	2916.441	2577	1.132		
Total	2919.089	2584			

4.6.2 Posttest ANOVA Analysis

The responses to the survey question for the CSAT DV at posttest are shown in Table 4.51. The mean CSAT score across the 1,369 respondents was 1.85, with means varying from 1.80 to 1.90 across the treatment groups (Table 4.52).

One-way ANOVA results are presented in Table 4.53 and show a significance level of p = .942 (p > .05), indicating the mean responses for CSAT across the eight treatment groups are not statistically different following the conclusion of the treatments.

Table 4.51 Overall Posttest Responses for CSAT

Response Value	Frequency	Percent
1 – Definitely	605	44.2%
2 – Probably	479	35.0%
3 – Maybe	197	14.4%
4 – Probably Not	62	4.5%

5 – Definitely Not	26	1.9%
Total	1,369	100.0%

Note: A total of 10 responses of “Don’t Know” were excluded, and do not contribute to the quantitative interpretation of the responses

Table 4.52 Mean Posttest CSAT Results by Treatment Group

					95% Confidence Interval for Mean			
Treatment Group	**N**	**Mean**	**Std. Deviation**	**Std. Error**	**Lower Bound**	**Upper Bound**	**Minimum**	**Maximum**
1	165	1.88	.980	.079	1.73	2.04	1	5
2	163	1.90	.957	.075	1.75	2.05	1	5
3	136	1.84	.871	.075	1.69	1.99	1	5
4	188	1.84	.930	.068	1.70	1.97	1	5
5	170	1.86	.912	.070	1.72	2.00	1	5
6	191	1.80	.987	.071	1.65	1.94	1	5
7	175	1.90	1.062	.080	1.74	2.06	1	5
8	191	1.80	.942	.068	1.67	1.94	1	5
Total	1,369	1.85	.957	.026	1.80	1.90	1	5

Table 4.53 One-Way ANOVA Analysis - Posttest CSAT

	Sum of Squares	**dF**	**Mean Square**	**F**	**Sig.**
Between Groups	2.097	7	.300	.326	.942
Within Groups	1250.905	1361	.919		
Total	1253.002	1368			

4.6.3 Posttest Regression Analysis – PERCVALUE and CSAT

Three regression analyses are completed to evaluate the hypotheses describing the relationships between the three independent variables (PERCVALUE, SERVICE, MARKETING) and the dependent variable of CSAT. The outcome of each regression analysis is provided in turn below.

The first model examines the relationship between the PERCVALUE IV and the CSAT DV, and as shown in (4.5).

$$CSAT_i = b_0 + b_1 PERCVALUE \quad (4.5)$$

The descriptive statistics resulting from the linear regression are shown in Table 4.54, and show the expected values for the CSAT DV overall (mean = 1.84, N = 1349)[23].

Table 4.54 Descriptive Statistics - Linear Regression Analysis of CSAT DV

	Mean	Std. Deviation	N
CSAT	1.84	.957	1349
PERCVALUE	3.00	1.175	1349

The summary results of the multiple linear regression model are shown in Table 4.55, and the regression model coefficients are shown in Table 4.56. A significant regression equation was found (F(1, 1347) = 774.834, $p < .001$), with an R^2 of .365 and an adjusted R^2 of .365.

Table 4.55 Regression Model Summary – CSAT DV

				Change Statistics				
R	R Square	Adj. R Square	Std. Error of Est.	R Square Change	F Change	df1	df2	Sig. F Change
.604[a]	.365	.365	.763	.365	774.834	1	1347	<.001

a. Dependent Variable: CSAT; Independent Variable: PERCVALUE

[23] The value of N = 1349 represents the subset of data with valid survey responses for both the PERCVALUE and CSAT variables, which is slightly lower than 1358 and 1369 responses for the same variables independently.

Table 4.56 Regression Model Coefficients – CSAT DV

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations		
	B	Std Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
(Constant)	.368	.057		6.459	<.001	.256	.479			
PERCVALUE	.492	.018	.604	27.836	<.001	.457	.527	.604	.604	.604

Replacing the *b*-values in the model (4.5) with the values shown in Table 4.56, the model estimating the relationship between the predictors and the outcome of CSAT then becomes Equation(4.6):

$$CSAT_i = 0.368 + (0.492\ x\ PERCVALUE) \tag{4.6}$$

Participants' predicted level of satisfaction (CSAT) is .368 + .492 (PERCVALUE), where satisfaction is measured by likelihood to recommend (1 = definitely will to 5 = definitely will not) and perceived value is scored on a scale of 1 = excellent to 5 = poor. Participant's level of satisfaction increases (improves) by .492 for each one point better rating on perceived value. Perceived value is a significant predictor of satisfaction.

4.6.4 Posttest Regression Analysis – SERVICE and CSAT

The second model examines the relationship between the SERVICE IV and the CSAT DV, and as shown in (4.7).

$$CSAT_i = b_0 + b_1 SERVICE \tag{4.7}$$

The descriptive statistics resulting from the linear regression are shown in Table 4.57, and show the expected values for the CSAT DV overall (mean = 1.92, N = 609)[24].

Table 4.57 Descriptive Statistics - Linear Regression Analysis of CSAT DV

	Mean	Std. Deviation	N
CSAT	1.92	1.016	609
SERVICE	3.10	1.219	609

Summary results of the multiple linear regression model and the regression model coefficients are shown in Table 4.58 and Table 4.59, respectively. A significant regression equation was found ($F(1, 607) = 153.989$, $p < .001$), with an R^2 of .202 and an adjusted R^2 of .201.

Table 4.58 Regression Model Summary – CSAT DV

				Change Statistics				
R	R Square	Adj. R Square	Std. Error of Est.	R Square Change	F Change	df1	df2	Sig. F Change
.450[a]	.202	.201	.908	.202	153.989	1	607	<.001

a. Dependent Variable: CSAT; Independent Variable: SERVICE

Table 4.59 Regression Model Coefficients – CSAT DV

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations		
	B	Std Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
(Constant)	.756	.101		7.519	<.001	.559	.954			
SERVICE	.375	.030	.450	12.409	<.001	.316	.434	.450	.450	.450

[24] The value of N = 609 represents the subset of data with valid survey responses for both the SERVICE and CSAT variables, which is slightly lower than 612 and 1369 responses for the same variables independently.

Replacing the b-values in the model (4.7) with the values shown in Table 4.59, the model estimating the relationship between the predictors and the outcome of CSAT then becomes Equation (4.8):

$$CSAT_i = 0.756 + (0.375\ x\ SERVICE) \quad (4.8)$$

Participants' predicted level of satisfaction (CSAT) is .756 + .375 (SERVICE), where perceived service quality is scored on a scale of 1 = excellent to 5 = poor. Participant's level of satisfaction increases (improves) by .375 for each one point better rating on perceived service. Perceived service is a significant predictor of satisfaction.

4.6.5 Posttest Regression Analysis – MARKETING and CSAT

The third and final model of this section examines the relationship between the MARKETING IV and the CSAT DV, and as shown in (4.9).

$$CSAT_i = b_0 + b_1 MARKETING \quad (4.9)$$

The descriptive statistics resulting from the linear regression are shown in Table 4.60, and show the expected values for the CSAT DV overall (mean = 1.82, N = 1278)[25].

Table 4.60 Descriptive Statistics - Linear Regression Analysis of CSAT DV

	Mean	Std. Deviation	N
CSAT	1.82	.940	1278
MARKETING	3.98	1.040	1278

[25] The value of N = 1278 represents the subset of data with valid survey responses for both the MARKETING and CSAT variables, which is slightly lower than 1285 and 1369 responses for the same variables independently.

The summary results of the multiple linear regression model are shown in Table 4.61, with the regression model coefficients shown in Table 4.62. A significant regression equation was found ($F(1, 1276) = 227.164$, $p < .001$), with an R^2 of .151 and an adjusted R^2 of .150.

Table 4.61 Regression Model Summary – CSAT DV

				Change Statistics				
R	R Square	Adj. R Square	Std. Error of Est.	R Square Change	F Change	df1	df2	Sig. F Change
.389[a]	.151	.150	.867	.151	227.164	1	1276	<.001

a. Dependent Variable: CSAT; Independent Variable: MARKETING

Table 4.62 Regression Model Coefficients – CSAT DV

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations		
	B	Std Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
(Constant)	3.223	.096		33.597	<.001	3.035	3.411			
MARKETING	-.351	.023	-.389	-15.072	<.001	-.397	-.306	-.389	-.389	-.389

Replacing the *b*-values in the model (4.9) with the values shown in Table 4.62, the model estimating the relationship between the predictors and the outcome of CSAT then becomes Equation (4.10):

$$CSAT_i = 3.223 - (0.351\ x\ MARKETING) \quad (4.10)$$

Participants' predicted level of satisfaction (CSAT) is 3.223 - .3519 (MARKETING), where frustration with marketing messaging is scored on a scale of 1 = very often frustrated to 5 = not at all frustrated. Participant's level of satisfaction increases (improves) by .351 for each

one point better (higher) rating on perceived frustration with marketing messaging, indicating that this IV is a significant predictor of satisfaction.

4.6.6 Results of Hypothesis Tests – CSAT

The null hypotheses to be tested for the relationships between the independent variables of PERCVALUE, SERVICE and MARKETING and the dependent variable CSAT were stated in Section 3.4 as follows:

H_{04}: There is no relationship between perception of value and customer satisfaction

H_{05}: There is no relationship between perception of customer service and customer satisfaction

H_{06}: There is no relationship between frustration with marketing messaging and customer satisfaction

The linear regression analyses provide support for a significant relationship between each of the IVs and the DV of CSAT; therefore, the null hypotheses that there is no relationship between those IVs and the DV are each rejected (Table 4.63).

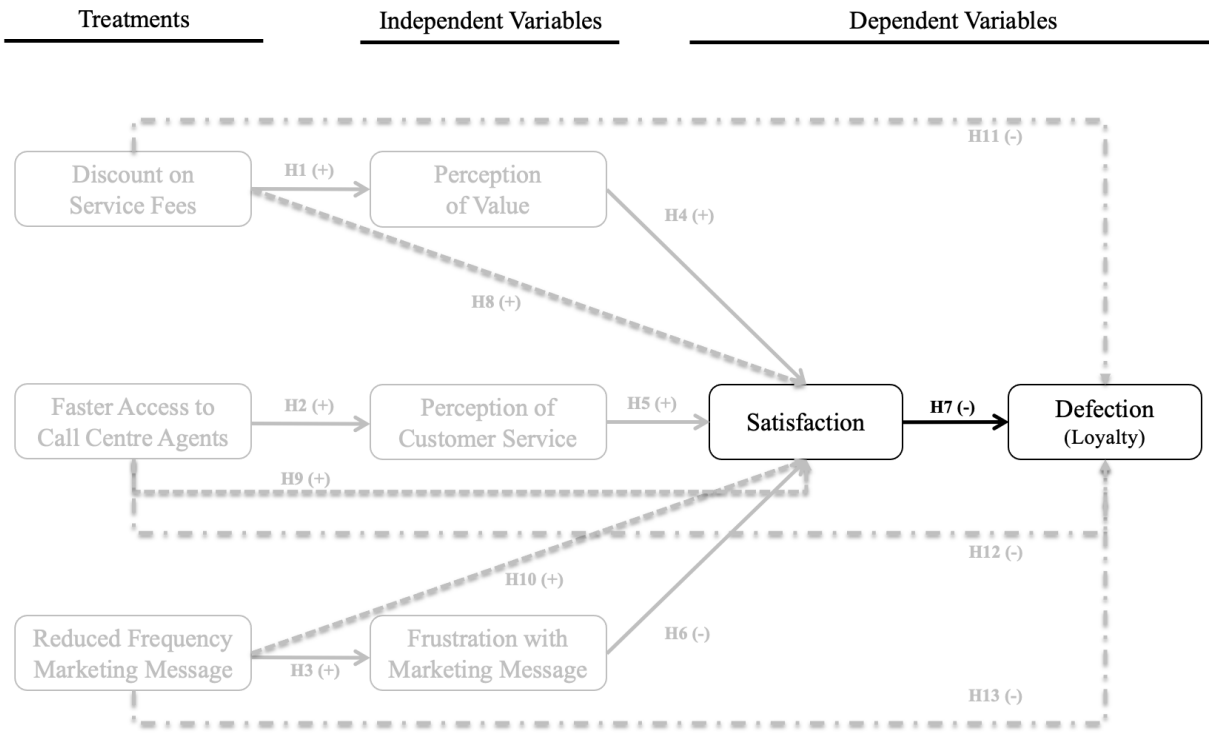
Table 4.63 Summary of CSAT DV Hypothesis Outcomes

Hypothesis Number	Null Hypothesis Statement	Outcome
H_{04}	***There is no relationship between perception of value and customer satisfaction***	Rejected
H_{05}	***There is no relationship between perception of customer service and customer satisfaction***	Rejected
H_{06}	***There is no relationship between frustration with marketing messaging and customer satisfaction***	Rejected

4.7 Analysis of DEFECT Dependent Variable

The next analysis to be completed is the examination of the relationship between the satisfaction variable of CSAT and the final dependent variable of DEFECT, as shown in Figure 20.

Figure 20 Hypothesis of Relationship Between CSAT and DEFECT DV



The CSAT variable was measured for the 1,369 study participants that responded to the posttest survey, whilst the outcome for deactivations was measured for all 35,000 study participants. Of the participants that responded to the posttest survey, there were no observed instances of deactivation during the experiment timeframe, precluding analysis of this relationship.

The null hypothesis that was to be tested for the dependent variable DEFECT was stated in Section 3.4 as follows:

H_{07}: Higher levels of customer satisfaction have no impact on customer defection (i.e. improved customer loyalty).

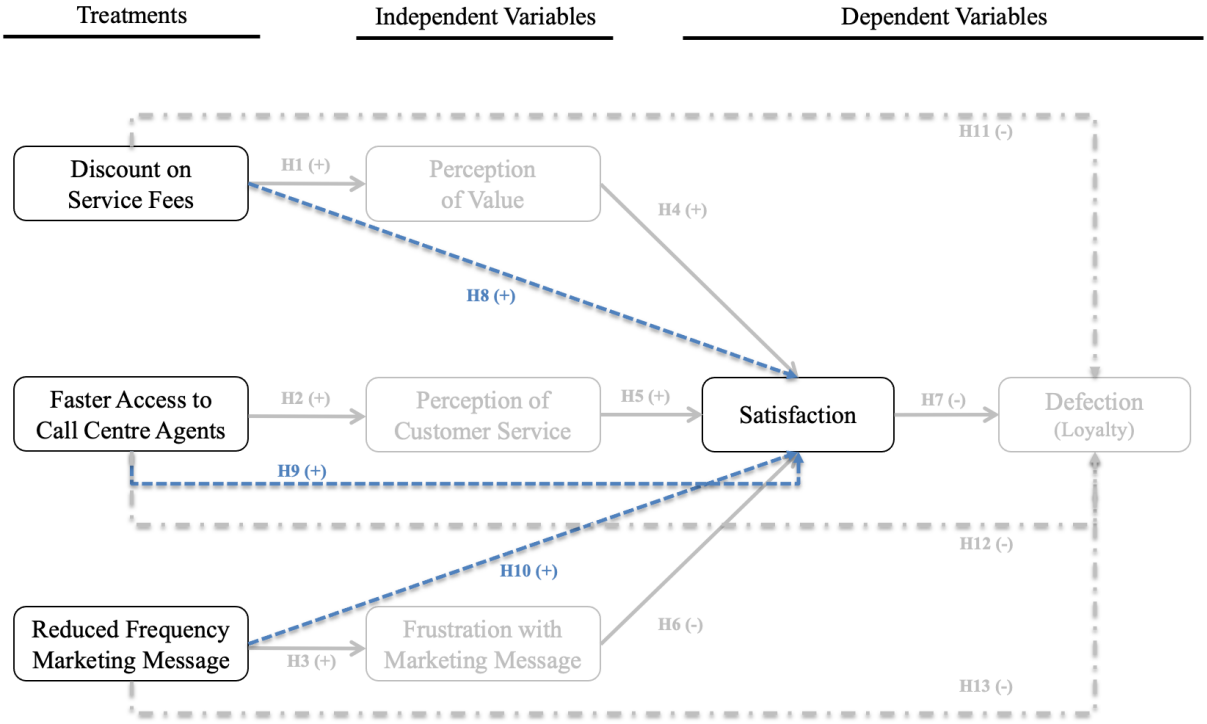
The null hypothesis cannot be rejected as there is insufficient data to determine if there is a significant relationship between the CSAT and DEFECT variables.

There is, however, sufficient data to evaluate the remaining hypotheses regarding the relationships between the experimental treatments and their effect, if any, on the DEFECT variable. This analysis is completed in Section 4.9.

4.8 Direct Relationship of Treatments and CSAT DV

The study model proposed the potential for the treatments implemented in the experiment to have a direct impact on the dependent variable of CSAT, instead of or in addition to, an influence through the perceptions measured by the independent variables of PERCVALUE, SERVICE and MARKETING. These direct relationships are highlighted in the model shown in Figure 21.

Figure 21 Hypotheses for Direct Relationships Between Treatments and CSAT DV



Linear regression analysis was conducted to model the relationship between the treatments (explanatory variables) and the CSAT dependent variable, and to show the degree (and strength) to which variation in the CSAT variable can be attributed to those explanatory variables.

The model that is examined in this regression analysis is shown in (4.9).

$$CSAT_i = b_0 + b_1 VALUE + b_2 CSERV + b_3 VALUE\_CSERV + b_4 CONTACT + b_5 VALUE\_CONTACT + b_6 CSERV\_CONTACT + b_7 VALUE\_CSERV\_CONTACT \quad (4.11)$$

This model defines the outcome of the prediction as $CSAT_i$ with regression coefficients for each of the predictors $VALUE$, $CSERV$, $VALUE\_CSERV$, $CONTACT$, $VALUE\_CONTACT$, $CSERV\_CONTACT$ and $VALUE\_CSERV\_CONTACT$. The predictors correspond to the seven treatment combinations utilized in the study, and were defined as dummy variables in Section 4.5.2.2.

4.8.1 Regression Analysis – Descriptive Results

The descriptive statistics shown in Table 4.64 show the expected values for CSAT (mean = 1.85, N = 1369), consistent with the analysis in Section 4.6.2.

Table 4.64 Descriptive Statistics - Linear Regression Analysis of CSAT

	Mean	Std. Deviation	N
CSAT	1.85	.957	1369
VALUE	0.12	.324	1369
CSERV	0.10	.299	1369
VALUE_CSERV	.014	.344	1369
CONTACT	0.12	.330	1369
VALUE_CONTACT	0.14	.347	1369
CSERV_CONTACT	0.13	.334	1369
VALUE_CSERV_CONTACT	0.14	.347	1369

4.8.1.1 Regression Model Summary and Coefficients

The summary results of the multiple linear regression model are shown in Table 4.65, and the regression model coefficients are shown in Table 4.66. A non-significant regression equation was found ($F(7, 1361) = .326$, $p > .05$), with an R^2 of .002 and an adjusted R^2 of -.003.

F-test results yielded an F-statistic of 0.311, which was not greater than the Critical F-Value of 2.01630. As a result, it is also concluded that the treatments do not jointly predict CSAT in a statistically significant manner.

None of the treatment variables were a significant direct predictor of CSAT, and therefore subsequent analysis of the results is not conducted. The regression analysis yields a result similar to the ANOVA analysis in Section 4.6.1, specifically that there does not appear to be a statistically significant relationship between the treatments employed in the study and the dependent or outcome variable of CSAT.

Table 4.65 Regression Model Summary

				Change Statistics				
R	R Square	Adj. R Square	Std. Error of Est.	R Square Change	F Change	df1	df2	Sig. F Change
.041[a]	.002	-.003	.959	.002	.326	7	1361	.942

a. Dependent Variable: CSAT, Independent Variables: VALUE, CSERV, VALUE_CSERV, CONTACT, VALUE_CONTACT, CSERV_CONTACT and VALUE_CSERV_CONTACT

Table 4.66 Regression Model Coefficients

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations		
	B	Std Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
(Constant)	1.884	.077		24.464	.000	1.733	2.035			
dum_VALUE	.018	.108	.006	.167	.867	-.193	.229	.020	.005	.005
dum_CSERV	-.046	.113	-.014	-.405	.685	-.267	.175	-.004	-.011	-.011
dum_VALUE_CSERV	-.049	.104	-.018	-.469	.639	-.253	.155	-.006	-.013	-.013
dum_CONTACT	-.025	.106	-.009	-.235	.814	-.234	.184	.004	-.006	-.006

dum_VALUE_CONTACT	-.088	.104	-.032	-.850	.396	-.291	.115	-.023	-.023	-.023
dum_CSERV_CONTACT	.013	.106	.005	.126	.900	-.194	.221	.019	.003	.003
dum_VALUE_CSERV_CONTACT	-.083	.104	-.030	-.799	.424	-.286	.120	-.020	-.022	-.022

4.8.2 Regression Analysis of CSAT DV – Additional Predictor Variables

The linear regression presented in Section 4.8.1 does not take into account the influence, if any, of any additional predictor (also known as control or extraneous) variables. In order to confirm the results observed in this initial linear regression, a subsequent linear regression was conducted with the addition of several additional predictor variables in order to determine if the outcome of the analysis for the treatment variables is any different.

A hierarchical regression analysis was conducted to determine the model that describes the relationship between several variables (including both potential predictor and treatment variables) and that of the outcome variable (in this case, CSAT) as observed in the experiment.

The hierarchical regression analysis was conducted by adding the variables to the analysis in SPSS in two stages or blocks. In block one, variables that are known to be potential predictors (based on the experience of the researcher and the subject organization) are loaded first. The second block consists of dummy variables that represent the various treatment combinations employed in the experiment. This approach allows for control of the role that the predictor variables play themselves in determining the outcome variable.

The potential predictor variables used in block one represent six demographic or profile attributes of the study participants. All but one variable (tenure_years) are dummy variables, and each variable is coded according to the definition shown in Table 4.67. The treatment variables used in block two are dummy variables, as previously defined in Section 4.6.2. Where dummy variables are used, one variable created from the categorical data for each dummy group is left out of the regression analysis (noted with *) as it would be redundant in the analysis.

The selection of the omitted dummy variable was based on experience within the subject organization as to what dummy variable would be consistent with references made internally when examining the same variables in actual business performance. The specifics of the dummy variables omitted, and the rationale, are detailed in Table 4.68.

The model that is examined in this hierarchical regression analysis is shown in Equation (4.10):

$$CSAT_i = b_0 + b_1 tenure\_years + b_2 CreditRiskMed + b_3 CreditRiskHigh + b_4 RevenueMed + b_5 RevenueHi + b_6 RevenueVeryHi + b_7 RevenueOther + b_8 RegionPrairies + b_9 RegionOntario + b_{10} RegionQuebec + b_{11} RegionAtlantic + b_{12} RegionOther + b_{13} Rural + b_{14} ContractNo + b_{15} VALUE + b_{16} CSERV + b_{17} VALUE\_CSERV + b_{18} CONTACT + b_{19} VALUE\_CONTACT + b_{20} CSERV\_CONTACT + b_{21} VALUE\_CSERV\_CONTACT \quad (4.12)$$

Table 4.67 Predictor Variables Employed in Block One

Variable Group	Variable	Represents
Tenure	tenure_days	Tenure of study participant in days at time of experiment start
Credit Risk	CreditRiskLow*	Credit risk profile at time of experiment start is Low
	CreditRiskMed	Credit risk profile at time of experiment start is Medium
	CreditRiskHigh	Credit risk profile at time of experiment start is High
Average Revenue	RevenueLow*	Average monthly revenue less than $35
	RevenueMed	Average monthly revenue between $35 to $64.99
	RevenueHi	Average monthly revenue between $65 to $149.99
	RevenueVeryHi	Average monthly revenue $150 or more
	RevenueOther	Other non-categorized revenue
Geographical Region	RegionWest*	Study participant lives in BC or AB
	RegionPrairies	Study participant lives in SK or MB
	RegionOntario	Study participant lives in Ontario
	RegionQuebec	Study participant lives in Quebec

	RegionAtlantic	Study participant lives in NL, NB, NS or PEI
	RegionOther	Study participant lives in Territories (NU, NT, YT) or outside Canada
Location Type	Urban*	Study participant lives in an Urban location
	Rural	Study participant lives in a Rural location
Contract Status	ContractNo*	Study participant was not under contract at start of experiment
	ContractYes	Study participant was under contract at start of experiment

Note: Dummy variables marked with an asterisk (*) were excluded from the model

Table 4.68 Omitted Dummy Variable Rationales

Variable Group	Omitted Variable	Rationale
Credit Risk	CreditRiskLow*	The majority (95%) of customers are assessed a low credit risk, and is considered the “normal” credit status for customers starting a business relationship with the company. The higher credit risk values indicate some issues with the customer’s credit profile (for example, a history of late payments with other companies).
Average Revenue	RevenueLow*	Objective of the subject organization is to grow the revenue earned from each customer. Using the low revenue variable as the reference variable in the regression supports analysis to understand the relationship of higher revenue values on deactivations.
Geographical Region	RegionWest*	The subject organization is based in BC and AB (defined as RegionWest value for this variable). As the “home territory”, comparisons of business performance are most commonly made as references to this geographical region.
Location Type	Urban*	The majority of customers (79%) are located in areas defined as Urban, and most competitive and business activities are focused in these areas.
Contract Status	ContractNo*	The subject organization assumes that customers without a contract (63% of study participants) are at greatest risk for deactivation or churn, and therefore form the baseline for the model.

4.8.2.1 Regression Model Summary, Coefficients and Collinearity Results

The model summary in Table 4.69 describes the overall model and provides an indication of whether the model is successful in predicting the outcome (in this case, customer satisfaction defined by L2R or $CSAT_i$).

The first stage of the hierarchical regression utilized the block of potential predictor variables. The results of this stage are shown as Model 1 in Table 4.69. The R value is .222, with R Square and Adjusted R Square values of .049 and .040, respectively. This indicates that approximately 4.0% of the variation in the outcome of the model is attributed to these predictors (using the Adjusted R Square value). The similar values between R Square and Adjusted R Square indicates that model generalizes well. The predictors yield a change in the *F*-ratio of 5.020, which is significant ($p = .000 < .05$).

Adding the treatment variables in the second stage of the regression analysis yields the results shown for Model 2. After adding the treatment variables, the R value increases slightly to .224. The increase from Model 1 to Model 2 for the R Square is .001 and Adjusted R Square decreases .004 reflecting the addition of additional variables in the second stage. Model 2 yields a change in the *F*-ratio of .200, which is not significant ($p = .986 > .05$).

Detailed results of this second linear regression analysis of the CSAT DV are provided in Appendix G. The model collinearity results are shown in this appendix indicate that no multicollinearity issues exist. The variance inflation factor (VIF) measures the correlation and strength of correlation between the predictor variables. Howell et al. (2003) suggest that the VIF should not be greater than 10; the VIF values in this regression analysis range from 1.011 to 4.860 and the average VIF for all attributes in Model 2 is 1.720, not much greater than 1, confirming no concerns regarding multicollinearity.

Table 4.69 Model Summary – CSAT Regression with Predictor Variables

Model	R	R Square	Adj. R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.222[a]	.049	.040	.938	.049	5.020	14	1354	.000
2	.224[b]	.050	.036	.940	.001	.200	7	1347	.986

a. Predictors: (Constant), ContractYes, RegionAtlantic, CrditRiskHigh, Rural, RegionOntario, RevenueVeryHi, RevenueOther, RegionPrairies, tenure_years, RegionQuebec, RevenueMed, CreditRiskMed, RevenueHi, RegionOther

b. Predictors: (Constant), ContractYes, RegionAtlantic, CrditRiskHighw, Rural, RegionOntario, RevenueVeryHi, RevenueOther, RegionPrairies, tenure_years, RegionQuebec, RevenueMed, CreditRiskMed, RevenueHi, RegionOther, VALUE, VALUE_CSERV, VALUE_CSERV_CONTACT, CSERV, VALUE_CONTACT, CSERV_CONTACT, CONTACT

4.8.2.2 Regression Model Parameters

The model coefficients shown in Appendix G (Model 2) allows for the replacement of the *b*-values in Equation (4.12), resulting in the definition for the model shown in (4.13).

$$\begin{aligned} CSAT_i = 2.236 &- (.007\ x\ tenure\_years) - (.048\ x\ CreditRiskMed) \\ &- (.482\ x\ CreditRiskHigh - (.025\ x\ RevenueMed) \\ &- (.117\ x\ RevenueHi) + (.138\ x\ RevenueVeryHi) \\ &- (.243\ x\ RevenueOther) - (.347\ x\ RegionPrairies) \\ &- (.270\ x\ RegionOntario) - (.511\ x\ RegionQuebec) \\ &- (.362\ x\ RegionAtlantic) - (.719\ x\ RegionOther) \\ &- (.031\ x\ Rural) - (.049\ x\ ContractYes) \\ &+ (.025\ x\ VALUE) - (.035\ x\ CSERV) \\ &- (.017\ x\ VALUE\_CSERV) - (.009\ x\ CONTACT) \\ &- (.006\ x\ VALUE\_CONTACT) \\ &+ (.021\ x\ CSERV\_CONTACT) \\ &- (.054\ x\ VALUE\_CSERV\_CONTACT) \end{aligned} \tag{4.13}$$

The degree to which each predictor variable affects the outcome variable of $CSAT_i$ (when the effects of all other variables are held constant) is:

- **tenure_years (b = -.007, p = .130 > .05):** This predictor does not play a statistically significantly role in predicting the outcome variable.
- **CreditRiskMed (b = -.048 , p = .767 > .05):** This predictor does not play a statistically significantly role in predicting the outcome variable.
- **CreditRiskHigh (b = -.482, p = .132 > .05):** This predictor does not play a statistically significantly role in predicting the outcome variable.
- **RevenueMed (b = -.025, p = .827 > .05):** This predictor does not play a statistically significantly role in predicting the outcome variable.
- **RevenueHi (b = -.117, p = .310 > .05):** This predictor does not play a statistically significantly role in predicting the outcome variable.
- **RevenueVeryHi (b = .138, p = .448 > .05):** This predictor does not play a statistically significantly role in predicting the outcome variable.
- **RevenueOther (b = -.243, p = .306 > .05):** This predictor does not play a statistically significantly role in predicting the outcome variable.
- **RegionPrairies (b = -.347, p = .001 < .05):** This predictor plays a statistically significant role in predicting the outcome variable. Customers in the Prairie region would be expected to exhibit a lower CSAT rate (that is, a higher likelihood to recommend) by .347 relative to customers in the West region (the reference variable).
- **RegionOntario (b = -.027, p = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers in the Ontario region would be expected to exhibit a lower CSAT rate by .027 relative to customers in the West region.
- **RegionQuebec (b = .511, p = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers in the Quebec region would be expected to exhibit a lower CSAT rate by .511 relative to customers in the West region.
- **RegionAtlantic (b = -.362, p = .003 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers in the Atlantic region would be expected to exhibit a lower CSAT rate by .362 relative to customers in the West region.
- **RegionOther (b = -.719, p = .189 > .05):** This predictor does not play a statistically significantly role in predicting the outcome variable.
- **Rural (b = -.031, p = .062 > .05):** This predictor does not play a statistically significantly role in predicting the outcome variable.
- **ContractYes (b = -.049, p = .385 > .05):** This predictor does not play a statistically significantly role in predicting the outcome variable.

- **VALUE (*b* = .025, *p* = .817 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **CSERV (*b* = -.035, *p* = .753 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **VALUE_CSERV (*b* = -.017, *p* = .865 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **CONTACT (*b* = -.009, *p* = .931 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **VALUE_CONTACT (*b* = -.060, *p* = .554 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **CSERV_CONTACT (*b* = .021, *p* = .844 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **VALUE_CSERV_CONTACT (*b* = -.054, *p* = .597 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.

The updated linear regression analysis (with the inclusion of the predictor variables) yields a result similar to the regression analysis in Section 4.8.1, specifically that there does not appear to be a statistically significant relationship between the treatments employed in the study and the dependent or outcome variable of CSAT.

4.8.2.3 Comparison of Linear Regression Models

To determine whether the inclusion of the predictor variables changes the outcome of the regression analysis as it relates to the influence of the treatment variable in the model, a comparison of the coefficient outcomes of the two analyses is show in Table 4.70.

The sign and significance of the coefficients for the dummy treatment variables agree between the two linear regression analyses, regardless of whether the predictor variables are included in the analysis or note. This indicates that predictor variables do not affect the outcome of the model of the treatments.

Table 4.70 Comparison of Linear Regression Outcomes

Model Summary	First Regression (no predictors)		Second Regression (with predictors)	
	B	Sig.	B	Sig.
(Constant)	1.884	.000	2.236	.000
tenure_years	-	-	-.007	.130

CreditRiskMed	-	-	-.048	.767
CreditRiskHigh	-	-	-.482	.132
RevenueMed	-	-	-.025	.827
RevenueHi	-	-	-.117	.310
RevenueVeryHi	-	-	.138	.448
RevenueOther	-	-	-.243	.306
RegionPrairies	-	-	-.347	.001
RegionOntario	-	-	-.270	.000
RegionQuebec	-	-	-.511	.000
RegionAtlantic	-	-	-.362	.003
RegionOther	-	-	-.719	.189
Rural	-	-	-.031	.620
ContractYes	-	-	-.049	.385
VALUE	.018	.867	.025	.817
CSERV	-.046	.685	-.035	.753
VALUE_CSERV	-.049	.639	-.017	.865
CONTACT	-.025	.814	-.009	.931
VALUE_CONTACT	-.088	.396	-.060	.554
CSERV_CONTACT	.013	.900	.021	.844
VALUE_CSERV_CONTACT	-.083	.424	-.054	.597

4.8.3 Hypothesis Tests – Direct Relationship of Treatments on CSAT DV

The null hypotheses to be tested for the relationships between the treatment (predictor) variables of VALUE, CSERV and CONTACT and the dependent variable CSAT were stated in Section 3.4 as follows:

H_{08}: ***Receiving a discount on monthly service fees has no impact on customer satisfaction***

H_{09}: Faster access to call centre agents has no impact on customer satisfaction

H_{010}: Reducing the frequency of marketing messages has no impact on customer satisfaction

The linear regression analyses do not indicate any significant relationship between the treatments and the outcome DV of CSAT; therefore, the null hypotheses that there is no impact between those IVs and the DV cannot be rejected (Table 4.71).

Table 4.71 Summary of CSAT DV Hypothesis Outcomes

Hypothesis Number	Null Hypothesis Statement	Outcome
H_{08}	***Receiving a discount on monthly service fees has no impact on customer satisfaction***	Cannot be rejected
H_{09}	***Faster access to call centre agents has no impact on customer satisfaction***	Cannot be rejected
H_{010}	***Reducing the frequency of marketing messages has no impact on customer satisfaction***	Cannot be rejected

4.8.4 Difference-in-Differences Analysis

A difference-in-differences (DID) analysis was conducted to determine if the difference between the mean CSAT score of the treatment versus non-treatment groups changed from the time period prior to the treatments being applied compared to the time period after the treatments were completed.

Table 4.72 shows the mean CSAT score for the various groups, pre- and post-treatment, while

Table 4.73 shows the mean CSAT score for the one non-treatment group (the CONTROL group) compared to the mean score for the seven treatment groups combined. The difference-in-differences between the non-treatment group and that of the combined groups that did receive treatment is 0.04.

Figure 22 illustrates the outcome of the DID analysis, showing the expected CSAT outcome if the treatments had no impact on the score for the treatment groups (noted as the "unobserved counterfactual outcome trend for treatment groups"). Had the treatments had no effect, the expected outcome for the treatment groups would have been a mean CSAT score of 1.89, maintaining the same 0.01 difference between the combined treatment groups and the CONTROL group prior to the experiment. Instead, the combined treatment groups show a mean CSAT score of 1.85, a difference of 0.04.

Within the treatment groups, six of the seven groups receiving treatment show a greater improvement in the mean CSAT score than the CONTROL group over the period of the experiment (ranging from 0.02 to 0.12 lower, or better). One treatment group (receiving only the CONTACT treatment) showed a higher, or worse, mean CSAT score by 0.02.

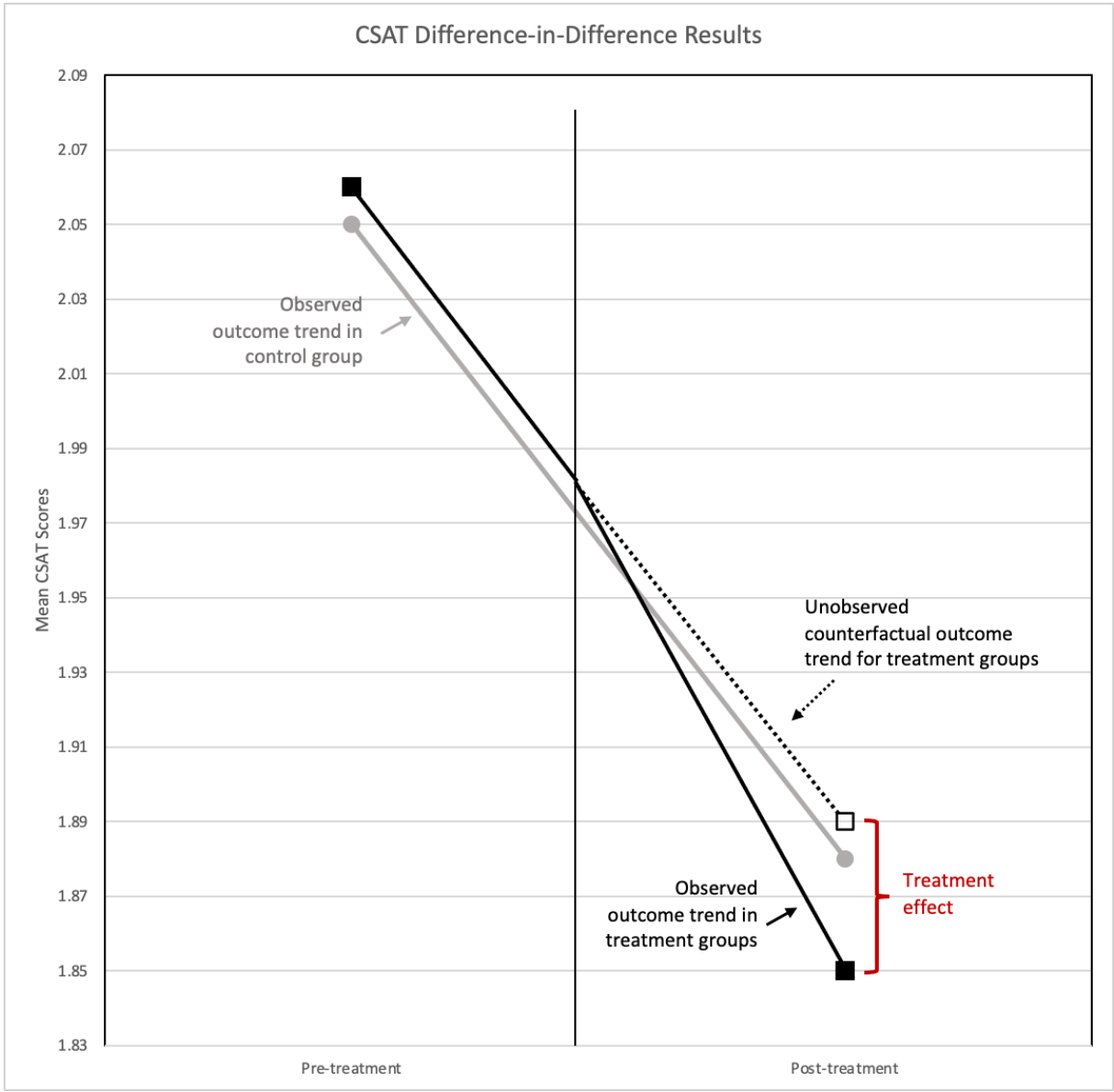
Table 4.72 Pre and Post CSAT Results by Group

Treatment Group	Pre-treatment Mean CSAT	Post-treatment Mean CSAT	Difference Post vs Pre	Difference-in-Difference vs. Control
CONTROL	2.05	1.88	-0.17	
VALUE	2.10	1.90	-0.20	-0.03
CSERV	2.08	1.84	-0.24	-0.07
VALUE + CSERV	2.08	1.84	-0.24	-0.07
CONTACT	2.01	1.86	-0.15	+0.02
VALUE + CONTACT	2.02	1.80	-0.22	-0.05
CSERV + CONTACT	2.09	1.90	-.019	-0.02
VALUE + CSERV + CONTACT	2.09	1.80	-0.29	-0.12

Table 4.73 Pre and Post CSAT Results: Treatment versus Non-Treatment

Treatment Group	Pre-treatment Mean CSAT	Post-treatment Mean CSAT	Difference
No Treatment (CONTROL)	2.05	1.88	-0.17
Received Treatment (ALL OTHER GROUPS)	2.06	1.85	-0.21
		Difference-in-Differences	**-0.04**

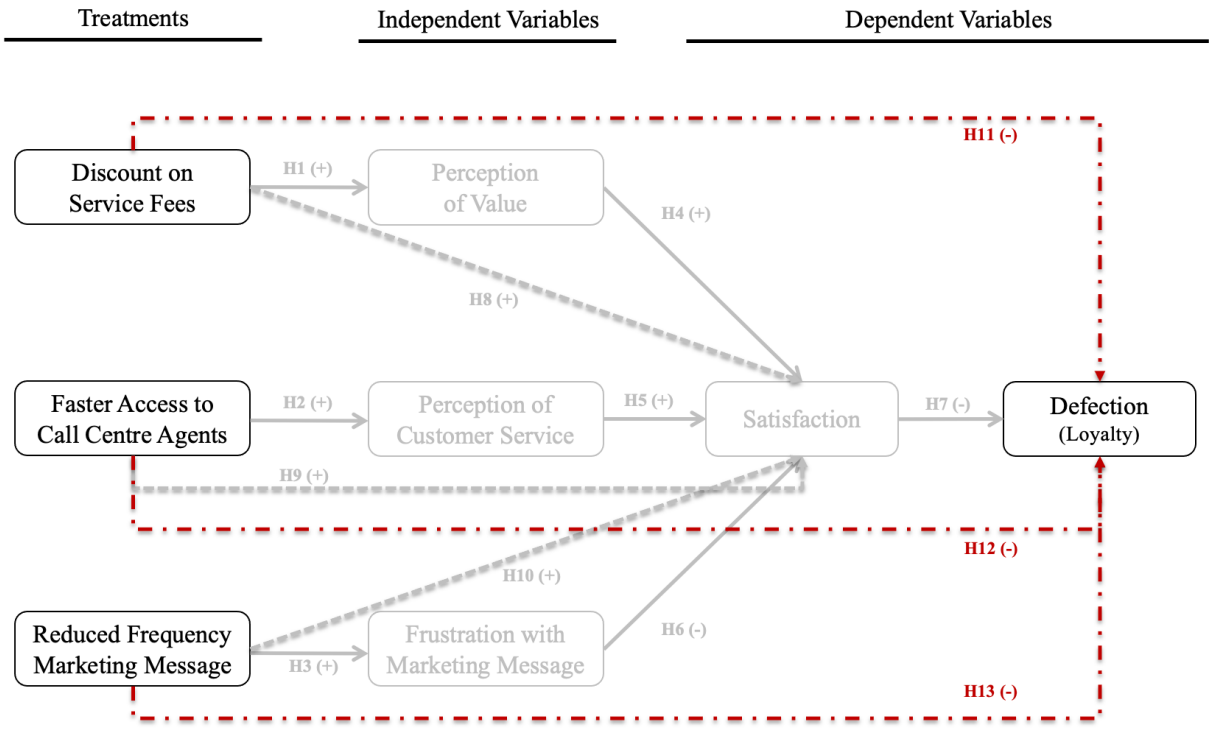
Figure 22 CSAT Difference-in-Difference Results



4.9 Direct Relationship of Treatments and DEFECT DV

The final section of this chapter evaluates the results of the experimental outcomes from the three treatments directly on the DEFECT dependent variable (DV), highlighted in the model shown in Figure 23. The analysis of the DEFECT DV begins with an examination of the DV levels across the treatment groups at the end of the experiment. Both ANOVA and linear regressions methods are used to test the impact, if any, of the treatments and changes in the IVs on the DEFECT DV. Logistic and probit regression analyses are also completed to confirm the use of the DEFECT DV as a continuous variable in the linear regression analysis is not an issue.

Figure 23 Hypotheses for Direct Relationships Between Treatments and DEFECT DV



4.9.1 One-Way ANOVA Analysis of DV: DEFECT

While the dependent variable DEFECT was measured on a monthly basis throughout the 12-month study period, for the purposes of the analysis presented here the cumulative results of deactivations across the study timeframe are analysed to determine the cumulative impact, if any, of the treatments on that DV.

The variable value for any single study participant is either 0 (if they have not deactivated over the course of the study) or 1 (if they have deactivated). The deactivation rate is the measure, for an individual treatment group or the study cohort as a whole, of the percentage of that group that deactivated during the study period. The total deactivation rate for all

participants in the study over the 12-month period was 11.58% (that is, 11.58% of the original 35,000 participants selected for the study had deactivated their mobile service by the end of the study). The deactivation rate varied from 10.54% to 12.76% across the treatment groups (Table 4.74).

One-way ANOVA results are presented in Table 4.75. The results of this analysis show a significance level of $p = .026$ ($p < .05$), indicating the mean responses for DEFECT vary between at least two of the treatment groups[26]. Additional post hoc analysis is required to determine which treatment groups have a significant difference in their mean DEFECT values. This analysis is presented in Section 4.9.1.1.

Table 4.74 Mean Cumulative Deactivation Rates by Treatment Group

					95% Confidence Interval for Mean			
Treatment Group	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
1	4400	.1186	.32340	.00488	.1091	.1282	0.00	1.00
2	4284	.1102	.31315	.00478	.1008	.1196	0.00	1.00
3	4397	.1217	.32695	.00493	.1120	.1313	0.00	1.00
4	4335	.1114	.31469	.00478	.1020	.1208	0.00	1.00
5	4412	.1276	.33369	.00502	.1178	.1375	0.00	1.00
6	4384	.1054	.30708	.00464	.0933	.1145	0.00	1.00
7	4404	.1197	.32461	.00489	.1101	.1293	0.00	1.00
8	4384	.1118	.31512	.00476	.1024	.1211	0.00	1.00
Total	35000	.1158	.32002	.00171	.1125	.1192	0.00	1.00

[26] Analysis of participant posttest responses to survey question 5 (measuring attitudinal loyalty) indicate no statistical difference between treatment groups with regards to i) intention to stay with subject organization, ii) willingness to switch for the right deal, or iii) declaration they are actively searching for a new provider.

Table 4.75 One-Way ANOVA Analysis – DEFECT

	Sum of Squares	dF	Mean Square	F	Sig.
Between Groups	1.633	7	.233	2.279	.026
Within Groups	3582.798	34992	.102		
Total	3584.431	34999			

4.9.1.1 Post-Hoc Analysis - DEFECT

A post hoc analysis using Tukey was completed to determine between which treatment groups the significant difference in means for the DEFECT DV can be observed. This analysis determined the subsets of treatment groups that exhibit a significant difference in reported means.

Table 4.76 shows the results of the Tukey post hoc analysis (additional details in Appendix H), showing a statistical difference in the DEFECT posttest results between treatment group 6 and treatment group 5. There was a statistically significant difference between groups as determined by one-way ANOVA ($F(7,34992) = 2.279$, $p = .026$).

The Tukey posthoc test revealed that the mean DEFECT result was statistically lower for group 6: VALUE+CONTACT (0.1054 ±0.307, $p = .025$) compared to group 5: CONTACT (0.1276 ±0.334).

Table 4.76 Tukey Post Hoc Analysis – MARKETING

		Subset for alpha = 0.05	
Treatment Group	N	1	2
6: VALUE + CONTACT	4384	.1054	
2: VALUE	4284	.1102	.1102
4: VALUE + CSERV	4335	.1114	.1114
8: VALUE + CSERV + CONTACT	4384	.1118	.1118
1: CONTROL	4400	.1186	.1186
7: CSERV + CONTACT	4404	.1197	.1197
3: CSERV	4397	.1217	.1217

5: CONTACT	4412		.1276
Sig.		.251	.176

4.9.2 Regression Analysis - DEFECT

A linear regression analysis across the treatment predictors was conducted to determine the model that describes the relationship between those predictor variables and that of the outcome variable (in this case, DEFECT) as observed in the experiment. The dummy variables defined in Section 4.6.2 were also used in this regression analysis.

Similar to the linear regression analysis conducted in Section 4.8 for the CSAT dependent variable, the linear regression analysis is completed twice: firstly, using only the dummy treatment variables, and secondly with the addition of the predictor variables.

Two additional regression analyses are completed to confirm that the use of the DEFECT DV as a continuous variable (though the underlying variable is dichotomous at the individual study participant level) is appropriate. A logistical regression and a PROBIT regression are used to validate the outcomes observed in the linear regression analyses.

The model that is examined in the first linear regression analysis is shown in (4.14).

$$DEFECT_i = b_0 + b_1 VALUE + b_2 CSERV + b_3 VALUE\_CSERV + b_4 CONTACT + b_5 VALUE\_CONTACT + b_6 CSERV\_CONTACT + b_7 VALUE\_CSERV\_CONTACT \quad (4.14)$$

This model defines the outcome of the prediction as $DEFECT_i$ with regression coefficients for each of the dummy predictor variables *VALUE*, *CSERV, VALUE_CSERV, CONTACT, VALUE_CONTACT, CSERV_CONTACT, and VALUE_CSERV_CONTACT*. The model would predict the deactivation rate ($DEFECT_i$) for a group as an outcome of whether they were included in one or more of the seven treatment combinations treatments (that is, whether b_1, b_2, b_3, b_4, b_5, b_6 or b_7 are 1 or not).

4.9.2.1 Descriptive Statistics for Regression Analysis

Descriptive statistics for the linear regression are shown in Table 4.77. The mean deactivation rate for all participants is 0.1158 with a standard deviation of .320. The mean and standard deviation results for the dummy variables used in the regression to represent the seven treatment combinations are as expected (as in Section 4.7.1).

Table 4.77 Descriptive Statistics from Linear Regression

	Mean	Std. Deviation	N
Cumulative Deacts	.1158	.320	35000
VALUE	.12	.328	35000
CSERV	.13	.331	35000
VALUE_CSERV	.12	.329	35000
CONTACT	.13	.332	35000
VALUE_CONTACT	.13	.331	35000
CSERV_CONTACT	.13	.332	35000
VALUE_CSERV_CONTACT	.13	.331	35000

4.9.2.2 Regression Model Summary, Coefficients and Collinearity Results

The model summary in Table 4.78 describes the overall model and provides an indication of whether the model is successful in predicting the outcome (in this case, deactivation rate or DEFECT). The model coefficients are shown in Table 4.79.

A significant regression equation was found ($F(7, 34992) = 2.279$, $p < .05$), with an R^2 of .000456 and an adjusted R^2 of .000256. The R value is low at 0.021, indicating that little to none of the variation in the outcome of the model is attributed to the predictors.

F-test results yielded an F-statistic of 1.62769, which was not greater than the Critical F-Value of 2.09886. As a result, it is concluded that the treatments do not jointly predict DEFECT in a statistically significant manner.

Table 4.78 Model Summary - DEFECT Regression

R	R Square	Adj. R Square	Std. Error of the Estimate	Change Statistics				
				R Square Change	F Change	df1	df2	Sig. F Change
.021[a]	.000456	.000256	.320	.000	2.279	7	34992	.026

a. Dependent Variable: DEFECT; Independent Variables: VALUE, CSERV, VALUE_CSERV, CONTACT, VALUE_CONTACT, CSERV_CONTACT and VALUE_CSERV_CONTACT

Table 4.79 Regression Model Coefficients

	Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations		
	B	Std Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
(Constant)	.119	.005		24.593	.000	.109	.128			
dum_VALUE	-.008	.007	-.009	-1.232	.218	-.022	.005	-.007	-.007	-.007
dum_CSERV	.003	.007	.003	.445	.656	-.010	.016	.007	.002	.002
dum_VALUE_CSERV	-.007	.007	-.007	-1.054	.292	-.021	.006	-.005	-.006	-.006
dum_CONTACT	.009	.007	.009	1.316	.188	-.004	0022	.014	.007	.007
dum_VALUE_CONTACT	-.013	.007	-.014	-1.941	.052	-.027	.000	-.012	-.010	-.010
dum_CSERV_CONTACT	.001	.007	.001	.151	.880	-.012	.014	.005	.001	.001
dum_VALUE_CSERV_CONTACT	-.007	.007	-.007	-1.006	.315	-.020	.017	-.005	-.005	-.005

4.9.2.3 Model Parameters

The model coefficients shown in Table 4.79 allow for the replacement of the *b*-values in (4.14), resulting in the definition for the model shown in Equation (4.15).

$$\begin{aligned} DEFECT_i = {} & 0.119 - (0.008\ x\ VALUE) + (0.003\ x\ CSERV) \\ & - (0.007\ x\ VALUE\_CSERV) + (0.009\ x\ CONTACT) \\ & - (0.013\ x\ VALUE\_CONTACT) \\ & + (0.001\ x\ CSERV\_CONTACT) \\ & - (0.007\ x\ VALUE\_CSERV\_CONTACT) \end{aligned} \quad (4.15)$$

The degree to which each predictor variable affects the outcome variable of $DEFECT_i$ (when the effects of all other predictor variables are held constant) is:

- **VALUE (*b* = -0.008, *p* = .218 > .05):** This variable does not play a statistically significant role in predicting the outcome variable
- **CSERV (*b* = 0.003, *p* = .656 > .05):** This variable does not play a statistically significant role in predicting the outcome variable
- **VALUE_CSERV (*b* = -0.007, *p* = .292 > .05):** This variable does not play a statistically significant role in predicting the outcome variable
- **CONTACT (*b* = 0.009, *p* = .188 > .05):** This variable does not play a statistically significant role in predicting the outcome variable
- **VALUE_CONTACT (*b* = -0.013, *p* = .052 > .05):** The *p* value is just slightly over the threshold defined in this study for statistical significance (that of .05). If taken as significant, study participants receiving both the VALUE and CONTACT treatments would be expected to exhibit a lower DEFECT rate by .013 relative to the control group (the excluded reference group).
- **CSERV_CONTACT (*b* = 0.001, *p* = .880 > .05):** This variable does not play a statistically significant role in predicting the outcome variable
- **VALUE_CSERV_CONTACT (*b* = -0.007, *p* = .315 > .05):** This variable does not play a statistically significant role in predicting the outcome variable

The regression analysis indicates a similar outcome to the ANOVA analysis in section 4.7.1, which indicated a statistically significant difference for the treatment group VALUE_CONTACT compared to the other treatment groups.

If the VALUE_CONTACT treatment is taken as a significant predictor in the model (a reasonable assumption, as the same treatment is found to be significant with p < .05 in the PROBIT and logisitic regression analyses to be presented in Sections 4.9.4 and 4.9.5, respectively), participants' deactivation rate (DEFECT) is .119 -.013 (VALUE_CONTACT). A participant's deactivation rate is predicted to be lower (from 11.9% to 10.6%) if they received both the VALUE and CONTACT treatments.

4.9.3 Linear Regression Analysis of DEFECT DV – Additional Predictor Variables

Following the same procedures described in Section 4.6.3, the linear regression analysis completed in the preceding section was conducted a second time with the addition of predictor variables.

The model that is examined in this hierarchical regression analysis is shown in (4.16):

$$DEFECT_i = b_0 + b_1 tenure\_years + b_2 CreditRiskMed + b_3 CreditRiskHigh + b_4 RevenueMed + b_5 RevenueHi + b_6 RevenueVeryHi + b_7 RevenueOther + b_8 RegionPrairies + b_9 RegionOntario + b_{10} RegionQuebec + b_{11} RegionAtlantic + b_{12} RegionOther + b_{13} Rural + b_{14} ContractNo + b_{15} VALUE + b_{16} CSERV + b_{17} VALUE\_CSERV + b_{18} CONTACT + b_{19} VALUE\_CONTACT + b_{20} CSERV\_CONTACT + b_{21} VALUE\_CSERV\_CONTACT \quad (4.16)$$

4.9.3.1 Descriptive Statistics for Linear Regression Analysis with Descriptor Variables

Descriptive statistics for the linear regression are shown in Table 4.80. The mean deactivation rate for all participants is 0.1158 with a standard deviation of .320, the same results observed in the ANOVA analysis in Section 4.7.1 and the prior linear regression analysis in Section 4.7.2.1. The mean and standard deviation results for the predictor and treatment variables are as expected.

Table 4.80 Descriptive Statistics from DEFECT Linear Regression with Predictor Variables

	Mean	Std. Deviation	N
Cumulative Deacts	.1158	.320	35,000
tenure_years	7.685	6.02759	35,000
CreditRiskLow	.95	.223	35,000
CreditRiskMed	.04	.186	35,000
CreditRiskHi	.02	.128	35,000
RevenueLow	.09	.283	35,000
RevenueMed	.27	.442	35,000
RevenueHi	.58	.493	35,000

	Mean	Std. Deviation	N
RevenueVeryHi	.04	.206	35,000
RevenueOther	.02	.142	35,000
RegionWest	.51	.500	35,000
RegionPrairies	.07	.259	35,000
RegionOntario	.21	.405	35,000
RegionQuebec	.16	.366	35,000
RegionAtlantic	.05	.212	35,000
RegionOther	.00	.050	35,000
Urban	.79	.410	35,000
Rural	.21	.410	35.000
ContactNo	.63	.484	35,000
ContractYes	.37	.484	35,000
CONTROL	.13	.332	35,000
VALUE	.12	.328	35,000
CSERV	.13	.331	35,000
VALUE_CSERV	.12	.329	35,000
CONTACT	.13	.332	35,000
VALUE_CONTACT	.13	.331	35,000
CSERV_CONTACT	.13	.332	35,000
VALUE_CSERV_CONTACT	.13	.331	35,000

4.9.3.2 Regression Model Summary, Coefficients and Collinearity Results

The model summary in Table 4.81 describes the overall model and provides an indication of whether the model is successful in predicting the outcome (in this case, deactivation rate or $DEFECT_i$).

The first stage of the hierarchical regression utilized the block of potential predictor variables. The results of this stage are shown as Model 1 in Table 4.81. The *R* value is .153, with R^2 and adjusted R^2 values of .023 each, indicating that approximately 2.3% of the variation in the outcome of the model is attributed to these predictors. The similar values between R^2 and adjusted R^2 indicates that model generalizes well. The predictors yield a change in the *F*-ratio of 60.029, which is significant (p = .000 < .05).

Adding the treatment variables in the second stage of the regression analysis yields the results shown for Model 2. After adding the treatment variables, the R value increases slightly to .155. The increase from Model 1 to Model 2 for the R^2 and adjusted R^2 values is also small at .001 and nil, respectively. Model 2 yields a change in the *F*-ratio of 2.176, which is significant (p = .034 < .05).

Detailed results of this second linear regression analysis of the DEFECT DV are provided in Appendix I. The model collinearity results are shown in this appendix indicate that no multicollinearity issues exist. The VIF values in this regression analysis range from 1.005 to 3.437 and the average VIF for all attributes in Model 2 is 1.55, not much greater than 1, confirming no concerns regarding multicollinearity.

Table 4.81 Model Summary – DEFECT Linear Regression with Predictor Variables

Model	R	R Square	Adj. R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.153[a]	.023	.023	.316	.023	60.029	14	34985	.000
2	.155[b]	.024	.023	.316	.000	2.176	7	34978	.034

a. Predictors: (Constant), ContractYes, RegionAtlantic, CrditRiskHigh, Rural, RegionOntario, RevenueVeryHi, RevenueOther, RegionPrairies, tenure_years, RegionQuebec, RevenueMed, CreditRiskMed, RevenueHi, RegionOther

b. Predictors: (Constant), ContractYes, RegionAtlantic, CrditRiskHighw, Rural, RegionOntario, RevenueVeryHi, RevenueOther, RegionPrairies, tenure_years, RegionQuebec, RevenueMed, CreditRiskMed, RevenueHi, RegionOther, VALUE, VALUE_CSERV, VALUE_CSERV_CONTACT, CSERV, VALUE_CONTACT, CSERV_CONTACT, CONTACT

4.9.3.3 Model Parameters

The model coefficients shown in Appendix I (Model 2) allows for the replacement of the *b*-values in (4.16), resulting in the definition for the model shown in Equation (4.17).

$$DEFECT_i = .195 - (.004\ x\ tenure\_years) + (.047\ x\ CreditRiskMed) + (.173\ x\ CreditRiskHigh - (.073\ x\ RevenueMed) - (.083\ x\ RevenueHi) - (.075\ x\ RevenueVeryHi) - (.035\ x\ RevenueOther) + (.029\ x\ RegionPrairies) + (.038\ x\ RegionOntario) + (.029\ x\ RegionQuebec) + (.042\ x\ RegionAtlantic) - (.031\ x\ RegionOther) - (.017\ x\ Rural) + (.024\ x\ ContractYes) - (.008\ x\ VALUE) + (.003\ x\ CSERV) - (.008\ x\ VALUE\_CSERV) + (.007\ x\ CONTACT) - (.013\ x\ VALUE\_CONTACT) + (.001\ x\ CSERV\_CONTACT) - (.008\ x\ VALUE\_CSERV\_CONTACT) \quad (4.17)$$

The degree to which each predictor variable affects the outcome variable of $DEFECT_i$ (when the effects of all other variables are held constant) is:

- **tenure_years (*b* = -.004, *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. For every one year of tenure, the expected DEFECT rate will be .004 lower.
- **CreditRiskMed (*b* = .047 , *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. For customers with an assigned credit risk of medium, the expected DEFECT rate will be .047 higher than customers with a credit risk of low (the reference variable).
- **CreditRiskHigh (*b* = .173, *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. For customers with an assigned credit risk of high, the expected DEFECT rate will be .173 higher than customers with a credit risk of low. Higher credit risk customers would be expected to exhibit a higher deactivation (churn rate).
- **RevenueMed (*b* = -.073, *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers with monthly revenue between $35 and $65 would be expected to exhibit a lower DEFECT rate by .083 relative to customers with a low (<$35) revenue level (the reference variable).
- **RevenueHi (*b* = -.083, *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers with monthly revenue between

$65 and $150 would be expected to exhibit a lower DEFECT rate by .083 relative to customers with a low revenue value.

- **RevenueVeryHi (*b* = -.075, *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers with monthly revenue >$150 would be expected to exhibit a lower DEFECT rate by .075 relative to customers with a low revenue value.
- **RevenueOther (b = -.035, *p* = .012 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers with this revenue value would be expected to exhibit a lower DEFECT rate by .035 relative to customers with a low revenue value.
- **RegionPrairies (*b* = .029, *p* = .000 < .05):** This predictor plays a statistically significant role in predicting the outcome variable. Customers in the Prairie region would be expected to exhibit a higher DEFECT rate by .029 relative to customers in the West region (the reference variable).
- **RegionOntario (*b* = .038, *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers in the Ontario region would be expected to exhibit a higher DEFECT rate by .038 relative to customers in the West region.
- **RegionQuebec (*b* = .029, *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers in the Quebec region would be expected to exhibit a higher DEFECT rate by .029 relative to customers in the West region.
- **RegionAtlantic (*b* = .042, *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers in the Atlantic region would be expected to exhibit a higher DEFECT rate by .042 relative to customers in the West region.
- **RegionOther (*b* = -.031, *p* = .364 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable.
- **Rural (*b* = -.017, *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customer living in areas designated as Rural would be expected to exhibit a lower DEFECT rate by .017 relative to customers in Urban areas (the reference variable).
- **ContractYes (*b* = .024, *p* = .000 < .05):** This predictor plays a statistically significantly role in predicting the outcome variable. Customers with a contract at the start of the study period would be expected to exhibit a higher DEFECT rate by .024 relative to customers without a contract.

- **VALUE (*b* = -0.008, *p* = .225 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **CSERV (*b* = 0.003, *p* = .644 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **VALUE_CSERV (*b* = -0.008, *p* = .235 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **CONTACT (*b* = 0.007, *p* = .271 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **VALUE_CONTACT (*b* = -0.013, *p* = .058 > .05):** The *p* value is just slightly over the threshold defined in this study for statistical significance (that of .05). If taken as significant, customers receiving both the VALUE and CONTACT treatments would be expected to exhibit a lower DEFECT rate by 0.013 relative to customers in the control group (the reference variable).
- **CSERV_CONTACT (*b* = 0.001, *p* = .829 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.
- **VALUE_CSERV_CONTACT (*b* = -0.008, *p* = .241 > .05):** This variable does not play a statistically significant role in predicting the outcome variable.

The updated linear regression analysis of the DEFECT DV (with the inclusion of the predictor variables) yields a result similar to the regression analysis in Section 4.7.2.3, specifically that – with one exception – there does not appear to be a statistically significant relationship between the treatments employed in the study and the dependent or outcome variable of DEACT. The one exception is the near significance of the VALUE_CONTACT treatment group.

4.9.3.4 Comparison of Linear Regression Models

Following the same approach taken in Section 4.6.3.3, a comparison of the coefficient outcomes of the two linear regression analyses is shown in Table 4.82.

The sign and significance of the coefficients for the dummy treatment variables agree between the two linear regression analyses, regardless of whether the predictor variables are included in the analysis or note. Consistent with the same analysis for the linear regression analyses of the CSAT DV, this indicates that the predictor variables do not affect the outcome of the model of the treatments.

Table 4.82 Comparison of Linear Regression Outcomes

Model Summary	First Regression (no predictors)		Second Regression (with predictors)	
	B	Sig.	B	Sig.
(Constant)	.119	.000	.195	.000
tenure_years	-	-	-.004	.000
CreditRiskMed	-	-	.047	.000
CreditRiskHigh	-	-	.173	.000
RevenueMed	-	-	-.073	.000
RevenueHi	-	-	-.083	.000
RevenueVeryHi	-	-	-.075	.000
RevenueOther	-	-	-.035	.012
RegionPrairies	-	-	.029	.000
RegionOntario	-	-	.038	.000
RegionQuebec	-	-	.029	.000
RegionAtlantic	-	-	.042	.000
RegionOther	-	-	-.031	.364
Rural	-	-	-.017	.000
ContractYes	-	-	.024	.000
VALUE	-.008	.218	-.008	.225
CSERV	.003	.656	.003	.644
VALUE_CSERV	-.007	.292	-.008	.235
CONTACT	.009	.188	.007	.271
VALUE_CONTACT	-.013	.052	-.013	.058
CSERV_CONTACT	.001	.880	.001	.829
VALUE_CSERV_CONTACT	-.007	.315	-.008	.241

4.9.4 PROBIT Regression Analysis of DEFECT DV

A PROBIT regression is the first of two additional analyses conducted to validate that the use of the DEFECT dependent variable as a continuous variable was appropriate, although the underlying variable at the study participant level is dichotomous (that is, 0 if the participant did not deactivate during the experiment and 1 if they did).

As with the first linear regression analysis conducted in Section 4.9.2, the dependent variable for the PROBIT regression was set as the DEFECT outcome; seven dummy variables were included representing the treatments applied in the experiment (the dummy variable for the control group was excluded).

The overall model (shown in Table 4.83) shows that the model with the treatment variables is a significantly (p = .026, < .05) better fit than a model without those variables.

The model summary is shown in Table 4.84, and a side-by-side comparison of the PROBIT analysis outcomes with the linear regression analysis from Section 4.9.2 is shown in

Table 4.85. The signs of the PROBIT coefficients and their respective significance levels are consistent with those observed in the two linear regression analyses. There is also a consistent outcome as to which treatment variables are significant in the models (in both cases, the VALUE_CONTACT variable is significant, or very close). These results indicate that there are no issues with treating the DEFECT DV as a continuous variable and therefore using linear regression as the model for the analysis.

Table 4.83 PROBIT Analysis Output - Omnibus Test

Likelihood Ratio Chi-Square	df	Sig.
15.916	7	.026

Dependent Variable: cumulative_churn_Apr19_Mar20
Model: (Intercept), dum_VALUE, dum_CSERV, dum_VALUE_CSERV, dum_CONTACT, dum_VALUE_CONTACT, dum_CSERV_CONTACT, dum_VALUE_CSERV_CONTACT

Table 4.84 PROBIT Analysis Output - Model Summary

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for (Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-1.182	.0246	-1.230	-1.134	2314.306	1	.000	.307	.292	.322
VALUE	-.044	.0353	-.113	.026	1.532	1	.216	.957	.893	1.026
CSERV	.015	.0346	-.053	.083	.192	1	.661	1.015	.949	1.087
VALUE_CSERV	-.037	.0352	-.106	.032	1.117	1	.291	.963	.899	1.032
CONTACT	.044	.0344	-.023	.111	1.642	1	.200	1.045	.977	1.118
VALUE_CONTACT	-.070	.0354	-.139	.000	3.877	1	.049	.933	.870	1.000
CSERV_CONTACT	.005	.0347	-.063	.073	.022	1	.882	1.005	.939	1.076
VALUE_CSERV_CONTACT	-.035	.0351	-.104	.033	1.016	1	.314	.965	.901	1.034

Dependent Variable: cumulative_churn_Apr19_Mar20

Model: (Intercept), dum_VALUE, dum_CSERV, dum_VALUE_CSERV, dum_CONTACT, dum_VALUE_CONTACT, dum_CSERV_CONTACT, dum_VALUE_CSERV_CONTACT

Table 4.85 Comparison of Estimation Results (Signs and Statistical Significance) from Linear and Probit Regressions (DEFECT DV)

	Linear Regression			Probit Regression		
	B	Std Error	Sig.	B	Std Error	Sig.
(Constant) / (Intercept)	.119	.005	.000	-1.182	.0246	.000
dum_VALUE	-.008	.005	.218	-.044	.0353	.216
dum_CSERV	.003	.007	.656	.015	.0346	.661
dum_VALUE_CSERV	-.007	.007	.292	-.037	.0352	.291
dum_CONTACT	.009	.007	.188	.044	.0344	.200

dum_VALUE_CONTACT	-.013	.007	.052	-.070	.0354	.049
dum_CSERV_CONTACT	.001	.007	.880	.005	.0347	.882
dum_VALUE_CSERV_CONTACT	-.007	.007	.315	-.035	.0351	.314

4.9.5 Logistical Regression Analysis of DEFECT DV

A final examination of the experiment results is conducted using a logistical regression analysis. As with the PROBIT analysis, the logistical regression utilizes the DEFECT DV as a dichotomous variable, allowing for examination of the results between the two types of regression analysis to confirm the findings are the same.

The logistical regression analysis provides additional support for the use of the DEFECT DV as a continuous variable in a linear regression model since the signs and significance levels of each of the treatment variables is the same between the two methods of regression analysis.

The logistical regression method provides additional value in that it allows for the determination of the probability of an outcome for the dependent variable given the state of a given input variable. This is different than the regression model examined in Section 4.9.2, but the inference of the probability outcome should align with that of the linear regression models.

To examine whether this is true, the probability of a deactivation outcome can be predicted using the results from Table 4.86 together with the formula shown in Equation (4.18). This formula predicts the probability of *Y* occurring given known values of *X*s.

$$P(Y) = \frac{1}{1 + e^{-(b_0 + b_1 X_1 + \cdots b_n X_n)}} \tag{4.18}$$

The treatment variable VALUE_CONTACT is statistically significant (p = .049, < .05) and is used to complete the probability calculation shown in Equation (4.19), starting with the determination of *Y* (a deactivation event) occurring for a study participant receiving no treatments (equal to the omitted dummy treatment variable, or the control group). The probability of a deactivation for a customer receiving no treatments is .1187, or 11.87%.

$$P(DEFECT) = \frac{1}{1 + e^{-(b_0)}} \tag{4.19}$$

$$P(DEFECT) = \frac{1}{1 + e^{-(-2.005)}}$$

$$P(DEFECT) = \frac{1}{1 + 7.426093897}$$

$$P(DEFECT) = .1187$$

Repeating the probability calculation with the inclusion of the treatment combination VALUE_CONTACT yields the results shown in Equation (4.20). The probability of a deactivation for a customer receiving the VALUE_CONTACT treatment is .1055, or 10.55%. This is 1.32 percentage points lower than that of the control group. This is the same result obtained from the linear regression in Section 4.9.2, demonstrating that the DEFECT DV can be treated as a continuous or dichotomous variable for the purposes of conducting a regression analysis.

$$P(DEFECT) = \frac{1}{1 + e^{-(b_0 + b_5 X_5)}} \tag{4.20}$$

$$P(DEFECT) = \frac{1}{1 + e^{-(-2.005 + (-0.133\ x\ 1))}}$$

$$P(DEFECT) = \frac{1}{1 + 8.482455741}$$

$$P(DEFECT) = .1055$$

Table 4.86 Logistical Regression Analysis Output - Model Summary

							95% Wald Confidence Interval for (Exp(B))	
Parameter	**B**	**Std. Error**	**Wald**	**df**	**Sig.**	**Exp(B)**	**Lower**	**Upper**
(Constant)	-2.005	.047	1850.252	1	.000	.135		
VALUE	-.084	.067	1.535	1	.216	.920	.806	1.050
CSERV	.029	.066	.192	1	.661	1.029	.905	1.170

VALUE_CSERV	-.071	.067	1.117	1	.291	.932	.817	1.062
CONTACT	.083	.065	1.641	1	.200	1.087	.957	1.234
VALUE_CONTACT	-.133	.068	3.874	1	.049	.875	.766	.999
CSERV_CONTACT	.010	.066	.022	1	.882	1.010	.888	1.149
VALUE_CSERV_CONTACT	-.067	.067	1.015	1	.314	.935	.820	1.066

Dependent Variable: cumulative_churn_Apr19_Mar20

Model: (Intercept), dum_VALUE, dum_CSERV, dum_VALUE_CSERV, dum_CONTACT, dum_VALUE_CONTACT, dum_CSERV_CONTACT, dum_VALUE_CSERV_CONTACT

Table 4.87 Comparison of Estimation Results (Signs and Statistical Significance) from Linear and Logistic Regressions (DEFECT DV)

	Linear Regression			Logisitic Regression		
	B	**Std Error**	**Sig.**	**B**	**Std Error**	**Sig.**
(Constant) / (Intercept)	.119	.005	.000	-2.005	.047	.000
dum_VALUE	-.008	.005	.218	-.084	.067	.216
dum_CSERV	.003	.007	.656	.029	.066	.661
dum_VALUE_CSERV	-.007	.007	.292	-.071	.067	.291
dum_CONTACT	.009	.007	.188	.083	.065	.200
dum_VALUE_CONTACT	-.013	.007	.052	-.133	.068	.049
dum_CSERV_CONTACT	.001	.007	.880	.010	.066	.882
dum_VALUE_CSERV_CONTACT	-.007	.007	.315	-.067	.067	.314

4.9.6 Results of Hypothesis Test - DEFECT

The null hypotheses to be tested for the relationships between the treatment (predictor) variables and the DEFECT dependent variable was stated in section 3.4 as follows:

H_{011}: Receiving a discount on monthly service fees has no impact on customer defection

H_{012}: Faster access to call centre agents has no impact on customer defection

H_{013}: Reducing the frequency of marketing messages has no impact on customer defection

The null hypotheses ***H_{011}*** and ***H_{013}*** are rejected, based on the results of the various regression analyses conducted in Section 4.9, where the combination of two treatments (VALUE and CONTACT) was observed to have a significant impact on the prediction of the DEFECT DV. The third null hypothesis ***H_{012}*** cannot be rejected.

4.10 Summary

This chapter provided an analysis of the data from the experiments. These results will be discussed in more detail in Chapter 5 (Discussion).

The following is a high-level summary of the results presented in this chapter:

- The three treatments were successfully deployed and created measurable differences in the experience of the study participants (Section 4.3)
- The effect of the CSERV treatment may be muted by overall improvements in call centre service levels during the experiment for all customers of the subject organization
- A statistically significant impact of the treatments was observed on the MARKETING independent variable, with no statistically significant impact observed on the PERCVALUE and SERVICE IVs (Section 4.5)
- Statistically significant relationships were observed between each of the three independent variables and that of the CSAT dependent variable (Section 4.6)
- A lack of data from study participants that responded to both the posttest survey and exhibited a deactivation behaviour precludes analysis of the relationship between CSAT and the DV of DEFECT (Section 4.7)
- No statistically significant impact was observed on the DV of CSAT as a direct result of the treatments in the experiment (Section 4.8)
- A statistically significant impact was observed on the DEFECT DV as a result of one or more of the treatments employed in the experiment (Section 4.9)

In each section, the results of the analysis of the experiment were used to test the thirteen hypotheses proposed in Section 3.4 of this study. The results of these hypothesis tests are shown in Table 4.88.

Table 4.88 Summary of Null Hypothesis Test Outcomes

Hypothesis Reference	Null Hypothesis Statement	Outcome
H_{01}	*There is no relationship between receiving a discount on monthly service fees and a customer's perception of value.*	Cannot be rejected
H_{02}	*There is no relationship between faster access to call centre agents and a customer's perception of customer service quality.*	Cannot be rejected
H_{03}	*There is no relationship between reduced frequency of marketing messages and a customer's frustration with receiving such marketing messaging.* With the rejection of the null hypothesis H_{03}, the alternative hypothesis H_{a3} is true: *There is a positive relationship between reduced frequency of marketing messages and a customer's frustration with receiving such marketing messaging*	**Rejected**
H_{04}	*There is no relationship between perception of value and customer satisfaction.* With the rejection of the null hypothesis H_{04}, the alternative hypothesis H_{a4} is true: *There is a relationship between perception of value and customer satisfaction*	**Rejected**
H_{05}	*There is no relationship between perception of customer service and customer satisfaction.* With the rejection of the null hypothesis H_{05}, the alternative hypothesis H_{a5} is true: *There is a relationship between perception of customer service and customer satisfaction*	**Rejected**
H_{06}	*There is no relationship between frustration with marketing messaging and customer satisfaction.* With the rejection of the null hypothesis H_{06}, the alternative hypothesis H_{a6} is true:	**Rejected**

	There is a relationship between frustration with marketing messaging and customer satisfaction.	
H_{07}	*Higher levels of customer satisfaction have no impact on customer defection (i.e. improved customer loyalty).*	Cannot be rejected
H_{08}	*Receiving a discount on monthly service fees has no impact on customer satisfaction.*	Cannot be rejected
H_{09}	*Faster access to call centre agents has no impact on customer satisfaction.*	Cannot be rejected
H_{010}	*Reducing the frequency of marketing messages has no impact on customer satisfaction.*	Cannot be rejected
H_{011}	*Receiving a discount on monthly service fees has no impact on customer defection.* With the rejection of the null hypothesis H_{011}, the alternative hypothesis H_{a11} is true: *Receiving a discount on monthly service fees causes an improvement in customer defection (i.e. lower rates of defection)*	**Rejected**
H_{012}	*Faster access to call centre agents has no impact on customer defection*	Cannot be rejected
H_{013}	*Reducing the frequency of marketing messages has no impact on customer defection* With the rejection of the null hypothesis H_{013}, the alternative hypothesis H_{a13} is true: *Reducing the frequency of marketing messages causes an improvement in customer defection.*	**Rejected**

5 Discussion

5.1 Introduction

This chapter revisits each of the hypotheses outlined in Section 3.4 and provides an interpretation of the results observed in the prior chapter for each hypothesis, together with the potential implications of the results on management practice.

5.2 Independent Variables: PERCVALUE, SERVICE, MARKETING

The three IVs of PERCVALUE, SERVICE and MARKETING were measures established to support the first three hypotheses established in this study. Each is discussed in turn in the following subsections.

5.2.1 IV: PERCVALUE

The first null hypothesis of the study was presented as:

H_{01}: There is no relationship between receiving a discount on monthly service fees and a customer's perception of value.

The results of the experiment failed to reject this hypothesis, as there were no statistically significant differences observed in the posttest PERCVALUE scores between participants that received the VALUE treatment (a 10% reduction in their monthly service fees for 12 months) compared to those that did not receive the Treatment; see Sections 4.5.2.1 and 4.5.2.2 for detailed results.

Reviewing the survey question used to measure this IV ("Considering the overall quality and the overall price you pay, how would you rate (subject organization) for overall value in terms of what you pay for?"), it might be concluded that study participants did not perceive the 10% savings as sufficient to change their opinion of the overall value of what they pay for.

Previous researchers, as reviewed in Section 2.2.5.1, have not attempted to affect the level of perceived value through the use of a treatment in an experiment; rather, prior studies measured the levels of perceived value at a point in time without understanding how that level of perception came to be, or how or if it could be changed. The VALUE treatment aimed to affect the participant-determined perception of value (McDougall et al. 2000), altering the ratio between perceived costs and perceived benefits by reducing the former. The study's results indicate that the use of a discount to lower costs was not successful in altering the participants' perception of value. Practitioners looking to exploit the relationship between perceived value and customer satisfaction, as established in more than 50 studies (El-Adly

2019), are left without a proven method to causally influence this antecedent. It may be that of the two components of the perceived value ratio (cost versus benefits), benefits play a greater role (for example: keeping cost constant, does improving the benefits or utility of the service cause perceived value to increase?). Further research is warranted to understand whether or not these two drivers of perceived value can leveraged to alter that same perception.

The outcome of this experimental treatment on perceived value is important to the subject organization and its competitors in the telecommunications market. Discounts around the same 10% level used in this study are a standard offering to incentivize potential customers to join the organization, or to try and retain customers that the organization already has. For example, all three major mobile service providers in Canada offer a $15 discount when adding a second mobile phone service to their first service. While the value this $15 discount represents as a percentage of total cost will vary based on the mobile service plan chosen by a customer, the entry level plans promoted as of this writing[27] were $80/month for each mobile phone. The $15 discount would represent a 9.4% discount on the $160 monthly cost two mobile phone plans – nearly identical to the 10% discount studied in this experiment.

The subject organization also offers a discount to its existing mobility customers when they add other services provided by the company (for example, TV or Internet). This discount, in the amount of $10/month, applies to the new service added. Monthly rates for the new services range start at $120/month, putting the $10 discount at a value of 8.3% - again close to the 10% discount studied here.

While the primary goal of these discounts is to promote and reward loyalty, the organization also feels that enhanced perception of value is an important metric that is worth improving in order to contribute to customer satisfaction (the implication of this study's results on the dependent variable CSAT are explored in Section 5.3).

5.2.2 IV: SERVICE

The null hypothesis that utilized the independent variable SERVICE as its measure was presented as:

H_{02}: There is no relationship between faster access to call centre agents and a customer's perception of customer service quality.

[27] Lowest monthly fee for unlimited data plans as of March 7th, 2021 as published online by TELUS, Bell and Rogers

This null hypothesis cannot be rejected, as the ANOVA analysis in Section 4.5.2.3 did not reveal any statistically significant difference between the treatment groups (that received priority access to call centre agents) and those that did not receive the treatment. The linear regression analysis conducted in Section 4.5.2.4 similarly yielded a non-significant regression formula.

The analysis of the execution of the CSERV treatment in Section 4.3.2 highlighted that concurrent with this experiment, the subject organization undertook efforts to improve the timeliness of access to its call centre agents for all customers. As a result, the impact of the treatment – on a relative basis compared to a non-treatment experience – was muted, especially over the last seven months of the experiment where the mean wait times for an agent were almost identical between treatment and non-treatment groups.

5.2.3 IV: MARKETING

The independent variable MARKETING was defined as to test the null hypothesis of:

H_{03}: There is no relationship between reduced frequency of marketing messages and a customer's frustration with receiving such marketing.

This null hypothesis was rejected on the evidence presented in Sections 4.5.2.5 and 4.5.2.6. The ANOVA analysis, and subsequent post-hoc analysis, identified a statistically significant difference between treatment and non-treatment groups in response to the survey question:

> "How often are you frustrated with communication quality, e.g. relevance or frequency of emails, other marketing?"

The ANOVA analysis revealed that three of the four study groups receiving the CONTACT treatment (which substantially supressed the frequency of marketing contacts to the participant, such as email and SMS solicitations) were statistically less frustrated than one other group that did not receive the treatment. These three treatment groups had mean MARKETING IV scores of 4.11 to 4.12, compared to 3.76 for the non-treatment group that was statistically different (a higher score indicates lower levels of frustration). These results indicate that the suppression of marketing contacts in the treatment was causal to the reduction in reported participant frustration.

The linear regression analysis yielded a significant regression equation, which predicts an outcome of less frustration with marketing messaging when customers receive three of the seven treatment combinations where the CONTACT treatment is combined with the other

two treatments, specifically: VALUE_CONTACT, CSERV_CONTACT, and VALUE_CSERV_CONTACT.

These results indicate that it is possible to cause customers to be less frustrated by reducing the frequency of marketing messages sent to those same customers when that reduced frequency is accompanied by a discount on their services and/or an improvement in wait times to access a call centre agent. The results support the concept of advertising pressure proposed by Micheaux (2011) which hypothesized that irritation forms as a function of receiving too much marketing communications from a company.

5.2.4 Summary of Findings for the Independent Variables

The analysis of the impact of the experimental treatments on the three independent variables revealed that only the MARKETING IV was statistically impacted by the related CONTACT treatment (and then, only when in combination with either one or both of the VALUE and CSERV treatments); the remaining two IV's showed no statistically significant impact from the treatments.

In practice, organizations are not seeking to positively impact the customer perceptions represented by these three IVs in and of themselves. Instead, the organization expects improved financial outcomes as a result of the positive influence of the IVs on both overall customer satisfaction and loyalty. The outcome of the experiment on these two dependent variables will be examined in the following sections.

5.3 Dependent Variable: CSAT

The first of two dependent variables established in the research model is CSAT, a metric measured by the survey question:

> "If a colleague, friend or family member were looking for a new wireless provider, what is the likelihood that you would **recommend** (subject organization) to them?"

The study model hypothesizes the relationships between the independent variables PERCVALUE, SERVICE and MARKETING as follows (null hypothesis statements shown):

H_{04}: There is no relationship between perception of value and customer satisfaction.

H_{05}: There is no relationship between perception of customer service and customer satisfaction.

H_{06}: There is no relationship between frustration with marketing messaging and customer satisfaction.

Each of these three null hypothesis statements were rejected based on the results of the linear regression analysis conducted between each of the IV variables PERCVALUE, SERVICE, and MARKETING and the DV of CSAT (Section 4.6). Each of the findings of the relationships between these IVs and CSAT is discussed in turn, and in combination with the findings of the previous section that evaluate the impact, if any, of the treatments on those same IVs.

5.3.1 PERCVALUE Relationship to CSAT

The regression analysis in Section 4.6.3 yielded a statistically significant outcome, predicting the outcome of CSAT based on the participants' level of perceived value. Higher levels of perceived value indicate greater levels of satisfaction by the participants. For a participant with the lowest level of perceived value (a score of 5 where 1 = excellent perceived value and 5 = poor perception of value), the regression predicts a CSAT value of 2.869 (where 1 = definitely will recommend and 5 = definitely will not recommend) – a strong "maybe" as to whether the participant would recommend the subject organization to others.

Conversely, a participant with the highest level of perceived value (a score of 1) will have a predicted CSAT value of 0.869. This result would place the participant firmly in the "definitely" would recommend the organization.

These results support the findings of previous research establishing a strong relationship between perceived value and customer satisfaction (such as the findings of Bolton et al. 1991, Cronin et al. 2000, Grönroos et al. 2013).

Combining the analysis conducted between the relationship of the experiment treatment on the PERCVALUE IV, and that of the relationship between the same IV and the CSAT DV, the study does not yield an indication of how to cause customers to be more satisfied, at least through the attempted manipulation of perceived value employed in this study. The results yield an outcome reinforcing that there is a positive relationship between perception of value and customer satisfaction, but not a way to manipulate that outcome to the benefit of the customer or the organization. If the perception of value cannot be influenced by the practitioner in order to effect a change in customer satisfaction, the assertion by prior researchers that perceived value is an "…important means of generating customer satisfaction…" (El-Adly 2019, p. 2) may come into question and efforts to influence this antecedent may not be a viable means for organizations to improve the satisfaction of their customers.

As called for in Section 5.2.1, future research should consider further examination of the causal antecedents of perceived value. Extending the treatment utilized in this study, additional research may consider utilizing difference discount levels (for example, 15% or higher), providing the discounts as absolute dollar benefits rather than as percentages, and executing recurring messaging to the study participants to reinforce that they are receiving the discount. Periodic reminders of the cumulative value received through these discounts may also be explored to determine the role, if any, in influencing the perception of value.

5.3.2 SERVICE Relationship to CSAT

Similar to the relationship between PERCVALUE and CSAT, there was also a statistically significant relationship established between the SERVICE IV and the CSAT DV. Study participants who reported a more positive experience relative to the time they had to wait to speak to a customer service representative are more likely to be more satisfied. While the predicted outcome of the CSAT DV was not as varied for those with the best or worst levels of perceived service (CSAT outcomes of 1.179 and 2.630, respectively), participants on the two respective ends of the service perception scale would still fall into "definitely" would recommend versus "maybe".

The significant relationship between the antecedent SERVICE and the dependent variable of CSAT observed in this study supports the assertions made in prior research of a direct

linkage between waiting time and customer satisfaction (Durrande-Moreau 1999, Feinberg et al. 2000, Whiting et al. 2009).

Accepting the relationship between better perceived customer service and improved customer satisfaction, practitioners will look for evidence that such perceptions can be directly influenced through operational investments in their call centres (that is, hiring more agents to answer calls resulting in lower wait times for customers calling). The implementation of the CSERV treatment in the study did not provide a significant impact to the outcome of the SERVICE IV. As discussed in Section 4.3.2, improvements made to overall customer service by the subject organization concurrent with the study timeframe negated any potential influence that the CSERV treatment might have had on the study participants. As a result, it cannot be established whether or not such a treatment would have had an impact on perceived service, and through that same antecedent, on customer satisfaction itself.

Future research should seek opportunities where the baseline levels of customer service (in terms of time to access an agent) are materially worse than those observed in the study, providing an opportunity for a treatment similar to the CSERV treatment deployed here to have a potentially greater effect. Such research may fill the gap that existed prior to (Feinberg et al. 2000, Feinberg et al. 2002), and still remains following, this study as to whether wait times can be causally linked to customer satisfaction.

5.3.3 MARKETING Relationship to CSAT

Contrary to the findings of the experiment outcomes between the treatments and IVs for value and customer service and their relationships to customer satisfaction, the study results show there is an ability to cause a change in perceived frustration with marketing messaging received by a company, and that the level of perceived frustration is inversely related to the level of customer satisfaction.

The experiment demonstrated that the combination of reducing the frequency of marketing messaging together with either the provision of a value discount and/or improving the level of service when calling the call centre creates the outcome of reduced levels of frustration as exhibited in the MARKETING variable. Reducing the frequency of marketing messaging alone did not result in a statistically significant change in the IV.

The CSAT value predicted by the regression analysis in Section 4.6.4 improves from 2.917 for a participant that is highly frustrated with marketing messaging (equivalent to a 3 or a “maybe” likelihood to recommend result on the posttest survey) to 1.481 when they report

"not at all frustrated" with the frequency of such messaging (an improvement to somewhere between "probably" and "definitely" would recommend the organization to others).

Taken together, the significant relationships between the CONTACT treatment and the MARKETING IV, and between the MARKETING IV and the CSAT DV, a causal chain can be established as to a method to improve customer satisfaction. This finding adds to the existing literature which had previously established that irritation with communication frequency leads to withdrawal of permission for the organization to send future communications, as well as the development of negative attitudes by the customer towards the organization itself, its brand, and its products (Micheaux 2011).

This ability to reduce customer frustration may appear to be a clear-cut solution for marketers seeking to improve customer satisfaction. However, reducing the frequency of marketing messaging eliminates an important channel for the marketer to promote and sell additional products and services. The net benefit of reducing customer frustration must be examined by also measuring the impact, if any, on the future financial benefits that would have accrued from the ability to market via this method. This is an area for future research to examine in more detail.

Additionally, the CONTACT treatment employed in the study was "all or nothing" – participants either continued to receive the normal level of communications from the subject organization or received none. Future research should examine whether there are different levels of communication frequency that can still contribute to reduced customer frustration (and therefore improved levels of satisfaction) whilst providing an opportunity to continue to promote additional products and services to the customer.

5.4 Dependent Variable: CSAT – Direct Impact of Treatments

In addition to the relationships hypothesized between the independent variables (PERCVALUE, SERVICE, and MARKETING) and that of CSAT (as reviewed in the prior section), the model also puts forward three hypotheses related to the direct impact that the treatments employed in the experiment might have on CSAT.

Sections 5.2 and 5.3 together established that there was a causal chain between the CONTACT treatment → MARKETING IV → CSAT DV, but there was no similar chain of significance between the remaining two treatments through their respective IVs and ending with CSAT. The hypotheses evaluated in this section establish whether or not a direct relationship between the treatments and the DV exists independent of whether the three IVs themselves have such a relationship.

The null hypothesis statements linking the treatments VALUE, CSERV and CONTACT to the CSAT DV are:

H_{08}: Receiving a discount on monthly service fees has no impact on customer satisfaction.

H_{09}: Faster access to call centre agents has no impact on customer satisfaction.

H_{010}: Reducing the frequency of marketing messages has no impact on customer satisfaction.

The three hypotheses cannot be rejected, as no statistically significant differences were observed in the CSAT DV as a result of the treatments employed in the experiment (see analysis in Section 4.8).

The results of the hypothesis statements H_{08} and H_{09} are consistent with the findings in Sections 5.2 and 5.3; specifically that there was no significant relationship between the VALUE and CSERV treatments and the CSAT variable, regardless of whether the intervening IVs of PERCVALUE and SERVICE are included in the analysis.

The most important implication of this outcome for practitioners is that they should be wary of the use of discounts to improve customer satisfaction.

The results of the test for hypothesis H_{010} were not consistent with the findings in Sections 5.2 and 5.3. The significant relationship established between the CONTACT treatment and MARKETING, and between MARKETING and CSAT is not reflected in any direct relationship between the CONTACT treatment and CSAT. This raises questions not answered in the study, including:

- Does a change in the MARKETING IV as a result of the treatment take more time (that is, greater than the 12 months of the study) to significantly affect the CSAT DV?
- Are there different measures within the study's MARKETING IV construct that need to be isolated to determine their individual relationships with the treatments and the CSAT DV?

The study results do not provide the practitioner with a statistically significant mechanism to effect a change in customer satisfaction. But does this mean there is no value in the use of any of the treatments to improve customer satisfaction?

In practice, many organizations (including the subject organization of this study) will not seek the high degree of validity established by tests of statistical significance. Section 4.2.2

established that tests of practical significant might also be used to evaluate the outcome of the experiment. Results from these follow-up tests are provided in Appendix J for the practitioner to examine but are not reviewed in detail here.

5.4.1 Summary of Findings for the CSAT Dependent Variable

The results of the experiment as it pertains to the CSAT DV indicated that there were no statistically significant differences as a result of the various treatments employed in the study. However, when reviewing the results in a manner consistent with how the subject organization tracks and rewards its employees based on improvements on the same metric, the results are of practical significance. As a whole (and for the most part individually), the study groups receiving one or more treatments show greater improvements in their T2B CSAT scores (those participants indicating that they “Definitely” or “Probably” would recommend the subject organization to their colleagues, friends or family) relative to the control group which received no treatments.

The T2B results reinforce the findings from the Difference-in-Differences analysis conducted in Section 4.6.3, which showed an improvement in mean CSAT scores for six of seven groups receiving treatment over the course of the experiment.

The results of the experiment indicates that despite the numerous studies examined in Section 2.2.5 that have previously established a relationship between the customer satisfaction and the antecedents utilized as treatments in this experiment (and in particular, that of perceived value), there is no statistically significant evidence that the DV of customer satisfaction can be affected in a controlled manner.

The study shows that the treatments did not have a statistically significant impact on the IVs intended to measure the impact of those treatments, so it may be that the treatments themselves were insufficient in some way. For example, the VALUE treatment of a 10% discount may not have been sufficient, all other things being equal, to create a statistically significant difference of opinion on perceived value by the participants. Similarly, the study did not measure awareness of the treatment and the associated discount, so it could be that a lack of awareness of the treatment resulted in the lack of impact on the PERCVALUE IV or the CSAT DV. Further research should explore the role that differing levels of discounts (for example, 15% or 20%) in affecting the levels of a customer’s perceived value, as well as examining whether there is a relationship between awareness of such a discount and perceived value, and whether attempts to increase awareness have an impact on customer satisfaction.

The study is inconclusive in supporting prior research that has sought to establish a relationship between the satisfaction that a customer experiences calling into a call centre and that of overall customer satisfaction (Section 2.2.5.2). As noted in the analysis of the wait time experience of customers assigned to the CSERV treatment (Section 4.3.2.2) the anticipated difference between treatment and non-treatment groups dissipated after the first four months of the experiment, with no difference between treatment and non-treatment groups for the remaining eight months of the study. As a result, the lack of a statistically significant influence on the CSAT DV does not provide any evidence for or against the conclusions of the prior research. Further research should examine opportunities to repeat the experiment when greater differences in the wait times to access a contact centre agent can be sustained.

While the analysis of the CONTACT treatment revealed a statistically significant impact of that treatment on the MARKETING IV, there was no significant impact of that treatment on the CSAT DV. The influence of the treatment on the MARKETING IV indicates that the treatment itself was sufficiently different from the non-treatment experience to influence the IV (which was not the case of the VALUE and CSERV treatments). The results of the study support prior research (examined in Section 2.2.5.3) which found that while frequency of communications may be an irritant to customers, this sentiment was not found to impact subsequent behaviours of the customers.

Overall, these results provide practical evidence that the L2R customer satisfaction metric at the subject organization might be improved through further implementation of one or more of the treatments on a broader scale across the organization. However, the net financial benefit of doing so is yet to be determined; this will be examined in the following section.

5.5 Dependent Variable: DEFECT – Direct Impact of Treatments

The second and final dependent variable, DEFECT, was measured by tracking the percentage of participants within each treatment group that left the subject organization during the period of the study. Two sets of hypotheses were set out in the model that relate to this DV. The first (H_{07}) hypothesized the relationship between the CSAT variable and the DEFECT DV. As noted in Section 4.7, the experimental study did not yield data to allow for this hypothesis to be tested. As a result, this section focuses on the second set of hypotheses related to the DEFECT DV, those establishing a direct link between the treatments employed in the study and the DEFECT DV. The final three hypotheses were stated as:

H_{011}: Receiving a discount on monthly service fees as no impact on customer defection

H_{012}: Faster access to call centre agents has no impact on customer defection

H_{013}: Reducing the frequency of marketing messaging has no impact on customer defection

The analysis completed in Section 4.9 provided evidence sufficient to reject the null hypotheses H_{011} and H_{013}, as the regression analysis yielded statistically significant results predicting the outcome of the DEFECT DV based on the joint application of the VALUE and CONTACT treatments. Hypothesis H_{012} cannot be rejected based on the results of the study.

The finding that the combined treatment of VALUE + CONTACT together causally affect the DEFECT DV yields an important outcome of the study: it demonstrates an ability to causally influence the loyalty of study participants by providing them with a 10% monthly discount on their service and concurrently reducing the volume of marketing messages sent to those participants.

The regression model (4.15) predicts that the participants' cumulative deactivation rate over twelve months decreases from 11.9% to 10.6% if they receive both treatments, compared to those participants that receive no treatments.

The study partially answers the call made by Cronin (2016) for further research to link customer behaviour with customer perceptions of value, through the evidence that customer loyalty (as evidenced by the actual behaviour of staying or leaving a company) can be causally influenced by treatments that affect one-half of the value perception ratio. Left unanswered is whether the perception of value itself can be causally influenced (the study results could not reject the null hypothesis that this relationship did not exist). The evidence that customer behaviour can be causally affected by the manipulation of the price or cost of a service provides a foundation for further research into the relationship between price/costs and perception of value, to determine if it is necessary for a customer to cognitively recognize the change in value to change their loyalty behaviour, or whether the change in behaviour is affected more subconsciously.

The practitioner is left to determine if the benefit of greater loyalty and lower deactivations are sufficient to offset the costs associated with providing the two treatments (both the absolute cost of the 10% discount as well as the opportunity cost from reduced marketing

communications to sell additional products and services). The cost/benefit outcome of applying this treatment combination in practice is examined in the follow section 5.5.1.

5.5.1 Practical Net Financial Implications of VALUE Treatment

To determine the net financial implication of providing the VALUE treatment of a 10% discount to a group of customers, it is necessary to compare the cost of providing the discount to the benefit of fewer customers deactivating. The opportunity cost of withholding marketing communications (the CONTACT treatment) was not examined in this study; as a result, the following analysis focuses on the cost/benefit of the VALUE treatment alone.

The subject organization utilizes a simple approach to determine the cost/benefit outcome for a program such as the VALUE treatment employed in this study. This approach uses two inputs: the monthly revenue expected from each group (those receiving the discount treatment, and those who did not) as well as the monthly deactivation rate for each group (derived from the annual rate observed over the course of the study). This approach does not take into account the actual profit or margin that is left over from the revenue received from each customer, so it should be considered a generous approach to interpreting results. The result of this cost/benefit analysis is shown in Table 5.1.

Participants in the control group have an expected lifetime of 101 months, paying an average of $65.35 per month. This yields an expected lifetime revenue for each participant of $6,600. Participants receiving the VALUE+CONTACT treatment are expected to remain customers for 13 months longer (as a result of their lower deactivation or churn rate), but also pay $5.78 less per month as a result of the treatment discount. These participants yield an expected lifetime revenue of $6,790 which is $190 greater than that of the control group.

On a per-customer basis, the increase in lifetime value of $190 over the course of an eight or nine year relationship might not seem substantial. Extrapolation of this result across the target population of the study (3.5 million customers) yields an increase in total lifetime value of more than $600 million, demonstrating the substantial financial contribution that can be realized if the treatment utilized in this experiment were deployed fully.

As already noted, this estimation of increased lifetime value does not take into account the opportunity cost of removing participants from future marketing communications (the CONTACT treatment). The study does not measure whether the participants receiving this treatment purchased fewer other products and services during the course of the experiment. Any statistically significant decrease in such purchases relative to the control group might negate some or all of the financial benefits that result from a lower deactivation rate. Future

research should incorporate additional tracking of this customer behaviour and its potential impact on the financial outcome.

The conclusions summarized here assume that the treatment discount would need to be provided in perpetuity to sustain the lower deactivation rates used in the cost/benefit calculation. The study does not examine whether removal of the discount at the end of the experiment resulted in the deactivation rates of those receiving the VALUE treatment returning to their pre-experiment levels. Future research should examine the impact of providing both ongoing and fixed-duration discounts over an even longer study period.

It is also possible that a larger discount could engender a greater improvement in deactivation rates, sufficient to offset the even greater financial cost. Conversely, it may be that a lower discount could generate the same loyalty benefit, and thereby improve the net financial outcome of the treatment. Both of these price elasticity scenarios remain hypotheses to be tested in future research.

Table 5.1 Cost/Benefit Analysis of Discount Treatment

Group	Annual Deactivation Rate[a]	Monthly Deactivation Rate (monthly churn)	Expected Lifetime (months)[b]	Average Monthly Recurring Revenue[c]	Expected Lifetime Revenue Value
CONTROL	11.9%	0.99%	101	$65.35	$6,600
VALUE + CONTACT	10.6%	0.88%	114	$59.57	$6,790
Difference (Treatment vs Control)	-1.3 pts	-11 basis points	+13 months	-$5.78	+$190

[a] Taken from results of the linear regression analysis in Section 4.7, specifically Equation 4.6.

[b] Calculated as: $\frac{1}{monthly\ churn}$

[c] As reported in Section 4.3.1

5.6 Discussion Summary

The results presented and discussed in Chapters 4 and 5 provide the details of the outcomes observed following a 12-month randomized controlled trial to examine the impact, if any, of specific treatments on the customer satisfaction and loyalty of mobile service subscribers in Canada.

The study utilized a variety of statistical methods to examine the results of the experiment, including methods to identify both the statistical and practical significance of the results. The outcomes of the overall study can be summarized as:

- Statistically significant causal relationships were established between the CONTACT treatment and the MARKETING independent variable, and between the VALUE treatment and the DEFECT dependent variable.
- Factoring in the cost of the VALUE treatment, there may be a substantial financial benefit by creating greater customer loyalty (lower deactivations) through the use of a recurring VALUE + CONTACT discount treatment. Further investigation is required to determine if the opportunity cost of the CONTACT treatment outweighs the net value received through improved customer loyalty.

6 Conclusions

6.1 Introduction

This study had two objectives: 1) to determine if a causal linkage between customer satisfaction and loyalty (churn) can be found, and 2) to determine if there exists an ability to influence customer satisfaction and loyalty through specific, practical, and scalable treatments.

These objectives were motivated by the substantial efforts and resources expended by mobile phone service providers in the highly competitive Canadian telecommunications market. The providers spend over 20% of their revenues on costs to acquire and retain customers, and further costs are incurred in the building an operation of the wireless network, the provision of access to live call centre agents and the operation of retail stores in order to provide ongoing customer service.

One of the most critical customer measures that drives firm performance (both in terms of its financial performance as well as performance in the public equities market) is that of customer loyalty, reported by mobile phone providers as monthly churn or the percentage of the company's customer base that defects each month. Section 1.1.3 established that a single basis point of monthly churn (0.01% of the customer base) is equivalent to the three largest Canadian mobile phone providers each losing more than 10,000 customers per year, representing a lifetime revenue value of over $66 million.

Perceptions of customer satisfaction and other customer metrics related to loyalty also impact a company's ability to attract new customers, with the government providing additional scrutiny through requirements legislated in the highly regulated Canadian telecommunications industry.

6.2 Summary of Key Findings and Main Aims of the Study

To achieve the two aims of the study, an experimental, longitudinal, randomized controlled trial (RCT) was conducted with 35,000 customers of one of Canada's three largest mobile phone service providers. Over the course of 12 months, three treatments were deployed across eight study groups to measure the effects of the treatments on a series of independent (IV) and dependent variables (DV), and to evaluate a series of hypothesis statements about the relationships between those IVs and DVs.

The experiment was not able to support claims that customer satisfaction is causal to customer loyalty (Aim #1) as there was insufficient data from the experiment to conduct a statistical analysis between those two variables. The experiment was, however, able to provide evidence of two treatments that causally influence customer satisfaction and customer loyalty independently (Aim #2).

This evidence includes:

- Ability to positively influence participant frustration with marketing messaging through a reduction in the frequency of such messaging, and demonstration that the level of frustration is significantly related to customer satisfaction.
- Demonstration that the combination of two treatments (a 10% discount and reduced marketing frequency) together cause greater loyalty, as evidenced by the lower deactivation rates of participants receiving the two treatments. The treatment impact on loyalty was observed even though there was no significant impact observed on the intervening variables identified in the model between the treatment and the outcome (such as perceived value and customer satisfaction).

Partial evidence was also found for two hypotheses in the model:

- Consistent with prior research, a significant relationship was found between perceptions of value and perceptions of customer service to customer satisfaction, but the treatments employed in the experiment did not provide evidence that the perceptions themselves could be causally influenced.

6.3 Implications and Recommendations for Management Practice

The study suggests several important implications for managers of telecommunication companies providing mobile phone service; practitioners at other consumer service organizations may also want to note the outcomes of this study.

Many organizations measure the perceptions of their customers across a variety of attributes. Where practitioners implement programs intended to influence perceptions of value, practitioners should proceed with caution given the evidence provided in this study regarding the inability to causally influence perceived value through a discount similar in value to many customer discount programs.

If the organization does not see a change in the levels of perceived value reported by their customers despite expensive programs to do so, the inference may be that the programs are

unsuccessful. However, improvements in perceived value in and of themselves are usually not the end goal of the organization; rather it is the assumption (reasonably made based on common business experience and prior research reviewed in Chapter 2) that improving perceived value will result in greater customer satisfaction and greater customer loyalty. The study suggests that the latter outcome of improved loyalty may be achieved even if the intervening measures of perception do not show significant change.

While it can be shown that a treatment such as a 10% discount can causally improve customer loyalty, managers should ensure that they fully assess the net benefit of their efforts by calculating the financial gains made by the improvement to loyalty less the costs of providing the treatment. Measured with care and precision, the financial outcomes may be much closer to barely break-even (or even a loss) than they may think.

Marketing practitioners should take note of the evidence that an increased frequency of marketing messages sent to their customers significantly – and negatively – impacts customers' levels of frustration. This frustration has a significant relationship to customer satisfaction and is causally linked (together with the discount treatment) to customer loyalty. Since communicating to one's existing customer base is a primary tool to cross-sell and upsell additional products and services, marketers should try to determine the optimal level of communication frequency to their customers. They can do so by testing the sensitivity of their own customer base to communication frequency, following a similar experiment to that conducted in this study. They may also establish a contact strategy, informed by the results of their research, to specify how often each customer may be contacted in a given time period to ensure that customers are not over-communicated to and do not become frustrated. More advanced marketers may choose to invest in machine learning or artificial intelligence to develop propensity models that allow them to target the right offers to the right customers at the right time, with an expectation that this would allow them to manage customer frustration through reduced frequency of communication without forgoing the opportunities for additional sales.

6.4 Original Contributions of this Study

This thesis is one of the few large-scale experimental field studies of selected antecedents of customer satisfaction and customer loyalty, utilizing the "gold standard" of research design, the randomized controlled experiment. The study builds upon, and contributes new evidence to support, several findings made in previous research regarding the drivers of, and the relationship between, customer satisfaction and loyalty.

In the application of the three experimental treatments (VALUE, CSERV, and CONTACT), the study demonstrated how predicted antecedents of customer satisfaction and customer loyalty could be operationalized – at scale – in order to support a randomized controlled trial experiment involving 35,000 participants. The approach used here can now be replicated by future researchers, whether replicating this specific study or in pursuit of the examination of other antecedents in causal relationships. In addition, the study provides a "blueprint" of sorts for how researchers might approach an organization when proposing a study of that organization's customers and requiring support from that organization (financial or otherwise).

Notwithstanding the demonstration of the ability to implement treatments at scale, the study uniquely demonstrates the difficulty in causing participants to have different perceptions of value, customer service or frustration with marketing messaging (the three independent variables in this study). The examination of the ability of the treatments to affect these perceptions as factors impacting overall customer satisfaction and loyalty is a new contribution made by this study. Only one treatment (CONTACT) was shown to have a statistically significant impact on its related IV (MARKETING, or frustration with marketing message), indicating that researchers and practitioners alike face obstacles in determining how to affect customer perceptions.

The outcome of the CONTACT treatment as employed in this study is a new contribution to the field. Consumers are exposed to an increasing variety and volume of communications that they do business with (including both solicited and unsolicited communications). The evidence identified by this study that the frequency of those communications is negatively related to customer satisfaction (that is, more frequent communications makes customers unhappy) and perhaps most importantly, also to loyalty, should give practitioners pause to consider their marketing plans and provide researchers with an important area of future investigation.

The study also provides evidence to support previously documented relationships between the IVs of Perception of Value (PERCVALUE), Perception of Customer Service (SERVICE) and Frustration with Marketing Message (MARKETING) with that of the DV Satisfaction (CSAT). This additional evidence reinforces the value of the theoretical frameworks proposed by prior research that attempt to explain how customers come to be satisfied or not.

Perhaps most importantly, the outcome of the RCT experiment in this study demonstrated that customer loyalty can be causally influenced using selected treatments. While prior research has asserted the existence of these causal relationships, these same studies have not succeeded in meeting the rigid standards required to prove such causal relationships,

and not in a way that be generalized to as great a population of consumers as is the case with the study presented here.

Through the pursuit of its two aims, the study demonstrates that robust research methodologies, such as a randomized controlled study, can be deployed at scale in a real-world, field environment by a researcher or organization to determine its customers' behaviours in relation to customer satisfaction and loyalty. The study provides evidence that uniquely contributes to the examination of satisfaction and loyalty in the consumer marketplace, and does so through an examination of a random selection of participants from a large population, each randomly assigned to a series of treatments – a study design that is quite rare (Shadish et al. 2002).

A final contribution of the study is to demonstrate of how the results can be used to determine the net financial outcome of improved customer loyalty, using an analysis that compares the expected lifetime revenue of customers receiving and not receiving treatments. While small gains in the expected value of a single customer can translate into substantial benefits for an organization over the lifetime of that customer, this same value can be overstated if the cost of the treatment itself is not considered.

Overall, the study presents a unique approach to the examination of customer satisfaction and loyalty, and applies research methodologies not frequently seen in the fields of social sciences, brand and reputation management.

6.5 Limitations of the Study

There are two main limitations that constrained the ability of the experiment to fully respond to the aims of the study and the hypotheses set out in the model.

Firstly, and most impacting, was the lack of data to report on the relationship between the customer satisfaction and customer loyalty variables. None of the respondents to the posttest survey (providing the measure of the customer satisfaction variable) deactivated their service during the period of the study, which prevented the determination of whether a relationship (causal or otherwise) exists between customer satisfaction and customer loyalty. There is the potential presence of bias in the study if the only participants motivated to respond to the survey were those satisfied enough to maintain their service, thereby creating the situation encountered with the data.

The second limitation encountered was in the implementation of the treatment intended to influence the participants' perception of service by manipulating the wait time it took to reach

a live agent in the call centre. The overall effort by the subject organization to improve wait times for all customers at the same time that the study was conducted largely neutralized the treatment employed in the study.

Cost limitations prevented the expansion of the study beyond the design that was ultimately implemented. While the design of the study provided for substantive experimentation using multiple treatments across a large study group of 35,000 customers, the factorial design limited the levels employed in the discount treatment to two (no discount and 10%).

During the design phase of the study, it was contemplated to add a third level to this treatment (for example, a 15% discount level in addition to the two that were employed), which would have resulted in a 3x2x2 factorial design instead of a 2x2x2 design. The 3x2x2 factorial design would have required 50% more participants in the study, approximately 17,500 more than the 35,000 employed. This was not the limitation that prevented the pursuit of this design; rather it was the incremental cost of the VALUE treatment. With the increased number of participants in the study, the cost of the VALUE treatment with three levels would increase by 150%, resulting in a total cost of over $3,250,000. The researcher could not justify to the subject organization the increased expense over the already estimated $1,300,000 cost of the original design.

6.6 Agenda for Further Research

There are several areas where further research would be of benefit to researcher and practitioners by extending the insights derived from this study.

Further research is suggested to continue the investigation into the existence, if any, of a causal relationship between customer satisfaction and loyalty. Future research may elect to replicate the design methodology of this study but utilize either a larger study sample of participants or a longer period in the field in order to increase the opportunity for representation of deactivating customers in the posttest customer satisfaction survey.

Longer study periods should also be considered, especially if combined with repeated measures of customer satisfaction at regular intervals, as this would provide insights as to how customer behaviours evolve over time. An additional consideration for conducting a similar experiment over a longer timeframe would be the examination of whether the effect of any treatments on satisfaction or loyalty required the ongoing provision of the treatment (for example, the discount) in order to see sustained benefits, or whether there was a benefit of the treatment what would persist beyond the withdrawal of the treatment itself.

To address the issue of increasing costs of treatments when the number of participants or time the experiment is in the field are increased, future researchers may choose to use a fractional factorial design rather than the full factorial design employed in this study. The fractional factorial design allows for the use of a subset of treatment groups to be employed, thereby reducing the cost of the treatments.

The use of a fractional factorial design in future studies would also support the economical use of differing levels of discount treatments (such as 5%, 10%, 15%), which might provide insights as to the sensitivity of customers to varying levels of discounts and the subsequent differences in their loyalty. Future researchers should also contemplate the use of discount treatments that use absolute dollar values instead of the percentage discounts employed in this study. Future studies employing a discount treatment should also consider the use of repeated reminders throughout the study timeframe to reinforce the presence and value of the discount treatment (including calculations of cumulative savings).

On the other side of the perception of value ratio, future research should examine if the manipulation of the product or service benefits (rather than the cost) exert any influence on the customer's perception of value.

Future studies should also seek opportunities to re-test the hypothesis from this study regarding the ability to influence perceptions of customer service (and customer satisfaction and loyalty) by manipulating wait time to access a call centre agent. Such opportunities may exist wherever current (pre-experiment) wait times are sufficiently long enough that a treatment designed to reduce that wait time would be significantly different from the non-treatment experience. Alternatively, researchers may employ an different version of the treatment, increasing the wait time of customers in the treated groups relative to their non-treated peers in the study.

There are several opportunities to extend the findings in this study related to the impact of marketing communication frequency on customers' perceptions of frustration and behavioural loyalty. Future researchers should examine the impact of different levels of marketing frequency, such as 50% or 150% of that received by the control group of customers. The measurement of the opportunity cost of reduced marketing frequency should also be incorporated in future research to allow for the calculation of the net financial benefit of providing the treatment (similar to how the net financial benefit of the discount treatment and churn was calculated in this study). The examination of the impact of varying frequency of marketing communications combined with measurement of opportunity costs might provide insights into how organizations could optimize their financial outcomes.

Appendix A – Summary of Research Findings

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
Oliver (1980)	Health; Flu vaccination program	One time, following end of flu season	USA	Retrospective mail survey (2k residents + 1k students, random samples) Recursive path analysis	Satisfaction (6-item Likert scale) Questions were emotional in nature, including references to outright satisfaction, regret, happiness and general feelings about the decision	1) Pre-exposure variables for: - Expectations of Consequences (5-pt scale) - Attitude (9-pt semantic scale) - Behavioural Intent (11-pt scale) 2) Post-exposure variables for: - Disconfirmation of Expected Problems (2-item construct, 7-pt scale)	Disconfirmation is independent of all pre-exposure measures, and may be considered exogenous. Satisfaction is a function of disconfirmation and a linear combination of pre-exposure variables (predominantly disconfirmation). Satisfaction is the primary determinant of post-exposure variables. Disconfirmation appears to fit in the theoretical framework as proposed.
Churchill Jr et al. (1982)	Durable good (video disc player) and non-durable good (hybrid plant)	One time	USA	Experimental (126 subjects): 3 levels of expectations and 3 levels of performance	Satisfaction 5 variables on 7-point scales, incl. outright satisfaction, attribute-specific, faces and purchase probability.	1) Expectations 2) Performance 3) Disconfirmation (manipulated indirectly through manipulation of expectations	Manipulations in experiment of expectations and performance levels produced desired effects; F-tests were significant in each case, with good separation in mean response levels corresponding to each treatment condition.

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
				manipulated in a 3x3 factorial design Analysis of variance procedures (manipulation check) SEM LISREL		and performance; measured independently) Each construct consists of one global item ("Overall quality", for example) and one or more attribute-specific items ("Sound quality"; each using 7-pt scales	Non-durable findings: Relationships were as hypothesized; expectations had a negative effect on disconfirmation, and performance had a positive effect; disconfirmation positively affected satisfaction
Bolton et al. (1991)	Residential landline telephone	One-time 1985	USA	Survey Two-stage least squares estimation	Customer Satisfaction, Service Quality 5-pt, Likert-like scale	Multiple endogenous variables (5) and predictor variables (34) used in 5 model equations	Key determinant of service quality is gap between performance and expectations (i.e. disconfirmation) Primary determinants of Service Quality were attributes for BILLING, LOCAL service and LONG distance service; Most important determinant of Value was QUALITY
Anderson et al. (1993)	Database of telephone survey, administered to 22,300	Data collected in 1989 and 1990	Sweden	Phone survey data from database of 22,300 customers of 57 companies	Satisfaction 10-point scale from Low to High	Negative Disconfirmation, Positive Disconfirmation, Perceived Quality	Satisfaction is positively influenced by perceived quality and positive disconfirmation, and negatively influenced by negative disconfirmation; average estimated impact of perceived quality on

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
	customers of 57 companies; multiple industries spanned, covering 70% of firms			Seemingly-unrelated regression (SUR) method			satisfaction is greater than those for either positive or negative disconfirmation. Satisfaction is more sensitive to negative disconfirmation than that of positive disconfirmation in 80% of cases. Repurchase intentions positively influenced by level of satisfaction. Elasticity of repurchase intentions with respect to satisfaction is found to be lower for firms that provide high satisfaction
Rust et al. (1993)	Consumer banking	One-time	USA	Cross-sectional telephone survey of 100 customers Logit regression	Retention Binary measure of whether customer left the firm or not	Satisfaction Nine-item construct, 5-pt Likert scales	Different customer satisfaction elements have differing impacts on retention.
Andreassen et al. (1998)	Package tours (charters)	One-time telephone survey of customers	Norway	Reflective measurement model; structural	Loyalty 10-pt scale likelihood of positive word of mouth	Satisfaction, Value, Perceived Quality, Corporate Image	For complex and infrequently purchased services, corporate image rather than satisfaction is main predictor of loyalty (contrary to disconfirmation-of-expectations paradigm)

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
		who had travelled with tour operator in last 12 months		equation modelling (LISREL VIII, ML)			Value had no impact on satisfaction; corporate image had stronger effect on loyalty than satisfaction, as well as impact on satisfaction itself.
Bolton (1998)	Mobile phone users	Cross-sectional and time series data over 22-month period (1991-1993)		Two waves of telephone surveys, account information on deactivations; 599 complete records. Proportional hazards regression model.	Duration of provider-subscriber relationship	Cumulative Satisfaction	Customers who rate their prior cumulative satisfaction higher tend to have longer duration times. Customers with greater tenure with the company weight their cumulative satisfaction more heavily than shorter tenured customers. Satisfaction levels explain a substantial portion (26%) of explained variance in duration of relationships.
Bolton et al. (1999)	Cellular communications service, interactive television entertainment service	Panel studies for both groups, ranging from 6 to 12 months	USA	Two-stage least squares estimation; usage equation estimated using maximum likelihood procedure	Payment Equity, Satisfaction, Usage Satisfaction measured as self-reported on an	Various predictor variables for each dependent variable (payment equity = 7, satisfaction = 5, usage = 6)	Identified causal links among customers' prior usage, satisfaction and subsequent usage Strong positive relationship between payment equity and satisfaction

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
		Cellular service: Wave 1 Jan-Mar 1992 Wave 2 Jul-Sept 1992 Wave 3 Jan-Mar 1993			undefined scale (very dis/satisfied)		Customers with high levels of cumulative satisfaction will have higher usage levels in future
		Entertainment service: Wave 1 Nov 1993 Wave 2 Apr 1994					
Homburg et al. (2001)	Consumer durables (German cars)	One time questionnaire, 2 years post-purchase	Germany	Multiple-group LISREL	Loyalty Multi-item scales, measured as multi-dimensional construct of behavioural and attitudinal	Satisfaction; Various moderating variables on satisfaction-loyalty relationship (gender, age, income, etc.	Hypothesized effects of satisfaction on loyalty are supported. Particularly strong effects of product satisfaction on recommendation behaviour and repurchase intention. Moderator variables, in general, are relevant in the context of the satisfaction-loyalty link

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
					Likelihood to recommend, intention to repurchase		
Mittal et al. (2001)	Car market	One-time	USA	Mail survey	Repurchase Behaviour	Satisfaction, Repurchase Intent 5-point Likert-like scale Demographic Characteristics (sex, age, etc.)	Consumers of varying characteristics (gender, age, marital status, education, children, and area of residence) have different repurchase probabilities. Repurchase behaviour is less influenced by changes in satisfaction for customers with different characteristics.
Back et al. (2003)	Hotel	One-time	USA	Questionnaire of guests at a hotel; 194 usable responses	Behavioural Brand Loyalty Measured by proportion of all stays with hotel vs all others in past year	Customer Satisfaction 3-item, 7-point Likert-like scale	Customer Satisfaction positively influences Cognitive Brand Loyalty, which in turn positively influences Affective Brand Loyalty, and in turn positively influences Conative Brand Loyalty. The three attributes together (Attitudinal Brand Loyalty) positively influence Behavioural Brand Loyalty. Direct effect of Customer Satisfaction on Behaviour brand Loyalty was not significant.
Verhoef (2003)	Financial Services (insurance)	Unknown	Netherlands	Telephone survey over 2 waves, 1,128 respondents	Customer Retention, Customer Share	Satisfaction, Payment Equity, Affective Commitment Various multi-item scales	Affective commitment is an antecedent of Customer Retention and Customer Share.

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
							Satisfaction and Payment Equity had no effect on Retention or Share.
Lai (2004)	Mobile phone users (SMS service)			Mall interview; 150 respondents. PLS-SEM	Behaviour Intention 2-item measure, 3-point Likert-like scales	Customer Satisfaction, Perceived Value, Service Quality Multiple-item measures, various Likert-like scales	The tangibles, empathy and assurance dimensions of service quality are antecedents of customer satisfaction, and a positive relationship exists between customer satisfaction and customers' behavioural intentions. Perceived value was one of the determinants of customer satisfaction.
Tam (2004)	Family / chain restaurants		Hong Kong	Intercept sampling, self-administered questionnaire from 217 respondents. LISREL	Customer Satisfaction, Perceived Value, Post-Purchase Behaviour	Perceived Service Quality, Perceived Monetary Costs, Perceived Time Costs, Customer Satisfaction, Perceived Value	Perceived Value had relatively greater influence on Post-Purchase Behaviour than Customer Satisfaction
Yang et al. (2004)	Consumer banking	One-time	USA	Web-based survey, 235 respondents	Perceived Satisfaction, Customer Loyalty 5-point Likert-like scales	Customer Value, Perceived Satisfaction 5-point Likert-like scales	Customer Value and Perceived Satisfaction are powerful predictors of Customer Loyalty; Customer Value is also a predictor of Customer Satisfaction.

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
Aydin et al. (2005)	Mobile phone users	One-time questionnaire gathered over 6-week period	Turkey	Structural equation modelling (SEM)	Loyalty 5-item scale, repurchase intention, willingness to recommend, resistance to switching	Perceived service quality, Corporate image, Trust, Perceived switching cost	Path analysis shows that all factors have a positive effect on loyalty, however results reveal that trust is the most important determinant of customer loyalty. Findings show that perceived service quality is a necessary but not sufficient condition for customer loyalty to emerge and to exist.
Gustafsson et al. (2005)	Telecommunications (fixed phone, Internet, mobile phone)	Periodic survey & longitudinal account data, 9-month period 2003	Sweden	Survey; 2,715 respondents	Churn (deactivation of service)	Customer Satisfaction, Affective/Calculative Commitment 3-item measures, 10-point Likert-like scales	Customer satisfaction has a consistent negative effect on churn (a positive effect on retention). Calculative commitment, a construct not included in previous studies of retention, has a consistent negative effect on churn.
Yoon et al. (2005)	Tourism	One-time	Cyprus	Self-administered questionnaire; 148 usable responses. SEM	Destination Loyalty 3-item measure, 3- and 4-point Likert-like scales	Travel Satisfaction 4-item, with 4 or 5-point Likert-like scales	Tourist destination loyalty positively influenced by tourist satisfaction.

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
Bodet (2008)	Fitness clubs	One time questionnaire, with check for renewal of membership 3 months after end of subscription	France	Regression analyses and Logit analyses when the dependent variable was binary. SEM would have been preferred to test and compare the causal models, however two factors did not allow use of this method (moderating role of satisfaction in two of the models could not be tested, and binary nature of behavioural loyalty)	Attitudinal loyalty 3 items on 7-pt scale (probability of renewing, likelihood to recommend, would make same choice again) Repurchase behaviour Binary variable of whether customer re-purchased or not	Transaction-specific satisfaction, Overall satisfaction,	Overall satisfaction plays a mediating role between transaction-specific satisfaction and attitudinal loyalty. Contrary to Jones and Suh (2000), this study did not confirm that transaction-specific satisfaction directly influences customer attitudinal loyalty toward service providers. Contrary to Arrondo et al. (2002) and Bandyopadhyay and Martell (2007), this study found that attitudinal loyalty was unable to predict behavioural loyalty, in this case, repurchasing behaviour
Brunner et al. (2008)	Train travel	One-time, 2003	Germany	Self-administered questionnaire, randomly selected train runs over 6	Loyalty 2-item measure, 5-point Likert-like scales	Customer Satisfaction, Image Single item measures, 5-point Likert-like scales	Satisfaction positively influences loyalty; role of satisfaction is greater than image for new customers, while satisfaction is slightly less

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
				weeks; 941 usable responses. SEM			important (though still dominant) for experienced customers.
Hu et al. (2009)	Hotel industry		Mauritius	Interviews using structured questionnaire; 1,500 respondents.	Behavioural Intention, Image	Customer Satisfaction, Perceived Value, Service Quality	Service Quality positively impacts Perceived Value, and both in turn positively impact Customer Satisfaction. Customer Satisfaction positively impacts Image, which in turn positively impacts Behaviour Intention
Kuo et al. (2009)	Mobile phone users	One-time	Taiwan	Survey to university students; 387 respondents. SEM	Customer Satisfaction, Post-purchase Retention 5-point Likert-like scales	Customer Satisfaction, Service Quality, Perceived Value 5-point Likert-like scales	Perceived Value positively influences Customer Satisfaction. Perceived Value and Customer Satisfaction directly and positively influence Post-purchase Intention.
Flint et al. (2011)	Study 1: Purchasing managers across various industries	One time survey	USA	Structural equation modelling (SEM) LISREL	Loyalty Likert-like 4-pt scale (I feel loyal, I seek alternatives (R), I am committed to relationship)	Customer Value Anticipation (CVA), Satisfaction	Customer value anticipation does positively affect both customer satisfaction and customer loyalty, and also has a strong effect on customer loyalty by operating through customer satisfaction.

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
	Study 2: Purchasing agents and executives in manufacturers of analyzing instruments, electronic and electrical equipment						
Ha et al. (2013)	Mobile smartphones, netbooks	One-time online survey	Korea	Partial least squares (PLS)	Customer Satisfaction, Loyalty Satisfaction defined on 7-point scale across 4 questions (choice to purchase was a wise, overall satisfaction, did the right thing buying this, I feel bad about choice to buy this [reverse])	Alternative attractiveness, Utilitarian benefit, Hedonic benefit, Monetary cost, Nonmonetary cost	Effects of the two kinds of benefit, utilitarian benefit and hedonic benefit, on customer satisfaction are significant. The effect of nonmonetary cost and alternative attractiveness on customer satisfaction is significant; however, the effect of monetary cost on customer satisfaction is not. Customer satisfaction is a powerful predictor of customer loyalty; however, the relation between alternative attractiveness and customer loyalty is

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
							not proved.
					Loyalty defined on 7-pt scale across 4 items (product is first choice, will continue to buy, will say excellent/positive things to others)		
Chen et al. (2014)	Mobile telecommunications	One-time field survey	China	Partial least squares (PLS)	Loyalty 7-pt scale, 3 items, adapted from Zeithaml (1996); actual items unknown	Customer satisfaction (3 item measure), Customer dissatisfaction (2 item), Functional perception (4 item), Motivating perception (4 item), various service attributes (11 total) All constructs used 7-pt Likert scales	Customer satisfaction contributes significantly to customer loyalty whereas customer dissatisfaction has a negative influence on customer loyalty. The effect of CS on enhancing customer loyalty is greater than the diminishing influence of CDS on customer loyalty. The negative bias theory is thus not supported; perhaps there are some other reasons that restrain the negative influence of CDS on customer loyalty, such as switching cost, habit, and customer-perceived power imbalance. MP is found to have a strong influence on customer loyalty, while FP does not have a significant effect.

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
Wu (2014)	Casino patrons			Convenience sampling, self-administered questionnaire from 470 respondents Exploratory factor analysis, regression analysis	Customer Satisfaction, Perceived Value, Service Quality, Behavioural Intentions 7-pt Likert scales	Customer Satisfaction, Perceived Value, Service Quality, Interaction Quality, Physical Environment Quality, Outcome Quality 7-pt Likert scales	Perceived Value positively influences Customer Satisfaction. Customer Satisfaction positively influences Behavioural Intentions.
Calvo-Porral et al. (2015)	Mobile telecommunications	One time questionnaire, gathered in March 2013	Spain	Multiple group analysis through structural equation modelling (SEM)	Satisfaction, Loyalty, Switching intention Satisfaction measured through 5-pt Likert-like scale (overall satisfaction, decision was wise, receive what I expect, has satisfied my expectations)	Attractiveness of alternatives, Search effort, Service value, Corporate image	Evidence to propose a significant relationship among corporate image and service perceived value as sources of customer satisfaction. Relationship of customer satisfaction on loyalty is significantly positive. Lack of influence of switching costs on customer satisfaction for mobile network – traditional – mobile services.

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
					Loyalty 5-pt Likert-like scale for 2 items (would choose same company again, consider myself loyal to company)		
Gallarza et al. (2016)	Hotel	Two month period in 2012	Italy	Intercept sampling, interviewer-guided questionnaire; 585 respondents. PLS-SEM	Perceived Value (4 item construct), Customer Satisfaction (6 item), Customer Loyalty (4 item) All constructs used 7-pt Likert scales	Value Dimensions (4), Perceived Value, Customer Satisfaction	Significant and large links between Perceived Value and Customer Satisfaction, and between Customer Satisfaction and Customer Loyalty.
Roberts-Lombard et al. (2018)	Mobile phone users	One-time	South Africa	Non-probability sampling via field workers; self-administered questionnaire from 593 respondents. Exploratory factor analysis	Customer Satisfaction (7 item construct, semantic differential scale), Behavioural Intention (3 item construct, 7-pt Likert scale)	Perceived Value, Perceived Employee Service Delivery Skills, Customer Satisfaction	Customer Satisfaction positively influences Behavioural Intention

Study	Industry / Product	Time Period of Research	Country of Research	Methodological Approach	Dependent Variable	Independent Variables	Key Results
El-Adly (2019)	Hotel	One-time survey	UAE	Intercept sampling and online survey for 305 respondents. SEM	Customer Satisfaction, Loyalty Each construct with 3 items, measured on 5-pt Likert scale	Customer Satisfaction, Perceived Value Multi-dimensional construct comprised of 7 factors	Perceived Value and Customer Satisfaction are significant antecedents to Loyalty.
Dey et al. (2020)	Mobile telecom subscribers	One-time survey in 2016	UK	Email survey to 1.2K subscribers Partial Least Squares – Structural Equation Modelling	Customer Satisfaction, Switching Intention (both single item constructs measured on 5-pt Likert scales)	Network Performance, Brand Image, Perceived Value, Service Quality, Customer Service, Switching Cost & Barriers Various single construct, 3-point and 5-point scales	Only non-significant path coefficient was between Switching Cost & Barriers and Switching Intention. Perceived Value had greater impact on Customer Satisfaction than Customer Service. Customer Satisfaction negatively and significantly impacts Switching Intention.

Appendix B – DEFECT DV and Related Measures

The following data elements are obtained for each participant in the study on a monthly basis.

Data Field	Description	Possible Values
Deactivation Flag	Indicates if participant deactivated service	0 = did not deactivate 1 = deactivated service
If Deactivation Flag = 1, then the following fields are populated:		
Deactivation Date	Date service was deactivated	
Deactivation Type	Indicates whether the participant initiated the deactivation or the subject organization	Voluntary = participant deactivated service and/or moved to another carrier Involuntary = subject organization deactivated service (e.g. for non-payment)
Deactivation Reason	If participant calls subject organization to cancel their service, the call centre agent attempts to determine (based on customer response) the reason for the cancellation of service	Alternate Service Billing Cancel to get acquisition Deceased Expense / Financial Left Service Area Network Issues No Reason Given Non-Payment Not Required - Senior Not required / don't use Now Supplied by Job Ported Out Unknown Competitor - Hardware Unknown Competitor - Price

Deactivation channel	The type of channel that the deactivation request from the participant was received in	Independent Dealer Internal Direct Kiosk (Corp Stores) Retail Unknown
Number port-out flag	Indicates whether the deactivation was a result of the participant porting their wireless phone number to another carrier	N = did not port Y = ported number
Port-out carrier name	If the deactivation was a result of a port out to another carrier, this field lists the name of the carrier that service was ported to	Bell Rogers TELUS Fido Koodo Videotron Freedom

Appendix C – Customer Satisfaction Survey

This short survey is to understand your experience with TELUS as a wireless service provider.

Q1. How would you rate your **overall experience** with TELUS as a wireless service provider? Select one.

1. Excellent
2. Very good
3. Good
4. Fair
5. Poor

Q2. If a colleague, friend or family member were looking for a new wireless provider, what is the likelihood that you would **recommend** TELUS to them? Select one.

1. Definitely
2. Probably
3. Maybe
4. Probably Not
5. Definitely Not
6. Don't know

Q3. In the past year, **how many times** have you recommended TELUS for wireless service to colleagues, friends or family members? Select one.

1. Zero
2. Once
3. 2 or 3 times
4. 4 or 5 times
5. More than 5 times
6. Don't know

Q4. How has your **opinion** of TELUS as a wireless service provider changed in the past year? Select one.

1. Improved
2. Stayed the same
3. Worsened

Q5. Please select the statement below that best reflects your current situation with TELUS. Select one.

1. I plan on staying with TELUS
2. For the right deal, I would switch wireless service providers
3. I'm actively searching for a new wireless service provider

Q6. Considering the overall quality and the overall price you pay, how would you rate TELUS for **overall value** in terms of being worth what you pay for? Select one.

1. Excellent
2. Very good
3. Good
4. Fair
5. Poor
6. Unsure

Q7. How often are you **frustrated** with the following with TELUS? Choose one for each row.

COLUMNS

1. Very often
2. Often
3. Sometimes
4. Rarely
5. Not at all
6. Unsure / Not applicable

ROWS

- Network quality (e.g. fastest, coverage, reliability)
- Concern about exceeding data allowance
- How quickly you use your data
- Lack of communication on data overage
- Customer service quality (e.g. easy or friendly to deal with)
- Ability to self-serve
- Communication quality (e.g. relevance or frequency of emails, other marketing)
- Roaming (North American or International)
- Managing a share plan
- Billing
- Plan pricing compared to Rogers or Bell
- Features included in plans compared to Rogers or Bell
- Ease of dealing with them
- Flexibility to adapt to your needs
- How long you have to wait to speak to a representative

Q8. Please indicate which of the following you have done in the **past 6 months**. Select all that apply.

1. Contacted customer service over the phone
2. Activated a cellphone
3. Needed a cellphone repair
4. Visited a TELUS retail location
5. Changed your device / renewed your contract for wireless services
6. Used TELUS online chat to speak with a representative
7. Interacted with the TELUS Virtual Assistant / Chatbot
8. Visited the TELUS website or app to get help
9. Visited the TELUS website or app to check my account details
10. Visited the TELUS website or app to check my data usage
11. Self-served through the TELUS website or app to top up data limit
12. Used a TELUS Roaming service while travelling
13. Used a social media platform to get help or ask a question to TELUS
14. None of the above

Q9. Thinking of your most recent interaction, how would you rate your experience with each of the below? Select one for each.

ROWS – STATEMENTS DEPENDENT ON Q8

- **IF CODE 1** “Time you waited to speak to a representative”
- **IF CODE 1** “Automated voice attendant system used to answer/direct calls (e.g. easy to use/connect to a live person/provide relevant information)”
- **IF CODE 2** “Activating a cellphone”

- **IF CODE 3** “Repairing your cellphone”
- **IF CODE 1**, 4, 6 “The knowledge of the representative you spoke with”
- **IF CODE 4** “The overall store experience”
- **IF CODE 5** “changing your device / renewing your contract”
- **IF CODE 7** ”TELUS Virtual Assistant / Chatbot”
- **IF CODE 8** “Resolving your question through the TELUS Website or App”
- **IF CODE 9** “Ease of checking your account details”
- **IF CODE 10** “Ease of checking data usage”
- **IF CODE 11** “Ease of topping up data”
- **IF CODE 12** “Ease of roaming with TELUS”
- **IF CODE 13** “Resolved my issue through social media with TELUS”

COLUMNS

1. Excellent
2. Very good
3. Good
4. Fair
5. Poor

Q10. In the 6 months, have you reached out to TELUS with an issue you needed help with? Select one.

1. Yes
2. No

Q11. And was that issue resolved? Select one.

1. Yes
2. No
3. Somewhat

Thank you for your time today!

Appendix D – Satisfaction Survey Response Scales

Each of the 11 questions in the satisfaction survey utilizes one of the following response scales.

Q1 – Scale A

Q2 – Scale B

Q3 – Scale C

Q4 – Scale D

Q5 – Scale E

Q6 – Scale F

Q7 – Scale G

Q8 – Scale H

Q9 – Scale I

Q10 – Scale J

Q11 – Scale K

The values shown in the table below represent the values as recorded in the data tables of survey responses.

Scale A	Description	Value
	Excellent	1
	Very Good	2
	Good	3
	Fair	4
	Poor	5

Scale B	Description	Value
	Definitely	1
	Probably	2
	Maybe	3
	Probably not	4
	Definitely not	5

	Don't know	6

Scale C	Description	Value
	Zero	1
	Once	2
	2 or 3 times	3
	4 or 5 times	4
	More than 5 times	5
	Don't know	6

Scale D	Description	Value
	Improved	1
	Stayed the same	2
	Worsened	3

Scale E	Description	Value
	I plan on staying with TELUS	1
	For the right deal, I would switch wireless service providers	2
	I'm actively searching for a new wireless service provider	3

Scale F	Description	Value
	Excellent	1
	Very good	2
	Good	3
	Fair	4
	Poor	5

	Unsure	6

Scale G	Description	Value
	Very often	1
	Often	2
	Sometimes	3
	Rarely	4
	Not at all	5
	Unsure / Not applicable	6

Scale H	Description	Value
	Indicated Yes	1
	Indicated No	0

Scale I	Description	Value
	Excellent	1
	Very good	2
	Good	3
	Fair	4
	Poor	5

Scale J	Description	Value
	Yes	1
	No	2

Scale K	Description	Value
	Yes	1

No	2
Somewhat	3

Appendix E – Email Invitation to Satisfaction Survey

The following email was sent to each study participant for the pretest and posttest satisfaction survey.

From: **The TELUS Team** telus@confirmit.ca
Subject: Survey - Your opinion matters to us!
Date: March 11, 2019 at 9:41 PM
To: [redacted]

TT

TELUS®

Trouble viewing this email? Click here to start survey

At TELUS,
we're all ears.

Hello,

At TELUS, we're always looking for ways to improve things, and that starts with listening to our customers. We have a short, 11 question survey that should take 2 to 3 minutes of your time that we need your feedback on to understand your customer experience with TELUS and what we can do to make it better.

All responses are private, confidential, and will only be used for the purposes of enhancing our service.

Simply click on the link below to begin the survey.

https://web.cicicresearch.com/wix/p1079658.aspx?__sid__=Wvh_0c0WWXOlzOyHu_AQh3gSrn6CKXVZQ2VvOHezcll3FBPtnAxN4KTFY4x4gp0QVJf16u-8sfXYVIsoPmI1Ow2&dnc=0&l=9

We realize that taking a survey may be inconvenient for you, so as a gesture of appreciation we'd like to offer you the chance to win one of two $250 Visa gift cards!

Please visit here for contest rules and details.

Sincerely,

The TELUS Team

Responses to this survey, and existing account information, allow us to reach out to you to provide the best possible service. This survey is administered by **CICIC Research**, an independent research company.

You are subscribed as [redacted]
If you no longer wish to receive TELUS communications, please unsubscribe here
Marketing Preferences, 30th Floor, 25 York St. Toronto, ON M5J 2V5
TELUS Privacy Policy

Appendix F – ANOVA Post Hoc Analysis (MARKETING)

ANOVA Tukey Post Hoc Analysis – MARKETING Independent Variable

Multiple Comparisons

Dependent Variable: Posttest_Q7_7
Tukey HSD

(I) Treatment Group	(J) Treatment Group	Mean Difference (I–J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
CTRL	DISC	.023	.119	1.000	-.34	.38
	IVR	-.161	.125	.901	-.54	.22
	DISC+IVR	.126	.115	.957	-.22	.47
	DBM	-.027	.117	1.000	-.38	.33
	DISC+DBM	-.226	.115	.501	-.57	.12
	IVR+DBM	-.234	.118	.491	-.59	.12
	DISC+IVR+DBM	-.231	.115	.468	-.58	.12
DISC	CTRL	-.023	.119	1.000	-.38	.34
	IVR	-.184	.124	.814	-.56	.19
	DISC+IVR	.103	.114	.986	-.24	.45
	DBM	-.050	.117	1.000	-.40	.30
	DISC+DBM	-.249	.114	.358	-.59	.10
	IVR+DBM	-.258	.117	.353	-.61	.10
	DISC+IVR+DBM	-.255	.114	.329	-.60	.09
IVR	CTRL	.161	.125	.901	-.22	.54
	DISC	.184	.124	.814	-.19	.56
	DISC+IVR	.287	.120	.242	-.08	.65
	DBM	.134	.123	.958	-.24	.51
	DISC+DBM	-.064	.120	.999	-.43	.30
	IVR+DBM	-.073	.123	.999	-.45	.30
	DISC+IVR+DBM	-.070	.120	.999	-.43	.29
DISC+IVR	CTRL	-.126	.115	.957	-.47	.22
	DISC	-.103	.114	.986	-.45	.24
	IVR	-.287	.120	.242	-.65	.08
	DBM	-.153	.112	.872	-.49	.19
	DISC+DBM	-.352*	.109	.028	-.68	-.02
	IVR+DBM	-.360*	.113	.031	-.70	-.02
	DISC+IVR+DBM	-.358*	.109	.024	-.69	-.03
DBM	CTRL	.027	.117	1.000	-.33	.38
	DISC	.050	.117	1.000	-.30	.40
	IVR	-.134	.123	.958	-.51	.24
	DISC+IVR	.153	.112	.872	-.19	.49
	DISC+DBM	-.199	.112	.640	-.54	.14
	IVR+DBM	-.207	.116	.626	-.56	.14
	DISC+IVR+DBM	-.204	.112	.605	-.54	.14
DISC+DBM	CTRL	.226	.115	.501	-.12	.57
	DISC	.249	.114	.358	-.10	.59
	IVR	.064	.120	.999	-.30	.43
	DISC+IVR	.352*	.109	.028	.02	.68
	DBM	.199	.112	.640	-.14	.54
	IVR+DBM	-.009	.113	1.000	-.35	.33
	DISC+IVR+DBM	-.006	.109	1.000	-.34	.33
IVR+DBM	CTRL	.234	.118	.491	-.12	.59
	DISC	.258	.117	.353	-.10	.61
	IVR	.073	.123	.999	-.30	.45
	DISC+IVR	.360*	.113	.031	.02	.70
	DBM	.207	.116	.626	-.14	.56
	DISC+DBM	.009	.113	1.000	-.33	.35
	DISC+IVR+DBM	.003	.113	1.000	-.34	.35
DISC+IVR+DBM	CTRL	.231	.115	.468	-.12	.58
	DISC	.255	.114	.329	-.09	.60
	IVR	.070	.120	.999	-.29	.43
	DISC+IVR	.358*	.109	.024	.03	.69
	DBM	.204	.112	.605	-.14	.54
	DISC+DBM	.006	.109	1.000	-.33	.34
	IVR+DBM	-.003	.113	1.000	-.35	.34

Appendix G – CSAT Regression Model (with Predictor Variables) Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	2.219	0.121		18.298	0	1.981	2.456					
	tenure_years	-0.007	0.004	-0.042	-1.505	.133	-0.015	0.002	0.008	-0.041	-0.04	0.898	1.114
	CreditRiskMed	-0.058	0.162	-0.01	-0.358	.72	-0.376	0.26	-0.013	-0.01	-0.009	0.982	1.018
	CreditRiskHigh	-0.48	0.319	-0.041	-1.508	.132	-1.105	0.145	-0.034	-0.041	-0.04	0.969	1.032
	RevenueMed	-0.027	0.116	-0.013	-0.231	.818	-0.255	0.201	0.014	-0.006	-0.006	0.237	4.219
	RevenueHi	-0.116	0.115	-0.059	-1.011	.312	-0.341	0.109	-0.031	-0.027	-0.027	0.206	4.848
	RevenueVeryHi	0.139	0.181	0.026	0.771	.441	-0.215	0.494	0.05	0.021	0.02	0.633	1.581
	RevenueOther	-0.238	0.236	-0.031	-1.008	.313	-0.702	0.225	-0.011	-0.027	-0.027	0.762	1.313
	RegionPrairies	-0.345	0.101	-0.097	-3.399	.001	-0.544	-0.146	-0.036	-0.092	-0.09	0.859	1.164
	RegionOntario	-0.271	0.066	-0.118	-4.105	.000	-0.400	-0.141	-0.033	-0.111	-0.109	0.848	1.179
	RegionQuebec	-0.515	0.072	-0.212	-7.186	.000	-0.655	-0.374	-0.15	-0.192	-0.19	0.807	1.239
	RegionAtlantic	-0.366	0.123	-0.081	-2.978	.003	-0.607	-0.125	-0.04	-0.081	-0.079	0.942	1.061

	RegionOther	-0.726	0.544	-0.035	-1.334	.182	-1.793	0.341	-0.025	-0.036	-0.035	0.993	1.007
	Rural	-0.03	0.063	-0.013	-0.483	.629	-0.153	0.092	-0.031	-0.013	-0.013	0.971	1.03
	ContractYes	-0.05	0.056	-0.026	-0.879	.38	-0.16	0.061	-0.043	-0.024	-0.023	0.827	1.209
2	(Constant)	2.236	0.14		15.992	.000	1.961	2.51					
	tenure_years	-0.007	0.004	-0.042	-1.514	.130	-0.015	0.002	0.008	-0.041	-0.04	0.895	1.117
	CreditRiskMed	-0.048	0.163	-0.008	-0.297	.767	-0.368	0.271	-0.013	-0.008	-0.008	0.977	1.024
	CreditRiskHigh	-0.482	0.32	-0.041	-1.506	.132	-1.109	0.146	-0.034	-0.041	-0.04	0.966	1.035
	RevenueMed	-0.025	0.117	-0.012	-0.218	.827	-0.254	0.203	0.014	-0.006	-0.006	0.237	4.226
	RevenueHi	-0.117	0.115	-0.059	-1.015	.310	-0.343	0.109	-0.031	-0.028	-0.027	0.206	4.86
	RevenueVeryHi	0.138	0.181	0.025	0.759	.448	-0.218	0.493	0.05	0.021	0.02	0.631	1.585
	RevenueOther	-0.243	0.237	-0.031	-1.024	.306	-0.709	0.223	-0.011	-0.028	-0.027	0.758	1.32
	RegionPrairies	-0.347	0.102	-0.098	-3.404	.001	-0.546	-0.147	-0.036	-0.092	-0.09	0.856	1.168
	RegionOntario	-0.27	0.066	-0.118	-4.064	.000	-0.4	-0.14	-0.033	-0.11	-0.108	0.842	1.188
	RegionQuebec	-0.511	0.072	-0.21	-7.107	.000	-0.652	-0.37	-0.15	-0.19	-0.189	0.805	1.242
	RegionAtlantic	-0.362	0.123	-0.081	-2.941	.003	-0.604	-0.121	-0.04	-0.08	-0.078	0.939	1.065
	RegionOther	-0.719	0.546	-0.035	-1.315	.189	-1.79	0.353	-0.025	-0.036	-0.035	0.989	1.011
	Rural	-0.031	0.063	-0.013	-0.496	.620	-0.154	0.092	-0.031	-0.014	-0.013	0.967	1.034
	ContractYes	-0.049	0.057	-0.025	-0.87	.385	-0.16	0.062	-0.043	-0.024	-0.023	0.822	1.216

VALUE	0.025	0.106	0.008	0.231	.817	-0.183	0.232	0.02	0.006	0.006	0.548	1.826
CSERV	-0.035	0.111	-0.011	-0.315	.753	-0.252	0.182	-0.004	-0.009	-0.008	0.588	1.7
VALUE_CSERV	-0.017	0.103	-0.006	-0.17	.865	-0.219	0.184	-0.006	-0.005	-0.005	0.518	1.931
CONTACT	-0.009	0.105	-0.003	-0.087	.931	-0.215	0.197	0.004	-0.002	-0.002	0.539	1.855
VALUE_CONTACT	-0.06	0.102	-0.022	-0.593	.554	-0.26	0.139	-0.023	-0.016	-0.016	0.518	1.93
CSERV_CONTACT	0.021	0.104	0.007	0.197	.844	-0.184	0.225	0.019	0.005	0.005	0.535	1.869
VALUE_CSERV_CONTACT	-0.054	0.102	-0.02	-0.529	.597	-0.254	0.146	-0.02	-0.014	-0.014	0.516	1.936

Appendix H – ANOVA Post Hoc Analysis (DEFECT)

ANOVA Tukey Post Hoc Analysis – DEFECT Dependent Variable

Multiple Comparisons

Dependent Variable: cumulative_churn_Apr19_Mar20

	(I) study_group	(J) study_group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	CONTROL	VALUE	.00846	.00687	.923	-.0124	.0293
		CSERV	-.00304	.00682	1.000	-.0237	.0176
		VALUE_CSERV	.00722	.00685	.966	-.0135	.0280
		CONTACT	-.00897	.00682	.893	-.0296	.0117
		VALUE_CONTACT	.01325	.00683	.522	-.0074	.0340
		CSERV_CONTACT	-.00103	.00682	1.000	-.0217	.0196
		VALUE_CSERV_CONTACT	.00687	.00683	.974	-.0138	.0276
	VALUE	CONTROL	-.00846	.00687	.923	-.0293	.0124
		CSERV	-.01150	.00687	.705	-.0323	.0093
		VALUE_CSERV	-.00124	.00689	1.000	-.0221	.0197
		CONTACT	-.01743	.00686	.179	-.0382	.0034
		VALUE_CONTACT	.00479	.00687	.997	-.0160	.0256
		CSERV_CONTACT	-.00949	.00687	.866	-.0303	.0113
		VALUE_CSERV_CONTACT	-.00159	.00687	1.000	-.0224	.0192
	CSERV	CONTROL	.00304	.00682	1.000	-.0176	.0237
		VALUE	.01150	.00687	.705	-.0093	.0323
		VALUE_CSERV	.01026	.00685	.809	-.0105	.0310
		CONTACT	-.00593	.00682	.989	-.0266	.0147
		VALUE_CONTACT	.01629	.00683	.248	-.0044	.0370
		CSERV_CONTACT	.00201	.00682	1.000	-.0187	.0227
		VALUE_CSERV_CONTACT	.00990	.00683	.834	-.0108	.0306
	VALUE_CSERV	CONTROL	-.00722	.00685	.966	-.0280	.0135
		VALUE	.00124	.00689	1.000	-.0197	.0221
		CSERV	-.01026	.00685	.809	-.0310	.0105
		CONTACT	-.01619	.00684	.258	-.0369	.0046
		VALUE_CONTACT	.00604	.00685	.988	-.0147	.0268
		CSERV_CONTACT	-.00825	.00685	.931	-.0290	.0125
		VALUE_CSERV_CONTACT	-.00035	.00685	1.000	-.0211	.0204
	CONTACT	CONTROL	.00897	.00682	.893	-.0117	.0296
		VALUE	.01743	.00686	.179	-.0034	.0382
		CSERV	.00593	.00682	.989	-.0147	.0266
		VALUE_CSERV	.01619	.00684	.258	-.0046	.0369
		VALUE_CONTACT	.02222*	.00682	.025	.0015	.0429
		CSERV_CONTACT	.00794	.00682	.942	-.0127	.0286
		VALUE_CSERV_CONTACT	.01584	.00682	.282	-.0048	.0365
	VALUE_CONTACT	CONTROL	-.01325	.00683	.522	-.0340	.0074
		VALUE	-.00479	.00687	.997	-.0256	.0160
		CSERV	-.01629	.00683	.248	-.0370	.0044
		VALUE_CSERV	-.00604	.00685	.988	-.0268	.0147
		CONTACT	-.02222*	.00682	.025	-.0429	-.0015
		CSERV_CONTACT	-.01428	.00683	.420	-.0350	.0064
		VALUE_CSERV_CONTACT	-.00639	.00683	.983	-.0271	.0143
	CSERV_CONTACT	CONTROL	.00103	.00682	1.000	-.0196	.0217
		VALUE	.00949	.00687	.866	-.0113	.0303
		CSERV	-.00201	.00682	1.000	-.0227	.0187
		VALUE_CSERV	.00825	.00685	.931	-.0125	.0290
		CONTACT	-.00794	.00682	.942	-.0286	.0127
		VALUE_CONTACT	.01428	.00683	.420	-.0064	.0350
		VALUE_CSERV_CONTACT	.00789	.00683	.944	-.0128	.0286
	VALUE_CSERV_CONTACT	CONTROL	-.00687	.00683	.974	-.0276	.0138
		VALUE	.00159	.00687	1.000	-.0192	.0224
		CSERV	-.00990	.00683	.834	-.0306	.0108
		VALUE_CSERV	.00035	.00685	1.000	-.0204	.0211
		CONTACT	-.01584	.00682	.282	-.0365	.0048
		VALUE_CONTACT	.00639	.00683	.983	-.0143	.0271
		CSERV_CONTACT	-.00789	.00683	.944	-.0286	.0128

Appendix I – DEFECT Regression Model (with Predictor Variables) Coefficients

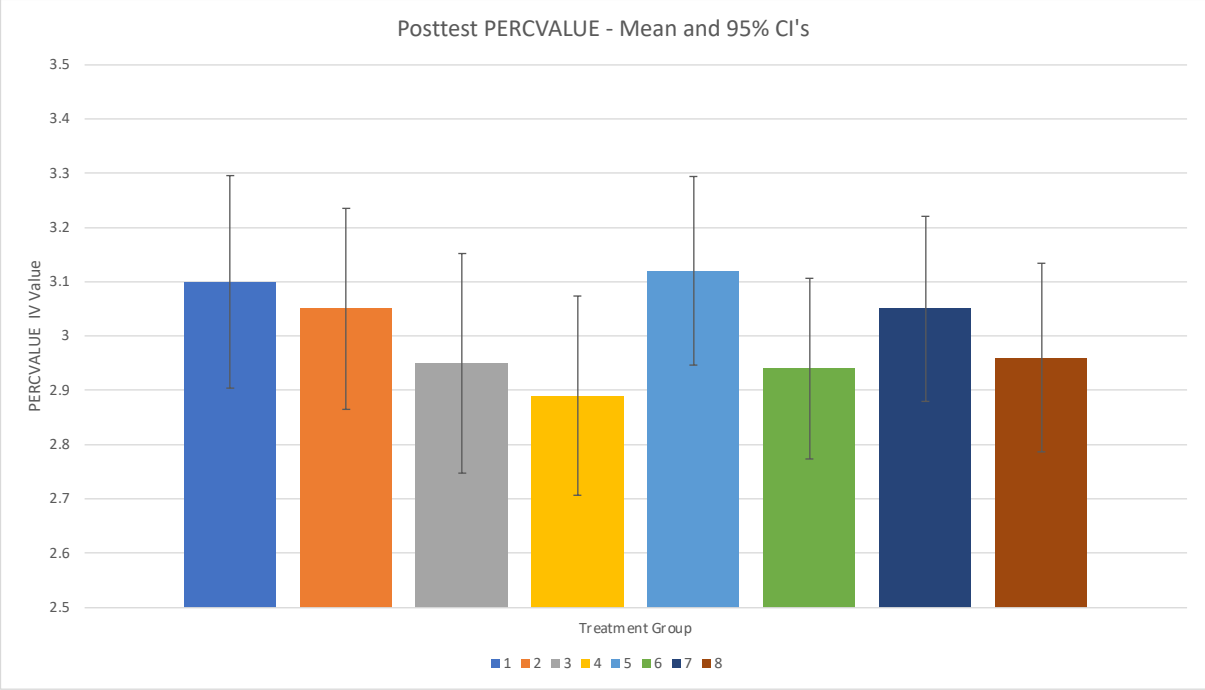
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.192	.007		27.862	.000	.178	.205					
	tenure_years	-.004	.000	-.076	-13.416	.000	-.005	-.003	-.096	-.072	-.071	0.876	1.141
	CreditRiskMed	.047	.009	.027	5.117	.000	.029	.065	.036	.027	.027	0.972	1.029
	CreditRiskHigh	.173	.014	.069	12.365	.000	.145	.200	.082	.066	.065	0.896	1.116
	RevenueMed	-.073	.007	-.101	-11.004	.000	-.086	-.060	-.009	-.059	-.058	0.331	3.018
	RevenueHi	-.083	.006	-.127	-13.007	.000	-.095	-.070	-.036	-.069	-.069	0.291	3.437
	RevenueVeryHi	-.075	.010	-.049	-7.429	.000	-.095	-.056	.005	-.040	-.039	0.654	1.528
	RevenueOther	-.035	.014	-.015	-2.528	.011	-.062	-.008	.046	-.014	-.013	0.744	1.344
	RegionPrairies	.029	.007	.024	4.169	.000	.015	.043	.028	.022	.022	0.878	1.139
	RegionOntario	.037	.004	.047	8.424	.000	.029	.046	.034	.045	.045	0.886	1.128
	RegionQuebec	.029	.005	.033	5.837	.000	.019	.039	.025	.031	.031	0.858	1.165

	RegionAtlantic	.042	.008	.028	5.171	.000	.026	.058	.023	.028	.027	0.953	1.049
	RegionOther	-.031	.034	-.005	-.916	.360	-.097	.035	-.004	-.005	-.005	0.995	1.005
	Rural	-.017	.004	-.021	-3.993	.000	-.025	-.008	-.012	-.021	-.021	0.982	1.018
	ContractYes	.024	.004	.036	6.276	.000	.017	.032	.028	.034	.033	0.832	1.202
2	(Constant)	0.195	0.008		23.751	.000	0.179	0.211					
	tenure_years	-0.004	0	-0.075	-13.369	.000	-0.005	-0.003	-0.096	-0.071	-0.071	0.876	1.142
	CreditRiskMed	0.047	0.009	0.027	5.107	.000	0.029	0.065	0.036	0.027	0.027	0.971	1.029
	CreditRiskHigh	0.173	0.014	0.069	12.37	.000	0.145	0.2	0.082	0.066	0.065	0.896	1.117
	RevenueMed	-0.073	0.007	-0.101	-10.995	.000	-0.086	-0.06	-0.009	-0.059	-0.058	0.331	3.019
	RevenueHi	-0.083	0.006	-0.127	-13.007	.000	-0.095	-0.07	-0.036	-0.069	-0.069	0.291	3.437
	RevenueVeryHi	-0.075	0.01	-0.048	-7.397	.000	-0.095	-0.055	0.005	-0.04	-0.039	0.654	1.529
	RevenueOther	-0.035	0.014	-0.015	-2.507	0.012	-0.062	-0.008	0.046	-0.013	-0.013	0.744	1.344
	RegionPrairies	0.029	0.007	0.024	4.176	.000	0.015	0.043	0.028	0.022	0.022	0.878	1.139
	RegionOntario	0.038	0.004	0.048	8.468	.000	0.029	0.046	0.034	0.045	0.045	0.886	1.128
	RegionQuebec	0.029	0.005	0.033	5.862	.000	0.019	0.039	0.025	0.031	0.031	0.858	1.166
	RegionAtlantic	0.042	0.008	0.028	5.162	.000	0.026	0.058	0.023	0.028	0.027	0.953	1.05
	RegionOther	-0.031	0.034	-0.005	-0.907	0.364	-0.097	0.035	-0.004	-0.005	-0.005	0.995	1.005
	Rural	-0.017	0.004	-0.021	-3.987	.000	-0.025	-0.008	-0.012	-0.021	-0.021	0.982	1.018

ContractYes	0.024	0.004	0.036	6.25	.000	0.016	0.031	0.028	0.033	0.033	0.832	1.202
VALUE	-0.008	0.007	-0.008	-1.213	0.225	-0.022	0.005	-0.007	-0.006	-0.006	0.577	1.732
CSERV	0.003	0.007	0.003	0.463	0.644	-0.01	0.016	0.007	0.002	0.002	0.572	1.749
VALUE_CSERV	-0.008	0.007	-0.008	-1.187	0.235	-0.021	0.005	-0.005	-0.006	-0.006	0.575	1.74
CONTACT	0.007	0.007	0.008	1.102	0.271	-0.006	0.021	0.014	0.006	0.006	0.571	1.751
VALUE_CONTACT	-0.013	0.007	-0.013	-1.894	0.058	-0.026	0	-0.012	-0.01	-0.01	0.572	1.747
CSERV_CONTACT	0.001	0.007	0.002	0.216	0.829	-0.012	0.015	0.005	0.001	0.001	0.572	1.75
VALUE_CSERV_CONTACT	-0.008	0.007	-0.008	-1.173	0.241	-0.021	0.005	-0.005	-0.006	-0.006	0.572	1.747

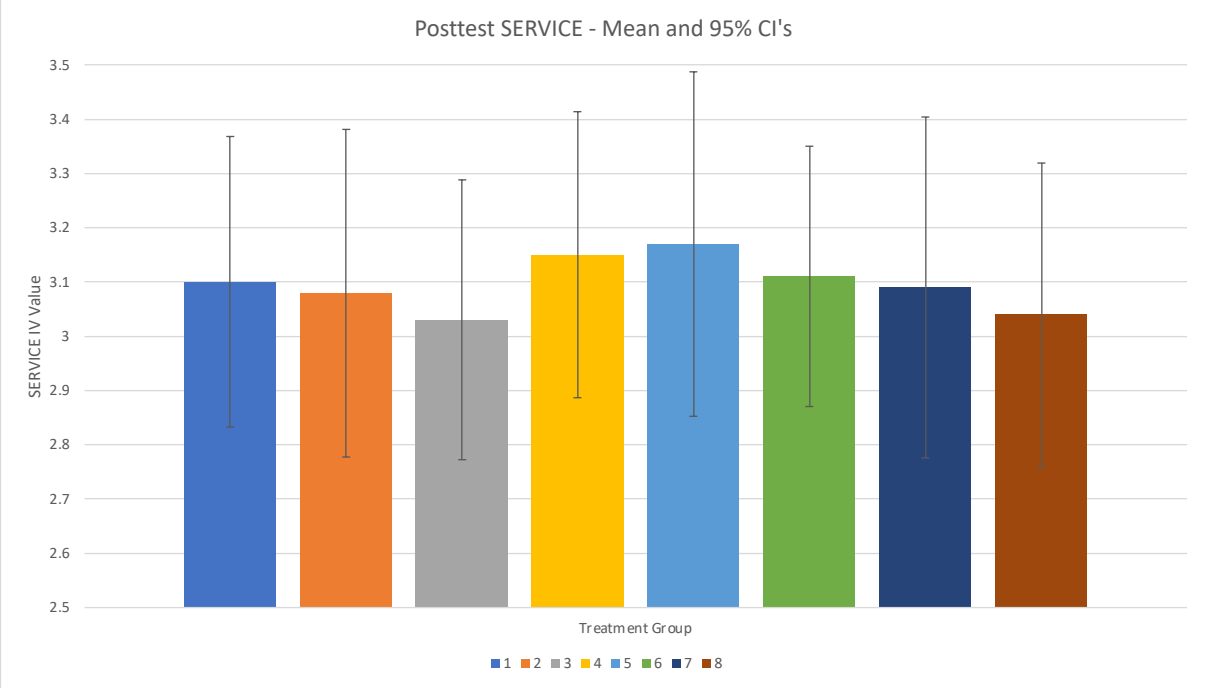
Appendix J – Analysis of Results of Practical Significance

Posttest Confidence Intervals – PERCVALUE IV



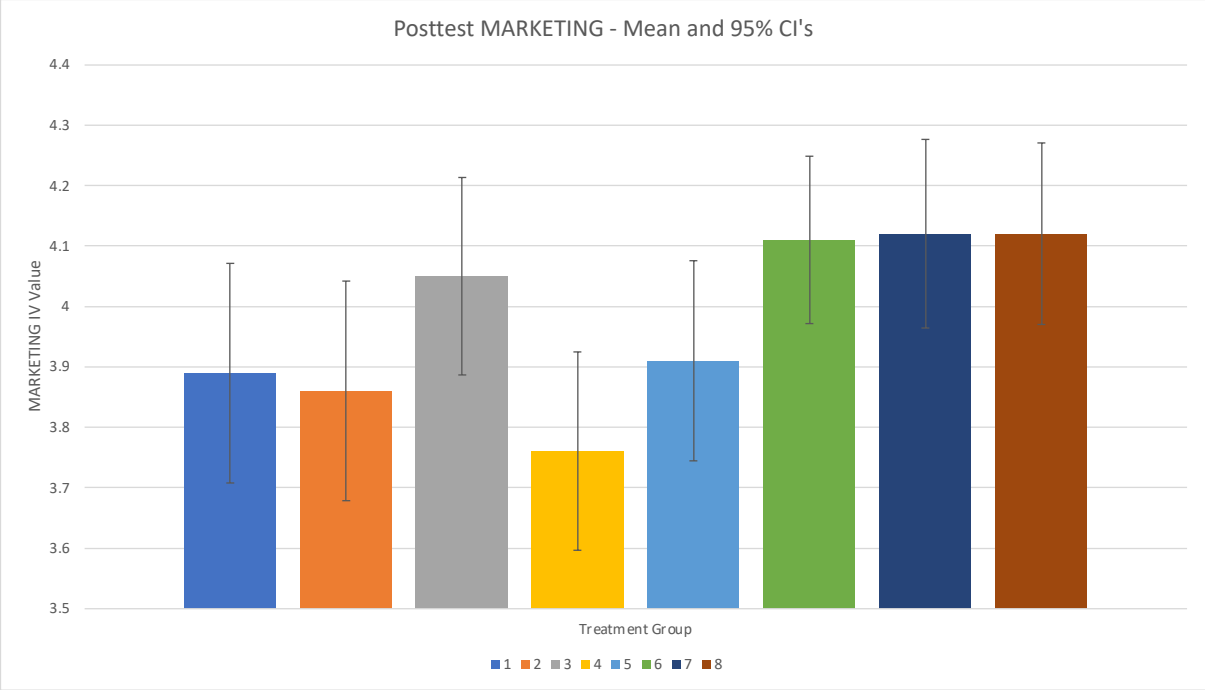
Note: Groups 2, 4, 6 & 8 received the VALUE treatment

Posttest Confidence Intervals – SERVICE IV



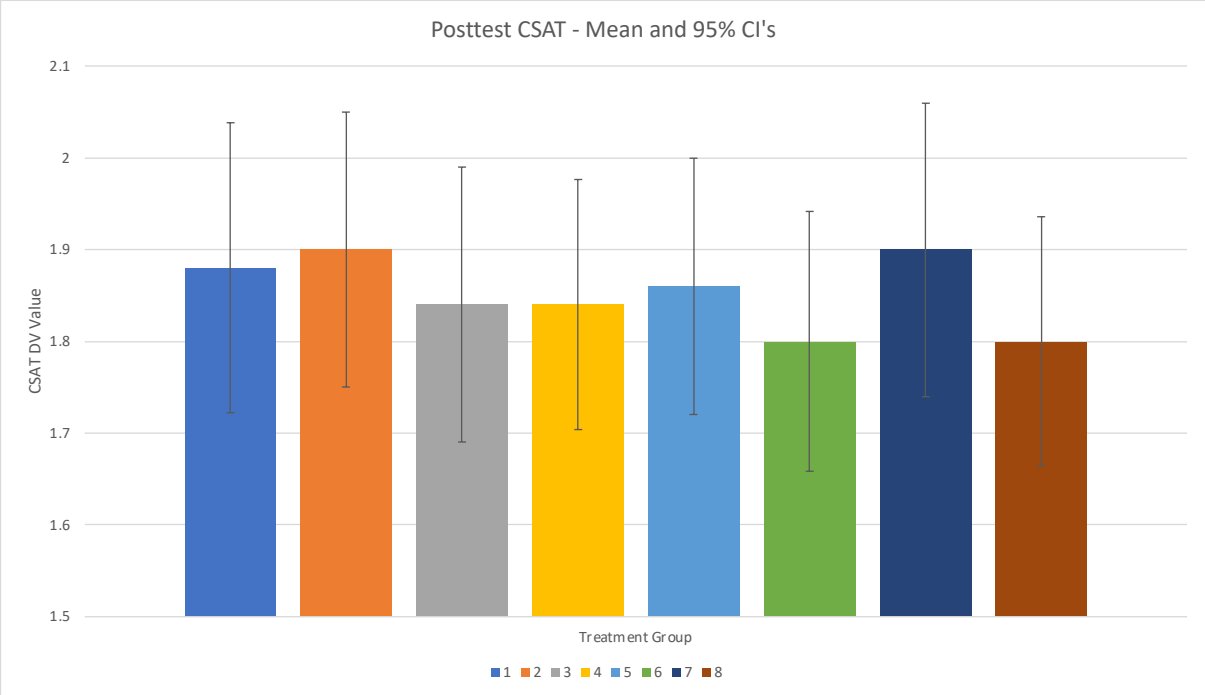
Note: Groups 3, 4, 7 & 8 received the CSERV treatment

Posttest Confidence Intervals – MARKETING IV

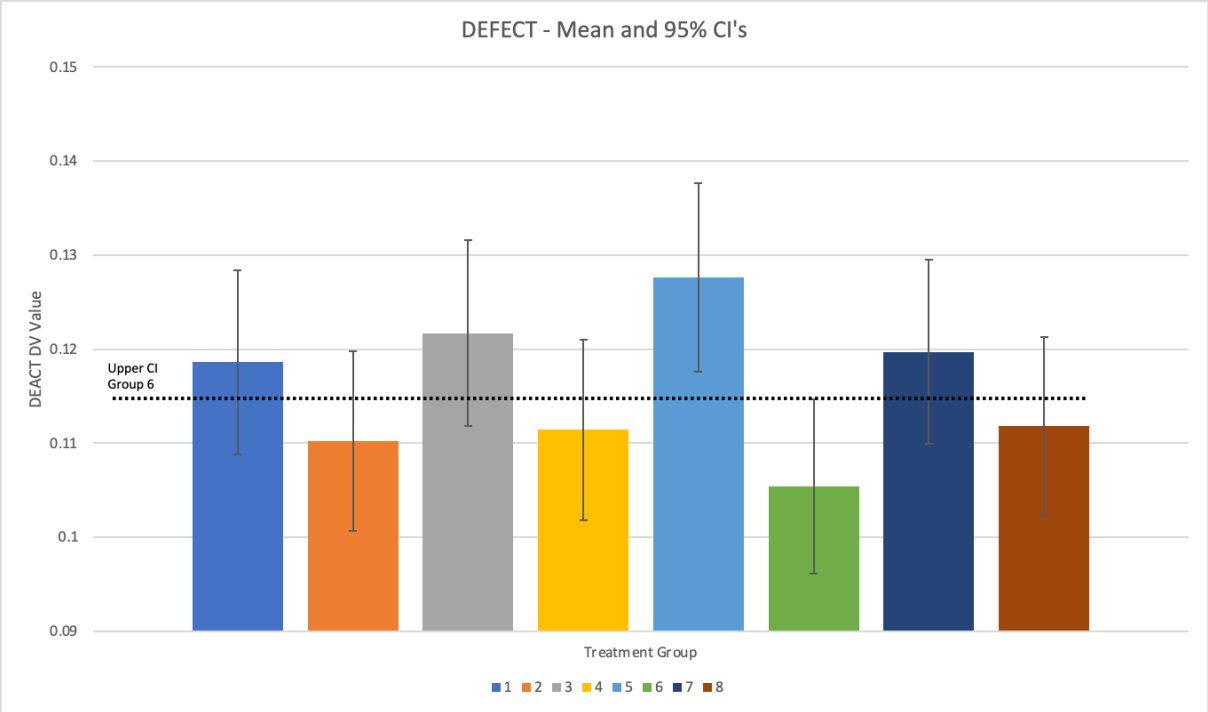


Note: Groups 5, 6, 7 & 8 received the CONTACT treatment

Posttest Confidence Intervals – CSAT DV



Posttest Confidence Intervals – DEFECT



Note: Groups 2, 4, 6 & 8 received the VALUE treatment

References

Anderson, E. W. (1994). "Cross-category variation in customer satisfaction and retention." Marketing letters **5**(1): 19-30.

Anderson, E. W., C. Fornell and D. R. Lehmann (1994). "Customer Satisfaction, Market Share, and Profitability: Findings from Sweden." Journal of Marketing **58**(3): 53-66.

Anderson, E. W. and V. Mittal (2000). "Strengthening the satisfaction-profit chain." Journal of Service research **3**(2): 107-120.

Anderson, E. W. and M. W. Sullivan (1993). "The antecedents and consequences of customer satisfaction for firms." Marketing science **12**(2): 125-143.

Anderson, R. E. and M. A. Jolson (1973). "Consumer expectations and the communications gap: Causes of consumer dissatisfaction." Business Horizons **16**(2): 11-16.

Andreassen, W. and B. Lindestad (1998). "Customer loyalty and complex services: The impact of corporate image on quality, customer satisfaction and loyalty for customers with varying degrees of service expertise." International Journal of Service Industry Management **9**(1): 7-23.

Arifin, W. N. (2012). "Random sampling and allocation using SPSS." Education in Medicine Journal **4**(1).

Aydin, S. and G. Özer (2005). "The analysis of antecedents of customer loyalty in the Turkish mobile telecommunication market." European Journal of Marketing **39**(7/8): 910-925.

Back, K.-J. and S. C. Parks (2003). "A brand loyalty model involving cognitive, affective, and conative brand loyalty and customer satisfaction." Journal of hospitality & tourism research **27**(4): 419-435.

Bae, Y. H. (2012). Three essays on the customer satisfaction-customer loyalty association. PhD (Doctor of Philosophy), University of Iowa.

Balnaves, M. and P. Caputi (2001). Introduction to Quantitative Research Methods - An Investigative Approach, SAGE Publications.

Bansal, H. S. and S. Taylor (2015). Investigating the relationship between service quality, satisfaction and switching intentions. Proceedings of the 1997 Academy of Marketing Science (AMS) Annual Conference, Springer.

Bertrand, M., E. Duflo and S. Mullainathan (2004). "How much should we trust differences-in-differences estimates?" The Quarterly journal of economics **119**(1): 249-275.

Bodet, G. (2008). "Customer satisfaction and loyalty in service: Two concepts, four constructs, several relationships." Journal of Retailing and Consumer Services **15**(3): 156-162.

Bolton, R. N. (1998). "A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction." Marketing Science **17**(1): 45-65.

Bolton, R. N. and J. H. Drew (1991). "A Longitudinal Analysis of the Impact of Service Changes on Customer Attitudes." Journal of Marketing **55**(1): 1.

Bolton, R. N. and J. H. Drew (1991). "A Multistage Model of Customers' Assessments of Service Quality and Value." Journal of Consumer Research **17**(4): 375-384.

Bolton, R. N. and K. N. Lemon (1999). "A dynamic model of customers' usage of services: Usage as an antecedent and consequence of satisfaction." Journal of marketing research: 171-186.

Brunner, T. A., M. Stöcklin and K. Opwis (2008). "Satisfaction, image and loyalty: new versus experienced customers." European Journal of Marketing **42**(9/10): 1095-1105.

Calvo-Porral, C. and J.-P. Lévy-Mangin (2015). "Switching behavior and customer satisfaction in mobile services: Analyzing virtual and traditional operators." Computers in Human Behavior **49**: 532-540.

Chen, A., Y. Lu, S. Gupta and Q. Xiaolin (2014). "Can customer satisfaction and dissatisfaction coexist?; An issue of telecommunication service in China." Journal of Information Technology **29**(3): 237-252.

Christensen, L. B., B. Johnson and L. A. Turner (2015). Research Methods, Design, and Analysis, Pearson Education Limited.

Churchill Jr, G. A. and C. Surprenant (1982). "An investigation into the determinants of customer satisfaction." Journal of marketing research: 491-504.

Cotiu, M.-A. (2013). "Satisfaction with goods and services: definitions and determinants. A critical review of the literature." Marketing From Information to Decision(6): 68-80.

Creswell, J. W. (2014). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, Sage.

Cronin, J. J. (2016). "Retrospective: a cross-sectional test of the effect and conceptualization of service value revisited." Journal of Services Marketing **30**(3): 261-265.

Cronin, J. J., M. K. Brady, R. R. Brand, R. Hightower and D. J. Shemwell (1997). "A cross-sectional test of the effect and conceptualization of service value." Journal of Services Marketing **11**(6): 375-391.

Cronin, J. J., M. K. Brady and G. T. M. Hult (2000). "Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments." Journal of Retailing **76**(2): 193-218.

Dabbolkar, P. A. (1995). "A Contingency Framework for Predicting Causality Between Customer Satisfaction and Service Quality." Advances in Consumer Research **22**(1): 101-108.

Dabholkar, P. A. and D. I. Thorpe (1994). "Does customer satisfaction predict shopper intentions." Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior **7**(1): 161-171.

Dahl, A. and J. Peltier (2015). "A Historical Review and Future Research Agenda for the Field of Consumer Satisfaction, Dissatisaction, & Complaining Behaviour." Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior **28**: 5-25.

Dean, A., M. Morris, J. Stufken and D. Bingham (2015). Handbook of Design and Analysis of Experiments, CRC Press.

Dey, B. L., W. Al-Karaghouli, S. Minov, M. M. Babu, A. Ayios, S. S. Mahammad and B. Binsardi (2020). "The Role of Speed on Customer Satisfaction and Switching Intention: A Study of the UK Mobile Telecom Market." Information Systems Management **37**(1): 2-15.

Dick, A. S. and K. Basu (1994). "Customer loyalty: toward an integrated conceptual framework." Journal of the academy of marketing science **22**(2): 99-113.

Dreze, X. and A. Bonfrer (2003). "To pester or leave alone: Lifetime value maximization through optimal communication timing."

Dubrovski, D. (2001). "The role of customer satisfaction in achieving business excellence." Total quality management **12**(7-8): 920-925.

Duflo, E. and A. Banerjee (2017). Handbook of field experiments, Elsevier.

Durrande-Moreau, A. (1999). "Waiting for service: ten years of empirical research." International Journal of Service Industry Management **10**(2): 171-194.

Easterby-Smith, M., R. Thorpe and P. R. Jackson (2015). Management & Business Research, SAGE Publications Ltd.

Eden, D. (2017). "Field experiments in organizations." Annual Review of Organizational Psychology and Organizational Behavior **4**: 91-122.

Edward, M., B. P. George and S. K. Sarkar (2010). "The Impact of Switching Costs Upon the Service Quality-Perceived Value-Customer Satisfaction-Service Loyalty Chain: A Study in the Context of Cellular Services in India." Services Marketing Quarterly **31**(2): 151-173.

El-Adly, M. I. (2019). "Modelling the relationship between hotel perceived value, customer satisfaction, and customer loyalty." Journal of Retailing and Consumer Services **50**: 322-332.

Feinberg, R. A., L. Hokama, R. Kadam and I. Kim (2002). "Operational determinants of caller satisfaction in the banking/financial services call center." International Journal of Bank Marketing **20**(4): 174-180.

Feinberg, R. A., I.-S. Kim, L. Hokama, R. Ko de and C. Keen (2000). "Operational determinants of caller satisfaction in the call center." International Journal of Service Industry Management **11**(2): 131-141.

Fisher, R. A. (1925). Statistical methods for research workers. Breakthroughs in statistics, Springer**:** 66-70.

Flint, D. J., C. P. Blocker and P. J. Boutin Jr (2011). "Customer value anticipation, customer satisfaction and loyalty: An empirical examination." Industrial Marketing Management **40**(2): 219-230.

Fornell, C. (1992). "A national customer satisfaction barometer: the Swedish experience." the Journal of Marketing: 6-21.

Fornell, C., M. D. Johnson, E. W. Anderson, C. Jaesung and B. E. Bryant (1996). "The American Customer Satisfaction Index: Nature, Purpose, and Findings." Journal of Marketing **60**(4): 7-18.

Gallarza, M. G., F. Arteaga-Moreno, G. Del Chiappa and I. Gil-Saura (2016). "Intrinsic value dimensions and the value-satisfaction-loyalty chain: a causal model for services." Journal of Services Marketing **30**(2): 165-185.

Gallarza, M. G. and I. Gil Saura (2006). "Value dimensions, perceived value, satisfaction and loyalty: an investigation of university students' travel behaviour." Tourism Management **27**(3): 437-452.

Gallarza, M. G., M. E. Ruiz-Molina and I. Gil-Saura (2016). "Stretching the value-satisfaction-loyalty chain by adding value dimensions and cognitive and affective satisfactions: A causal model for retailing." Management Decision **54**(4): 981-1003.

Giese, J. L. and J. A. Cote (2000). "Defining Consumer Satisfaction." Academy of Marketing Science Review **2000**: 1.

Grönroos, C. (1984). "A Service Quality Model and its Marketing Implications." European Journal of Marketing **18**(4): 36-44.

Grönroos, C. and P. Voima (2013). "Critical service logic: making sense of value creation and co-creation." Journal of the Academy of Marketing Science **41**(2): 133-150.

Gupta, S. and V. Zeithaml (2006). "Customer Metrics and Their Impact on Financial Performance." Marketing Science **25**(6): 718-739.

Gustafsson, A., M. D. Johnson and I. Roos (2005). "The Effects of Customer Satisfaction, Relationship Commitment Dimensions, and Triggers on Customer Retention." Journal of Marketing **69**(4): 210-218.

Ha, Y. W. and M. C. Park (2013). "Antecedents of Customer Satisfaction and Customer Loyalty for Emerging Devices in the Initial Market of Korea: An Equity Framework." Psychology & Marketing **30**(8): 676-689.

Homburg, C. and A. Giering (2001). "Personal characteristics as moderators of the relationship between customer satisfaction and loyalty—an empirical analysis." Psychology & Marketing **18**(1): 43-66.

Howell, D. C. and B. S. Everitt (2003). "Encyclopedia of statistics in behavioral science." Volume **1-4**.

Howitt, D. and D. Cramer (2017). Understanding statistics in psychology with SPSS, Pearson Higher Ed.

Hu, H.-H., J. Kandampully and T. Juwaheer (2009). "Relationships and impacts of service quality, perceived value, customer satisfaction, and image: an empirical study." Service Industries Journal **29**(2): 111-125.

Hung, H. and Y. Wong (2007). "Organisational perception of customer satisfaction: Theories and evidence." The Service Industries Journal **27**(4): 495-507.

Johnson, M. D. and C. Fornell (1991). "A framework for comparing customer satisfaction across individuals and product categories." Journal of economic psychology **12**(2): 267-286.

Jones, T. O. and W. E. Sasser (1995). "Why satisfied customers defect." Harvard business review **73**(6): 88-&.

Kenny, D. A. (1979). Correlation and Causality, John Wiley & Sons.

Kim, N., A. H. Fischer, B. Dyring-Andersen, B. Rosner and G. A. Okoye (2017). "Research techniques made simple: Choosing appropriate statistical methods for clinical research." Journal of Investigative Dermatology **137**(10): e173-e178.

Kothari, C. R. (2004). Research methodology: Methods and techniques, New Age International.

Kumar, V., I. Dalla Pozza and J. Ganesh (2013). "Revisiting the satisfaction–loyalty relationship: empirical generalizations and directions for future research." Journal of Retailing **89**(3): 246-262.

Kuo, Y.-F., C.-M. Wu and W.-J. Deng (2009). "The relationships among service quality, perceived value, customer satisfaction, and post-purchase intention in mobile value-added services." Computers in Human Behavior **25**(4): 887-896.

Lai, F., M. Griffin and B. J. Babin (2009). "How quality, value, image, and satisfaction create loyalty at a Chinese telecom." Journal of Business Research **62**(10): 980-986.

Lai, T. L. (2004). "Service quality and perceived value's impact on satisfaction, intention and usage of short message service (SMS)." Information Systems Frontiers **6**(4): 353-368.

Lim, H., R. Widdows and J. Park (2006). "M-loyalty: winning strategies for mobile carriers." Journal of Consumer Marketing **23**(4): 208-218.

Maner, J. K. (2016). "Into the wild: Field research can increase both replicability and real-world impact." Journal of Experimental Social Psychology **66**: 100-106.

Maxwell, S. E., H. D. Delaney and K. Kelley (2017). Designing experiments and analyzing data: A model comparison perspective, Routledge.

McDougall, G. H. G. and T. Levesque (2000). "Customer satisfaction with services: putting perceived value into the equation." Journal of Services Marketing **14**(5): 392-410.

Menidjel, C. and A. Bilgihan (2021). "How perceptions of relationship investment influence customer loyalty: the mediating role of perceived value and the moderating role of relationship proneness." Journal of Strategic Marketing: 1-24.

Micheaux, A. L. (2011). "Managing e-mail Advertising Frequency from the Consumer Perspective." Journal of Advertising **40**(4): 45-66.

Mittal, V. and W. A. Kamakura (2001). "Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effect of customer characteristics." Journal of marketing research **38**(1): 131-142.

Mohajeri, K., M. Mesgari and A. S. Lee (2020). "When Statistical Significance is Not Enough: Investigating Relevance, Practical Significance, and Statistical Significance." MIS Quarterly **44**(2): 525-559.

Montgomery, D. C. (2013). Design and Analysis of Experiments, John Wiley & Sons, Inc.

Oehlert, G. W. (2010). A First Course in Design and Analysis of Experiments, University of Minnesota Digital Conservancy.

Oliver, R. L. (1980). "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions." Journal of Marketing Research **17**(4): 460-469.

Oliver, R. L. (1999). "Whence Consumer Loyalty?" Journal of Marketing **63**: 33-44.

Oliver, R. L. (2010). Satisfaction: A Behavioral Perspective on the Consumer, Yee Dee Publishing Pvt Ltd.

Oliver, R. L. and W. O. Bearden (1985). "Disconfirmation processes and consumer evaluations in product usage." Journal of Business Research **13**(3): 235-246.

Oliver, R. L. and J. E. Swan (1989). "Equity and Disconfirmation Perceptions as Influences on Merchant and Product Satisfaction." Journal of Consumer Research **16**(3): 372-383.

Parasuraman, A., L. L. Berry and V. A. Zeithaml (1991). "Refinement and reassessment of the SERVQUAL scale." Journal of retailing **67**(4): 420.

Parasuraman, A., V. A. Zeithaml and L. L. Berry (1985). "A conceptual model of service quality and its implications for future research." the Journal of Marketing: 41-50.

Parasuraman, A., V. A. Zeithaml and L. L. Berry (1988). SERVQUAL - a multiple item scale for measuring customer perceptions of service quality, Journal of Retailing. **64:** 28.

Patterson, P. G. and R. A. Spreng (1997). "Modelling the relationship between perceived value, satisfaction and repurchase intentions in a business-to-business, services context: an empirical examination." International Journal of Service Industry Management **8**(5): 414-434.

Peeters, M. J. (2016). "Practical significance: Moving beyond statistical significance." Currents in Pharmacy Teaching and Learning **8**(1): 83-89.

Roberts-Lombard, M. and D. J. Petzer (2018). "Customer satisfaction/delight and behavioural intentions of cell phone network customers – an emerging market perspective." European Business Review(4): 427.

Rust, R. T., K. N. Lemon and V. A. Zeithaml (2004). "Return on Marketing: Using Customer Equity to Focus Marketing Strategy." Journal of Marketing **68**(1): 109-127.

Rust, R. T. and A. J. Zahorik (1993). "Customer satisfaction, customer retention, and market share." Journal of retailing **69**(2): 193-215.

Salkind, N. J. (2010). Encyclopedia of research design, Sage.

Sekaran, U. and R. Bougie (2016). Research methods for business: A skill building approach, John Wiley & Sons.

Shadish, W. R., T. D. Cook and D. T. Campbell (2002). Experimental and Quasi-Experimental Designs for Generalized Causal Inference, Houghton Mifflin Company.

Solomon, P., M. M. Cavanaugh and J. Draine (2009). Randomized controlled trials: Design and implementation for community-based psychosocial interventions, Oxford University Press.

Splawa-Neyman, J., D. M. Dabrowska and T. Speed (1923). "On the application of probability theory to agricultural experiments. Essay on principles. Section 9." Statistical Science: 465-472.

Spurlock, D. (2019). "Defining Practical Significance Is Hard, But We Should Do It Anyway." Journal of Nursing Education **58**(11): 623-626.

Tam, J. L. M. (2004). "Customer Satisfaction, Service Quality and Perceived Value: An Integrative Model." Journal of Marketing Management **20**(7/8): 897-917.

Thusyanthy, V. and V. Tharanikaran (2017). "Antecedents and Outcomes of Customer Satisfaction: A Comprehensive Review." International Journal of Business and Management **12**: 144.

Tunnell, G. B. (1977). "Three dimensions of naturalness: An expanded definition of field research." Psychological Bulletin **84**(3): 426.

van Diepen, M., B. Donkers and P. H. Franses (2009). "Does irritation induced by charitable direct mailings reduce donations?" International Journal of Research in Marketing **26**(3): 180-188.

Verhoef, P. C. (2003). "Understanding the Effect of Customer Relationship Management Efforts on Customer Retention and Customer Share Development." Journal of Marketing **67**(4): 30-45.

Wang, Y., H.-P. Lo and Y. Yang (2004). "An Integrated Framework for Service Quality, Customer Value, Satisfaction: Evidence from China's Telecommunication Industry." Information Systems Frontiers **6**(4): 325-340.

Whiting, A. and N. Donthu (2006). "Managing Voice-to-Voice Encounters: Reducing the Agony of Being Put on Hold." Journal of Service Research **8**(3): 234-244.

Whiting, A. and N. Donthu (2009). "Closing the gap between perceived and actual waiting times in a call center: results from a field study." Journal of Services Marketing **23**(5): 279-288.

Wing, C., K. Simon and R. A. Bello-Gomez (2018). "Designing difference in difference studies: best practices for public health policy research." Annual review of public health **39**.

Wu, H.-C. (2014). "The effects of customer satisfaction, perceived value, corporate image and service quality on behavioral intentions in gaming establishments." Asia Pacific Journal of Marketing and Logistics **26**(4): 540-565.

Yang, Z. and R. T. Peterson (2004). "Customer Perceived Value, Satisfaction, and Loyalty: The Role of Switching Costs." Psychology & Marketing **21**(10): 799-822.

Yoon, Y. and M. Uysal (2005). "An examination of the effects of motivation and satisfaction on destination loyalty: a structural model." Tourism management **26**(1): 45-56.

Zeithaml, V. A. (1988). "Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence." Journal of Marketing **52**(3): 2-22.

Zeithaml, V. A., L. L. Berry and A. Parasuraman (1993). "The Nature and Determinants of Customer Expectations of Service." Journal of the Academy of Marketing Science **21**(1): 1-12.

Zeithaml, V. A., L. L. Berry and A. Parasuraman (1996). "The behavioral consequences of service quality." the Journal of Marketing: 31-46.

Zeithaml, V. A., M. J. Bitner and D. D. Gremler (2018). Services marketing: Integrating customer focus across the firm, McGraw-Hill Education.