CUSTOMER SATISFACTION MEASUREMENT

Strategies, methodologies and factors influencing customer satisfaction measures

Universitat Oberta de Catalunya — Estudis d'Economia i Empresa

Grau de Màrqueting i Investigació de Mercats

Pau Giró Manzano
pgiro@uoc.edu
June 2021

Treball de final de grau
Tutor: Santi Ariste
Àmbit d'especialització: Investigació de mercats
Memòria final. Curs 2020-2021

"We have a million and one metrics to measure someone's performance, and negligible to no metrics to measure someone's trustworthiness. And so what we end up doing is promoting or bonusing toxicity in our business, which is bad for the long game."

Simon Sinek

"(the big data) Really should offer us proposals to make better decisions and patterns with non-obvious relationships between data in order to innovate. Having more data does not mean increase the purpose nor guarantee the impact of anything. Overfeed the cloud with massive data with no use does not mean innovation. The key is combine data with purpose, any other thing is living-room bull fighting."

Xavier Marcet

ABSTRACT

ENGLISH

Customer satisfaction measurement is currently within the most important metrics analysed by the business and used also to support the brand's positioning. However, measuring the actual customer's satisfaction is much more complex than it may look. First because its two parameters, customer expectations and actual performance, can evolve independently. But also because of the different parameters that can influence not just the scoring, but the fact of scoring.

From the analysis performed it's been possible to identify that in general terms customer tend not to answer the satisfaction surveys. However, this response rate fluctuates considering their satisfaction: there is a strong negative correlation between response rate when satisfied and dissatisfied. On top of that, customers that tend to answer with higher frequency when they are dissatisfied also tend to give lower scores on the scale.

There are abundant documented malpractices of employees or companies "hacking" the satisfaction score to get better results and have higher bonuses or priority access to inventory. But from the conducted analysis it's possible to establish that a soft guide (a colour scale) does not affect the satisfaction scores and seems to eliminate errors interpreting the scale wrong (giving low scores to high satisfaction and vice versa).

The customer context will also affect naturally the score given. From the evaluated parameters, when the customers are either satisfied or dissatisfied there's no particular variable influencing the results. However when the customer is neutral (not surprised positive or negatively), as per definition at high risk of churn for alternative solutions, there are multiple factors that influence on the score. One of the most relevant factors is price (the higher the importance the lower the score), but very interesting is that the respondents that work or worked recently on jobs connected to satisfaction surveys tend to give higher scores.

Keywords: Customer satisfaction, CSat, NPS, Satisfaction surveys

CATALÀ

La mesura de la satisfacció del client és actualment entre les mètriques més importants que es monitoritzen a nivell empresarial i utilitzada, també, per posicionament de marca. Tot i això, mesurar la satisfacció real dels clients és molt més complex del que podria semblar. Primer perquè els paràmetres de la satisfacció, les expectatives del client i l'ús del producte, poden variar de forma independent. Però també arran dels diferents paràmetres que poden influir no només en la puntuació sinó que en el fet de puntuar.

De l'anàlisi dut a terme ha estat possible identificar que en termes generals els clients acostumen a no contestar enquestes de satisfacció. Tot i això, aquesta taxa de resposta varia amb la seva satisfacció: hi ha una forta correlació negativa entre la taxa de resposta quan està satisfet i quan està insatisfet. A més a més, els clients que tendeixen a contestar més quan estan insatisfets també acostumen a puntuar amb notes més baixes.

Hi ha abundants casos documentats de treballadors o empreses alterant la puntuació de satisfacció per tal d'aconseguir millors resultats o accés prioritari a l'inventari. Però a partir de l'anàlisi dut a terme s'ha pogut veure que una guia subtil (una escala de color) no afecta la puntuació de la satisfacció i a més sembla eliminar errors interpretant l'escala (assignar baixes puntuacions a satisfaccions elevades i viceversa).

El context del client també afectarà de forma natural a les puntuacions negatives. Dels paràmetres analitzats quan el client està o bé satisfet o be insatisfet no hi ha cap variable que influeixi als resultats. Però quan el client és neutral (no sorprès negativament o positiva), que per definició estarà a risc de canviar a solucions alternatives, hi ha diversos factors que poden influir-hi. Un dels més rellevants és el preu (com més alta sigui la importància del preu, més baixa serà la puntuació), però un de molt interessant és el fet que quan l'enquestat ha treballat o treballa actualment en una feina relacionada amb enquestes de satisfacció tendeix a donar puntuacions més altes.

Paraules clau: Satisfacció del client, CSat, NPS, enquestes de satisfacció

INTRODUCTION

MOTIVE BEHIND THE RESEARCH

After working almost my entire life in front of the customer or managing customer care teams, more than 10 years in the contact centre industry and more than 8 managing teams across multiple departments (and being a customer), I have been able to see several methodologies, question formats as well as ways to skew my answer to the survey.

Knowing how the information is handled from the company side I have seen also many times people using the metric as an absolute truth without considering that it could be skewed by cultural or socio-demographic differences, errors, unclear user interfaces, fluctuations on product or issue mix as well as wrong usage of the results (accepting as valid extremely low samples or response rates, activating costly contingency plans after fluctuations that are statistically non-significant…).

As a customer I have also seen surprising behaviours, where I have had to choose between giving the score I though represented my satisfaction or choosing the score I though was fairer considering the impact it could have on the employee that served me.

That made me think if too many times the metric is more important as a metric rather than as an approximated representation of the customer's opinion and, therefore, with all the processes that can alter it, if at the end companies are, knowingly or not, masking key information to better understand their customers and their actual performance.

RESEARCH GOALS

1. Identify if there is any impact on the satisfaction scores with a soft guide on the survey (colour scale).
2. Identify if there are differences on response rate depending on the satisfaction or any other factor.
3. Identify if customer satisfaction (previous satisfaction) has an influence on future purchase decision making.
4. Identify if the different decision making factors generate different satisfaction scoring
5. Identify if the mindset towards scoring after a human interaction is the same than when scoring an automatic interaction (expectations and response rate)
6. Identify what would be the approach to monitor customer satisfaction on 3 scenarios: start-up, and mature companies with already a monitoring system and mature companies not monitoring customer satisfaction.

SCOPE AND METHODOLOGY OF THE RESEARCH

Even though initially the goal was to do the research across several countries in Western Europe, due to time and resource limitations it was reduced to Spanish market.

The qualitative analysis has been done through in-depth interviews to some members of the company or customer experience (CX) leadership team with focus on their customer satisfaction data gathering strategy as well as how is the data used internally. The companies interviewed allowed to understand the approach for 2 big companies with a mature customer satisfaction process; 2 start-ups, one with an initial customer satisfaction process in place and the other not in place yet; and a third company, a big company with many decades being leader in its market, that does not have a customer satisfaction monitoring process in place.

The quantitative analysis was conducted via online surveys being distributed through social networks (WhatsApp, LinkedIn and Twitter). The first section of the questionnaire ensured that the respondents currently lived (or until recently) in Spain and older than 18 years.

As one of the aims was to compare two different types of satisfaction surveys (as will be explained later in the document), there were two versions of the questionnaire, both created using Microsoft Forms. The respondents were randomly assigned to one or another version using *allocate.monster*, an online free resource that allows this type of survey testing.

The survey generated 181 answers from which 90 were on one type of survey and 91 on the other. The data originally was spread across 71 columns, from which only 23 were valid, the others were automatically generated by the form and were applicable only in case the questions were intended to be a quiz or exam (calculated the score). The dataset used for the final analysis had a total of 79 variables excluding the open text comments from the respondents, which were not analysed from a quantitative standpoint although will be used alongside the document to exemplify some customer behaviours.

CUSTOMER SATISFACTION ORIGINS

Customer or consumer satisfaction has not been always known as it is now. Initially, and for decades, it has been an area within the human behaviour field, more connected to philosophers. In some cases, it was understood as a part of human life, in other it was understood as the goal of human needs.

Later in time, and with the birth of modern social science during the 1940's, it started to gain relevance and economists started to use customer satisfaction as an important factor on supply and demand characteristics. By the same time the pioneers of the modern marketing started to measure it too.

Consumer behaviour as an independent field of study is considered to happen during the 1950's gaining importance during the next decade, when the classic consumer behaviour models recognized customer satisfaction as one of the factors influencing the purchase decision making.

In the late 1960's Edwin Locke theorized about defining customer (and employee) satisfaction using the combination of expectations, performance and importance of the different attributes.

A bit later, Ralph L. Day and other researchers wanted to understand better customer satisfaction and the related post-choice constructs. Having an initial definition in their first published paper on this field by Day in 1977, the first proper definition of customer satisfaction arrived in 1984, also by Day, leaving behind the attribute importance introduced by Locke:

> *"It is almost universally accepted in the literature that consumer satisfaction/dissatisfaction is the consumer's response in a particular consumption experience to the evaluation of the perceived discrepancy between prior expectations (or some other norm of performance) and the actual performance of the product as perceived after its acquisition."*
>
> *(DAY, 1984)*

The definition itself has been slightly reviewed a few times afterwards. One of the most recent updates, done by Uwe Peter Kanning and Nina Bergman published in 2009, confirmed Day's definition although they identified that performance itself is much more reliable predictor of satisfaction than the initial expectations and/or the attribute importance.

THE IMPORTANCE OF CUSTOMER SATISFACTION

CUSTOMER SATISFACTION AS A MEASURE OF VALUE PROVIDED

Customer satisfaction, even though it does not provide any direct revenue nor have any direct impact on the companies' market valuation, it has been gaining importance on the leadership teams of the business as an indicator of the value offered that their customers perceive.

However, one of the dangers of the customer satisfaction measurements, as will be analysed in depth later in the text, is putting the focus on the company rather than on the customer. In fact, in 2000 Reichheld, Markey and Hopton published a paper highlighting this discrepancy between their thinking and the standard understanding:

> *"Profits are important, not just as an end in themselves, but because they allow the company to improve value and provide incentives for employees, customers and investors to remain loyal to the company."*
>
> *(REICHHELD, MARKEY, & HOPTON, 2000)*

This current research tried identify whether quality and previous satisfaction influenced on future purchase decision making compared with other factors such as price, convenience, comments… After analysing the results (details regarding the quantitative analysis can be found later in the document), we could identify that:

1. Product / service / brand quality is the most important factor: 45% of the surveys placed quality as the top factor and 73% placed quality within the top 3 factors. According to Day's and Kanning & Bergman's researches, one of the factors that defines the customer satisfaction is placed on the top of the importance ranking.
2. Previous satisfaction is the 2nd most important factor: placed on top on 16% and on the top 3 in 49% of the surveys, the previous satisfaction, which is influenced by the top 1 factor as mentioned, is also a big influence on the future purchase decision making process.
3. Price: top factor on 12% of the surveys and put on the top 3 on 49% of them. One of the most obvious factors for purchase decision making (at the end of the day, when you have limited resources you will look for the best purchase available at the price you can pay). Using the definition of Anderson and Narus from their paper on 1998: "*The difference between value and price equals the customer's incentive to purchase."*(ANDERSON & NARUS, 1998). This factor is key in relationship with quality, as it will consciously limit the quality choosing some characteristics of the product over the other given the expected outcome.

All the other factors were significantly behind (more than 10 points in the top 3 analysis) in importance. This analysis confirms the importance of customer satisfaction as a measure to ensure a healthy growth, given the fact that 2 out of the top factors are what define your expectations on the purchase and the third one is the satisfaction in the last purchase.

RETENTION & PROFIT: LOYALTY & CUSTOMER LIFETIME VALUE

As the current one does, previous research that customer satisfaction should have a direct influence on future purchases except for purely convenience-driven decisions (for instance, fuel stations in the highway) or for commodities (such as toilet paper or batteries for devices), where the factor ranking will probably be different (but where previous satisfaction will have an impact).

In fact, there is abundant research on the field trying to connect customer retention and profit. By the end of the 1980's the concept of Customer Lifetime Value (CLV or CLTV) started to appear. It aims to measure or predict the net profit (difference between the revenue generated by the customer throughout its relationship with the company and costs associated with it, such as advertising, cost of sale, costs of serving the customer…) contributed by the customer considering its past and expected future relationship. It is important not to mistake with customer profit, which measures only past profit but without projecting into the future.

In 2012 a team from the Beijing University in collaboration with the University of Washington analysed the relationship between customer satisfaction, CLV, customer loyalty, revenue and relationship duration in the mobile data services sector across both US and China. Their data identified that there is a connection between customer satisfaction and customer loyalty (fig. 1). They also identified that customer revenue and relationship duration were both driven by customer loyalty but not by customer satisfaction.

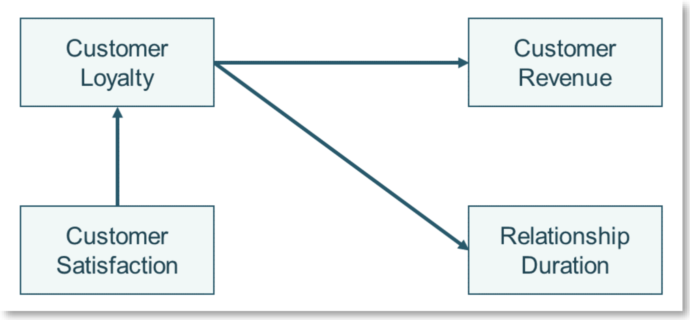

Fig. 1: Conceptual map of the relationships identified.
Source: "Are customer satisfaction and customer loyalty drivers of customer lifetime value in mobile data services: a comparative cross-country study" (QI, ZHOU, CHEN, & QU, 2012)

Although a few years later Bader Almohaimmeed, a researcher from the Qassim University (Saudi Arabia), published another paper analysing the relationship between CRM actions, customer satisfaction, loyalty, revenue and relationship duration. The results, published in 2019, reported that customer satisfaction influenced both loyalty and profitability, finding no mediation of loyalty between satisfaction and profitability, although no direct connection between satisfaction and retention.

The concept of Customer Loyalty was introduced by Reichheld, Markey and Hopton in 2000. Their article "The loyalty effect – The relationship between loyalty and profits" connected loyalty with the creation of value for the customer as it "*reliable measures whether superior value has been delivered*" (REICHHELD, MARKEY, & HOPTON, 2000) and therefore, the customers will naturally come for more. But they not only identified this connection, the document also specified other economic impacts that affect the entire business system:

> *"1. Revenues and market share grow as the best customers are swept into the company's book of business, building repeat sales and referrals.*
>
> *2. Costs shrink as the expense of acquiring and serving new customers and replacing old ones declines.*
>
> *3. Employee retention increases because job pride and job satisfaction increase, in turn creating a loop that reinforces customer retention through familiarity and better service to the customers. Increased productivity results from increasing employee tenure."*
>
> *(REICHHELD, MARKEY, & HOPTON, 2000)*

IMPORTANCE OF RETENTION

All this research on the relationship between customer satisfaction as indicator of CLV and customer retention is triggered by the initial work of Reichheld and Sasser in the early 1990's that was able to quantify that a 5% reduction on defection rate (churn) generated between 30% to 85% more benefit depending on the industry (their research analysed banking, insurance and self-service industries) (REICHHELD & SASSER, 1990).

This first research generated further research by Heskett, Sasser and Schlesinger who quantified the cost of acquiring a new customer as 5 times the cost of retaining an existing customer (HESKETT, SASSER, & SCHLESINGER, 1997). Not only that, the work by Reichheld, Markey and Hopton identified that "5 percentage points shift in customer retention consistently resulted in 25-100% profit swings" (REICHHELD, MARKEY, & HOPTON, 2000). Although this statement has been considered a myth more recently by some authors (DUNN, 2003) considering that it is an over-generalization, that the research was done almost 30 years ago and that both customer acquisition and product/service consumption has changed during the last decades with globalization, online services and the subscription economy. However, this ratio is still a generally accepted value.

The initial article published by Reichheld and Sasser in 1990 was able to identify not only direct profit from the original purchase but also related profit generated by parallel sources all coming from the original purchase (fig. 2). The fundamental concept of their article is that maximizing retention will maximize the return of the investment.

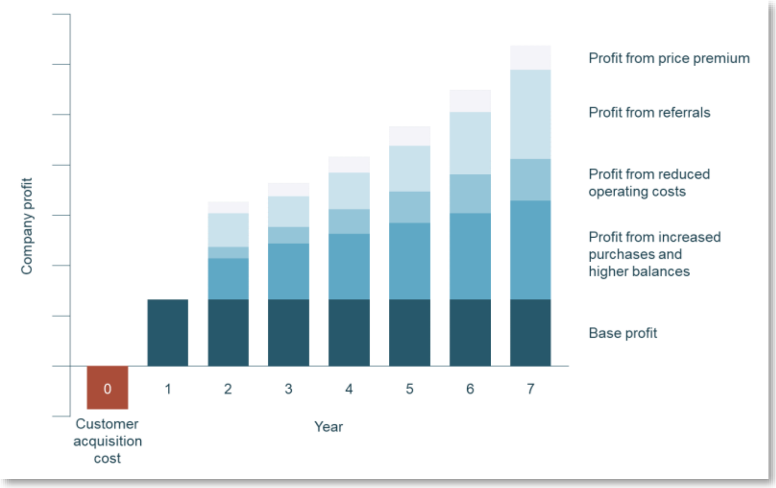

Fig. 2: Example of annual growth generated by a retained customer with the different profit streams.
Source: "Zero defections: Quality comes to service" (REICHHELD & SASSER, 1990)

This chart included what would be the foundation of one of the most used satisfaction metrics across all industries: Net Promoter Score. On his article in Harvard Business Review published in 2004, Reichheld described the importance of the metric as follows:

> *"Companies spend lots of time and money on complex tools to assess customer satisfaction. But they're measuring the wrong thing. The best predictor of top-line growth can usually be captured in a single survey question: Would you recommend this company to a friend?"*
>
> *(REICHHELD F. F., 2004)*

SATISFACTION AS A COST SAVER

Even though customer satisfaction has been identified as a driver of revenue, recent work conducted by Lim, Tuli and Grewal was able to proof that customer satisfaction has statistically and economically significant negative effect on future cost of sale (COS), which means: the higher the customer satisfaction, the lower the COS (LIM, TULI, & GREWAL, 2020).

Specifically, they identified that a 1 point increase of the American Customer Satisfaction Index (ACSI, explained in the next section) corresponded to a decrease of nearly 3% of the average COS in their sample (in revenue, they estimated it in $130 million in future COS). It is also mentioned that the effect is stronger on the cost of persuasion (compromising commissions and marketing and advertising expenses) and not that much on the cost of convenience (compromising freight-out and bad debt expenses). This work will help putting customer satisfaction as an important metric because a reduction on the COS is generally viewed as a positive signal in financial markets.

COMPLAINTS AND CRITICAL INCIDENTS

The most common customer experience throughout the customer journey is where the product works as intended (although there may be a mismatch of expectations and actual performance). However, on some cases the product, service or points of the journey do not work as intended. In 1984 Day published his work "Modelling choices among alternative responses to dissatisfaction" and with the chart in figure 3 summarised the different factors that drive the customer to complain.

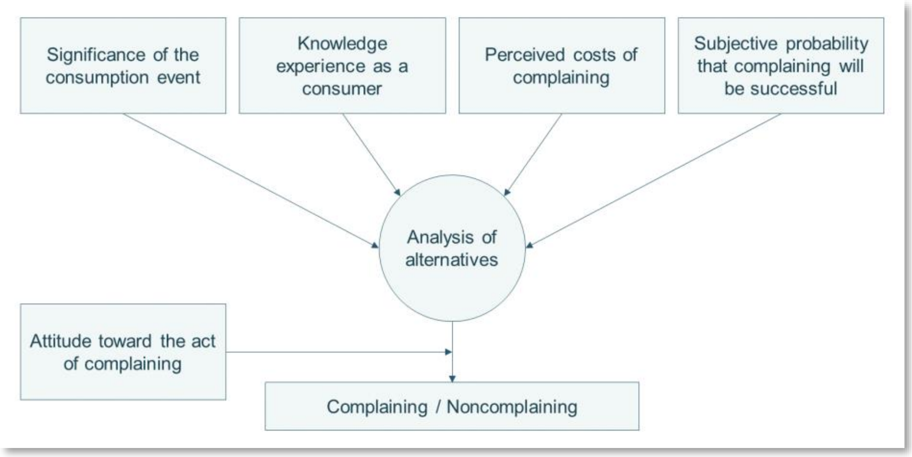

Fig. 3: Conceptual map of the process for complaining or not.
Source "Modelling choices among alternative responses to dissatisfaction": (DAY, 1984)

In the chart we can identify that 4 factors (significance of consumption event, knowledge experience as a customer, perceived costs of complaining and subjective probability that complaining will be successful) are considered by the customer. That, together with the attitude towards the act of complaining, will be the triggers of the complaint.

A work by Van Doorn and Verhoef published in 2008 evaluated how do incidents affect loyalty through customer satisfaction focusing, particularly, their work was focused on critical incidents, as they considered, supported on existing literature, that links between satisfaction (transactional satisfaction) and loyalty (a proxy for relational satisfaction) "*tend to be characterized by inertia that causes parties to conduct 'business as usual' and, in essence, maintain the status quo*" (VAN DOORN & VERHOEF, 2008). However, they considered that critical incidents could act as "triggers" that could destabilize and reconsider the relationship.

In fact, their work identified that the relationship outcome after a critical incident was non-linear, having different outcomes depending on the relationship quality. Their work suggested that for a satisfied customer with a high customer share, a critical incident has the potential to intensify the relationship. In the quadrant reproduced in figure 4 they outlined four different strategies depending on customer share and previous customer satisfaction.

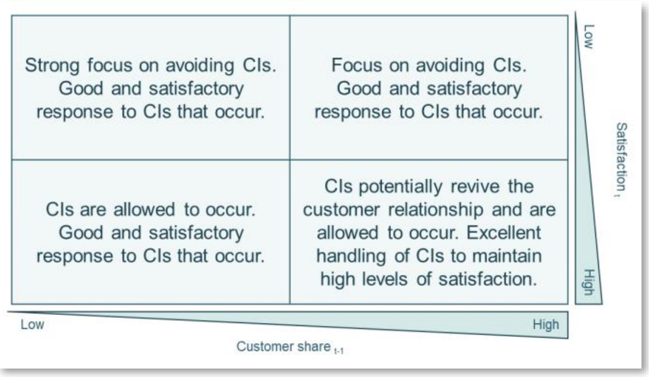

Fig. 4: Different strategies to implement depending on the customer's satisfaction and customer share.
Source: "Critical incidents and the impact of satisfaction on customer share" (VAN DOORN & VERHOEF, 2008)

From the interviewed companies, all of them reported to have processes to handle this type of incidents, no matter the company size, scope, business or even their customer experience area maturity. It was mentioned by all of them as can see in the following quotes:

> *"When there's an incidence we are extremely reactive (we react almost with a drop-everything strategy), we analyse the category (reason for the dissatisfaction) and we make the appropriate decisions whatever the cost and with very little questioning to the customer."*
>
> *B2B Company with no customer satisfaction process*
>
> *"The contact centre has an SLA of 24h to contact back all customers […]. Each 24h that it is not actioned, it goes up to the next level and as it is a big company but with a relatively flat structure, the 3rd or 4th day the alert can reach the Country Manager."*
>
> *B2B Company with a mature customer satisfaction process*
>
> *"We focus on showing the customer our intention to be there for them, which also makes easier the problem solving as the perception is that we do everything we can. We do a first call 3 days after they received the product."*
>
> *B2C Start-up at early stages of customer satisfaction process*

INDICATOR OF CHURN

Therefore, and highlighted the importance of retention and some of the factors that may influence it such as satisfaction or critical incidents, let's analyse churn, one of the most important metrics to monitor specially in subscription-based business models based on generating recurring revenue streams from customers instead of standard transactional models. In the recent years, with the standardization of broad band connections to the Internet and low cost data plans in the telecommunications industry, there has been a surge in companies whose business model is based on subscription.

For the subscription-based business model companies, retention is the default state: is what happens if the customer does not proactively request to cancel the service. Therefore, the measurement is based on the % of cancelled services over the total services active. This is the churn index.

This metric has been identified in this research as one of the most important metrics at the early stages of a start-up. In the in-depth interviews conducted during this present research to companies on this stage, it was mentioned in both that at the very first stages, they used churn as one of their product and customer experience indicator.

In one of the interviewed companies, monitoring churn at very early stages helped them identify opportunities of improvement not identified during the product research and validation process.

> *"The churn reduced a lot when we implemented the customer success, as it allowed us to identify some silly errors that had an easy fix."*
>
> *B2C Start-up with very early customer satisfaction process implemented*

In the other, they have been using churn to assess the growth quality: if during a significant sales-driven growth process churn has been relatively flat, it means that the sale process has been consistent even with the growth and that the product does not loose performance during the scale-up.

> *"The reason for the low focus on customer satisfaction monitoring is because the churn has remained stable and under control throughout the entire period and has been compensated with up-sell, cross-sell and share of wallet growth."*
>
> *B2B Start-up with no customer satisfaction monitoring process implemented*

MEASURING CUSTOMER SATISFACTION

GATHERING THE DATA

The standard approach to customer satisfaction measures is through a scale rating. This scale can be from 0 to 10, which is the most common not only because is the one used by the Net Promoter Score but also because it aligns with the standard scoring of exams and assessments, but also can be scales from 1 to 5, 1 to 7, thumbs up / thumbs down, traffic lights (red, amber, green)…

Below are a 4 examples of methodologies (the most common), but it does not mean that they are the only metrics: a quick search on Google will easily provide a dozen different indicators to evaluate the customer satisfaction.

AVERAGE SCORE

One of the most intuitive measurements is just calculating the average of the different scores given by the customers.

$$Average\ customer\ satisfaction = \frac{Sum\ of\ satisfaction\ scores}{Count\ of\ surveys\ received}$$

The metric can be calculated with any of the previous scales, although it will convert discrete scorings into continuous scorings. It can be found usually on online platforms like Google, Amazon or Airbnb (fig. 5).

Fig. 5: Different examples of average scored rates (on a 5 stars scale).
Source: screenshots from maps.google.com, amazon.com and airbnb.com websites

CSAT SCORE

The CSat score is an evolution from the previous metric. The CSat score is based on the proportion of customers with valid satisfaction over the total amount of surveys received. Therefore, requires assigning the threshold to define what is “valid satisfaction”. In reality, this metric will transform the scores into a binary score: if the result is above a given threshold it will count as satisfied (as a 1 for the effects of calculation) and, if it is below, it will count as dissatisfied (as a 0 for the effects of calculation).

The metric will be presented as a % of satisfied customers.

$$CSat\ score = \frac{Count\ of\ surveys\ scored\ above\ x}{Count\ of\ surveys\ received}$$

As with the previous, it can be used with all mentioned scales, although in binary scales (thumbs up / thumbs down) will have no practical difference with the average score.

NET PROMOTER SCORE

In 2003, Reichheld published an article in Harvard Business Review that meant the inception of a new metric: the Net Promoter Score (NPS). The article which title, "*The one number you need to grow*", was already a statement was based on the following analysis outcome:

> *"Companies spend lots of time and money on complex tools to assess customer satisfaction. But they're measuring the wrong thing. The best predictor of top-line growth can usually be captured in a single survey question: Would you recommend this company to a friend?"*
>
> *(REICHHELD F. F., 2004)*

After 2 years of research and analysis testing multiple survey questions and connecting the responses with actual customer behaviour (purchasing patterns and referrals) as well as company growth, what he found was that the most effective question in regards of correlation with growth rates among competitors question wasn't about customer satisfaction or even loyalty but the percentage of customers willing to refer the product/service to a friend or colleague.

It is very common to find surveys that include at the end a question that will look like the one in figure 6.

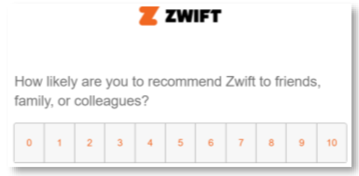

Fig. 6: Example of real survey sent by Zwift via email.
Source: example of survey received from Zwift (indoor e-cycling platform)

From a logic standpoint, the metric considers all customers who rated with a 9 or a 10 as satisfied and all the customers that rated with a 6 or lower score as dissatisfied. The customers that rated with a 7 or an 8 are considered as "neutral" or "passive". The calculation is as below and will be presented as a value that goes from 100 (all customers are satisfied) to -100 (all customers are dissatisfied):

$$NPS\ score = 100 \times \frac{Satisfied\ customers\ (promoters) - Dissatisfied\ customers\ (detractors)}{Count\ of\ surveys\ received}$$

This metric, though, presents an important weakness: it is based in the principle that the satisfied customers will speak favourably and the dissatisfied will speak unfavourably about the company. However, even though the metric overlooks what are defined as neutral customers, in reality they will have relatively low effect on the word-of-mouth but are customers that will easily change of provider as they are "not enough satisfied". At the end of the day, the passive or neutral customers are not loyal customers.

ACSI

During the first half of the 1990's a group of experts (Fornell, Johnson, Anderson, Cha and Bryant) from the United States worked on a customer satisfaction index (focused exclusively in the US market) that was presented to the world in 1996 with an article in the Journal of Marketing. This metric is significantly more complex than any other satisfaction metric in this research as it is based on an econometrics model but their research allowed them to create a conceptual map of customer satisfaction factors and relationship between them (fig. 7).

The measure is based on a nation-wide survey that collects data across multiple industries and an econometric approach to estimate the indices. The metric and the methodology has been active since the publication of the article and offers a consistent and standardized analysis that can be used not only to support decision making and help modelling the factors that end up giving a good understanding of customer loyalty following the below model, but also is used as a source for further research like the already mentioned research by Lim, Tuli and Grewal on the cost of sale impact of customer satisfaction.

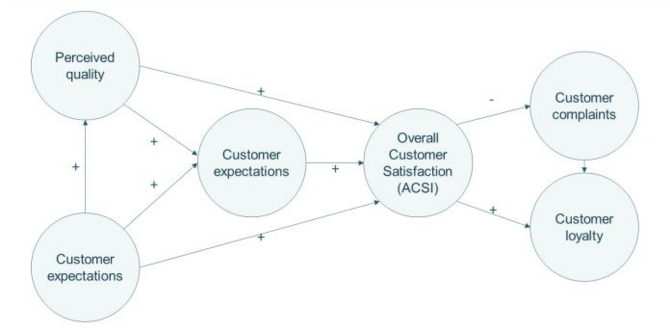

Fig. 7: Factors of satisfaction and relationship between them.
Source: "The American customer satisfaction index: nature, purpose and findings", (FORNELL, JOHNSON, ANDERSON, CHA, & BRYANT, 1996)

CUSTOMER EFFORT SCORE

A few years after the introduction of the NPS score the same journal published a new suggested customer satisfaction metric: the Customer Effort Score, designed by Dixon, Freeman and Toman. The article states that usually companies focus on delighting their customers by exceeding service expectations. However "*a large-scale study of contact-centre and self-service interactions finds out that what customers really want (but rarely get) is just a satisfactory solution to their service issue*" (DIXON, FREEMAN, & TOMAN, 2010).

Their research proved that the CES approach to customer satisfaction, based on the question "How much effort did you personally have to put forth to handle your request?" scored from 1 (very low effort) to 5 (very high effort), proved to have a very strong predictive power for both repurchasing intention and increased spending, outperforming CSat and NPS (fig. 8).

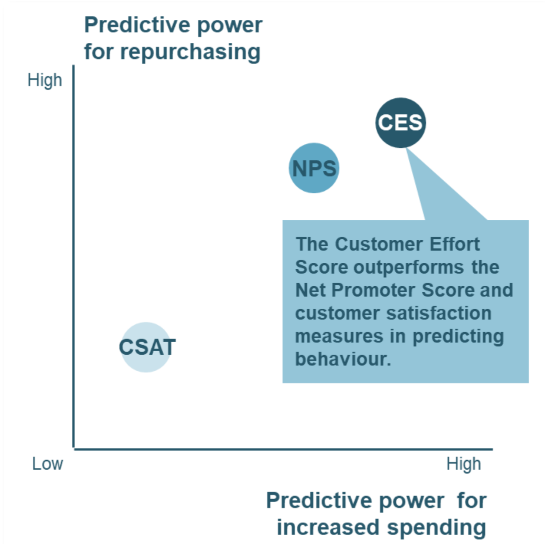

Fig. 8: Comparison of predictive power between CSat, NPS and CES according to the CES designers.
Source: "Stop trying to delight your customers" (DIXON, FREEMAN, & TOMAN, 2010)

One key difference of this metric towards the other is that instead of being company-centric measures how did the company do or how would the customer recommend the company puts the focus on the customer and how easy was the process for them, so moves the focus on a transactional level to a much more direct (and applicable) concept: from "*delight the customer*", which is a very vague statement, to "*solve the customer issue / need easily*", a much more direct and applicable statement.

SURVEY APPROACH

All the above mentioned methodologies, at the end, are not mutually exclusive. There are companies that use a combination of average and CSat or NPS scores so it helps quantifying the impact of outliers and the sub-standard experiences.

Also, some of the interviewed companies described their customer satisfaction monitoring approach as a hybrid methodology:

> *"We have 2 surveys: the customer voice, which is triggered after each touchpoint, and a relational survey, which is done twice a year.*
>
> *The OES has 2 annual waves, every 6 months, by an external company and covers not only customer satisfaction but also product usage, competitors… it is done over the phone (specially this year) or physically visiting the customers that have to be interviewed.*
>
> *The survey, that is sent by email, varies depending on the product (if it is after a delivery, it will ask about dates, status, order of the material…) and has between 10 and 16 questions, an NPS question and an open text field for additional comments."*
>
> *B2B Company with a mature customer satisfaction process*
>
> *"We have 2 surveys: an NPS survey done by an external company (2 waves, one each month, to actual and prospective customers) and a Voice of the Customer (VoC).*
>
> *The VoC survey always follows the same pattern: an NPS question, the satisfaction with the channel (0-10, measuring the average), an open text field and some channel related questions. In total it is around 10 questions.*
>
> *The VoC is triggered after all the different touchpoints as well as what we call moments of truth (for instance, a new customer, a mortgage signing…). It is sometimes used to understand better the impact and opinion of the customers on particular things, for instance, when we changed the ATMs for new ones.*
>
> *The NPS survey, which is longer than the VoC (around 15') allows us to apply very small segmentation […]. In this survey we do not provide any information, it is done by an external company by calling a random sample of population and gets data for the entire industry so it can provide a very valuable benchmark, but with very limited levers and specific insight."*
>
> *B2C / B2B Company with a mature customer satisfaction process*

As seen in the above quotes, the approach in regards to when to survey the customers also varies. Usually, there are relatively short and quick transactional measurements, that compensates the lack of depth with a high frequency of information, allowing to make tactical decision making and providing insights for strategic decision making, and longer surveys done with low frequency that compensate the lack of frequency with wide and in-depth information. As in a market research environment, the low frequency relational surveys would be the equivalent of the qualitative research while the high frequency transactional surveys are the quantitative research.

REALITY VS MEASUREMENT

THE STRENGTH AND RISK OF SIMPLE THINGS

One of the key benefits of metrics such as CSat, NPS or CES is that they are very simple and easy to calculate. This simplicity allows using the metric not just as a high level indicator, but are also easier to understand by the frontline staff than complex indexes such as the ACSI. The concept of satisfied / dissatisfied, promoter / detractor or effort is much simpler to explain and understand that an econometrics model. Not only that, but also will be easier for this type of employees to have the goal (delight the customer, make the process effortless…) in mind when dealing with each client, bringing the concept of customer satisfaction from a vague metric calculated from unknown variables to an actionable measurement that can be considered during every single customer interaction.

The fact that the NPS or the CSat metric are very easy to understand and to drive for the frontline members of a company was one of the key points that drove the high acceptance and application of the NPS as a metric. Actually, this was considered in the HBR article from Reichheld that presented the concept of the Net Promoter Score:

> *"Good luck to the branch manager who tries to help an employee interpret a score resulting from a complex weighting algorithm based on feedback from anonymous customers, many of whom were surveyed before the employee had his current job.*
>
> *Contrast that scenario with one in which a manager presents employees with numbers from the previous week (or day) showing the percentages (and names) of a branch office's customers who are promoters, passively satisfied, and detractors—and then issues the managerial charge, 'We need more promoters and fewer detractors in order to grow.' The goal is clear-cut, actionable, and motivating."*
>
> *(REICHHELD F. F., 2004)*

Even though all the metrics previously exposed have abundant research supporting the methodology, at the end of the day, they represent a subjective perception on a subjective scale on a given moment on time. This is an important consideration because if it is considered an "absolute truth", any decision based exclusively on this type of metrics will weaken the entire process. The best approach to identify areas that need improvement will always be adding context to each score, otherwise the direction suggested by the score will always be vague. In fact, this is stated in the article that introduced the CES to the world:

> *"Customer service organizations can use CES, along with operational measurements of such things as repeat calls, transfers, and channel switching, to conduct an "effort audit" and improve areas where customers are expending undue energy." (DIXON, FREEMAN, & TOMAN, 2010)*

Even Reichheld work prior to the launch of the NPS score highlighted the importance of not oversimplifying the analysis by assessing areas like customer experience and customer loyalty from just one metric, as it may be counterproductive, and his own company (Bain and Company) has morphed the acronym NPS from Net Promoter Score to Net Promoter System, transforming what was originally as simple as "all you need to know is the answer to one question" to what Fisher and Kordupleski define as "an expensive and intricate gathering of anecdotal information by personnel at all levels of the company, leading to a complex summary of non-statistically reliable customer complaints" (FISHER & KORDUPLESKI, 2019).

DECISION MAKING FACTORS

Another key area to identify when quantifying the impact of customer satisfaction is the impact that customer satisfaction has on future purchase decision making. As identified earlier, NPS and CES present a high correlation with intent. However it is very important to differentiate between the intent and the actual repurchase, because only the latter drives direct revenue (although the first one may generate referrals).

In the survey one of the areas asked the respondents asked them to rank 10 factors from the most important to the least important. Plotted into a distribution chart (the lower the score the higher the importance), the results can be found in figure 9.

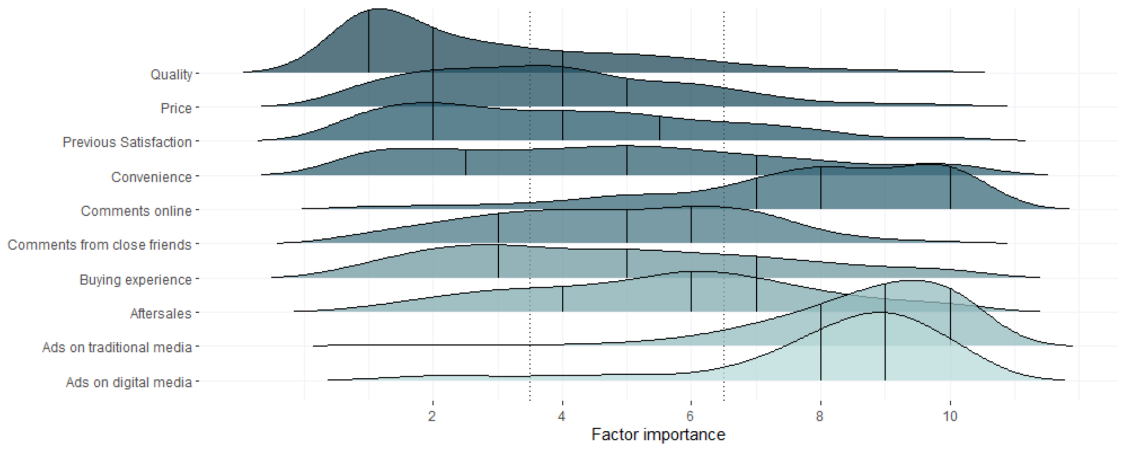

Fig. 9: Future purchase decision making factor importance distribution including the 3 quartiles (0.25, 0.5 and 0.75).
Source: Self-generated from conducted survey.

Where we can identify:

- 3 factors tending to be on top: quality, previous satisfaction and price.
- 4 more factors with a wide distribution but being usually out of the less important factors: convenience, comments from close people, buying experience and aftersales experience.
- 3 factors almost always at the bottom end of the ranking: comments online and advertisements (digital or traditional).

Checking the percentage of times each one of the factors has been placed on the top 3, we can have the following results (fig. 10).

Factor	#1	#2	#3	Top 3
Quality	45%	15%	13%	**73%**
Previous satisfaction	16%	19%	14%	**49%**
Convenience	13%	11%	15%	39%
Price	12%	21%	16%	**49%**
Buying experience	7%	15%	15%	37%
Comments from close	5%	11%	13%	29%
Comments online	4%	2%	2%	8%
Aftersales	3%	7%	13%	23%
Ads on traditional med	1%	2%	2%	5%
Ads on digital media	1%	3%	2%	5%

Fig. 10: Percentage of times each factor has been put on any of the top 3 positions and overall on the top 3.
Source: Self-generated from conducted survey.

Where we can identify quality, previous satisfaction and price as the top 3 factors followed by convenience and buying experience. The other factors are most frequently out of the top 3 factors (at least 70% of the times). From that, and analysing the interaction of the top 5 factors with different variables (touchpoint type, response rates, gender, education and age band), there are a few interactions significant:

- Quality: none of the mentioned variables have any influence on the position on the ranking.
- Previous Satisfaction: none of the mentioned variables have any influence on the position on the ranking.
- Price: the age band has an influence on the importance of price. In this case, we can see that while the customer gets old the price importance tends to go down, until retirement when the trend seems to get reverted (fig. 11).

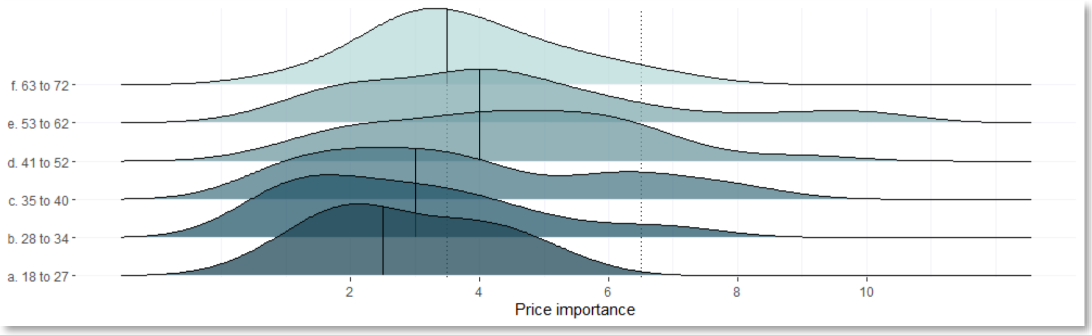

Fig. 11: Importance assigned to price for purchase decision making by age band.
Source: Self-generated from conducted survey.

- Convenience: response rate when dissatisfied as well as its interaction with general response rate. The interaction between general response rate and gender also have an influence (fig. 12).

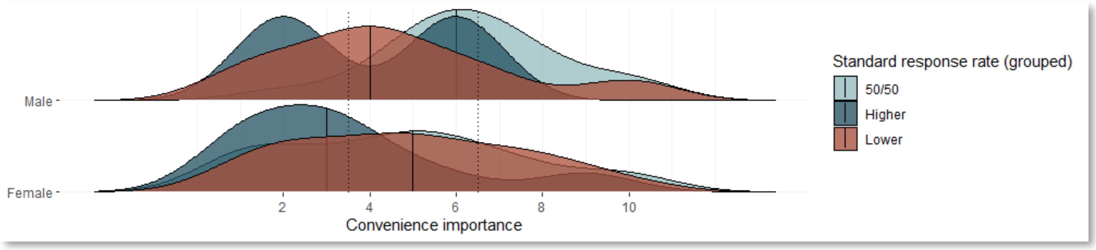

Fig. 12: Importance assigned to convenience for purchase decision making by gender and standard response rate (grouped into 3 bands).
Source: Self-generated from conducted survey.

- Buying Experience: the touchpoint interaction type together with response rate when dissatisfied have an influence on its position on the ranking. Although the touchpoint type and response rate when dissatisfied do not have any influence on the satisfaction scoring analysed, they do have an influence on the importance of the buying experience. The respondents who gave the buying experience higher importance tend to be either as strict or more strict with human touchpoints than with automated touchpoints and have a higher response rate when dissatisfied, which underlines the importance not just of automating efficiently, but of not forgetting the non-automated touchpoints as they are going to be critical for the customer satisfaction (fig. 13).

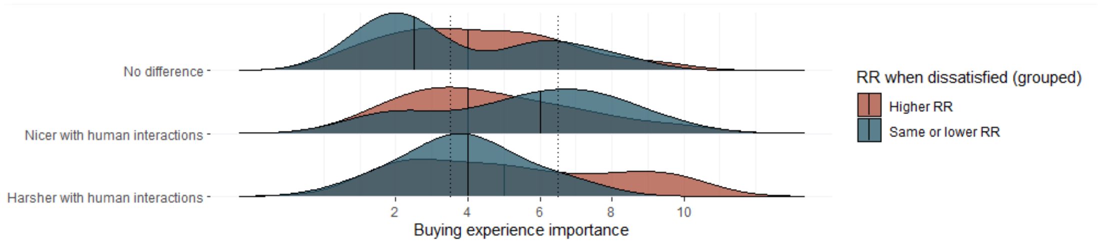

Fig. 13: Importance assigned to the buying experience for purchase decision making by how do they score different touchpoint types and response rate when satisfied.
Source: Self-generated from conducted survey.

MODELLING DECISION MAKING FACTORS AND PREDICTING SATISFACTION

Different strategies have been attempted: MCA, CHAID, Linear Regression… but none have been successful. That can be due to two reasons:

1. Sample size: the size of the dataset, both amount of surveys received and variables in the data, is a limiting factor, only extremely significant factors may end up generating any values to predict but with low modelling capabilities.
2. Correlated variables: some of the variables have correlations, as we have already seen. This fact will reduce the modelling power of the existing dataset, but eliminating them will reduce even further the amount of datapoints.

However, it is an analysis that can be done with a larger dataset both on rows and variables not only to predict the satisfaction, also with a wider variety of sociodemographic parameters as well as previous customer information (specially for already existing customers) can help predicting the importance that they give to the different decision making factors, which will help on designing a more tailor-made strategy to improve the experience for each identified customer segment.

NON STANDARDIZED APPROACH

On "The one number you need to grow", Reichheld already introduced an issue that has ended up becoming an issue of his own metric: result hacking.

> *"In most cases, dealers told me, the satisfaction survey is a charade that they play along with to remain in the good graces of the manufacturer and to ensure generous allocations of the hottest-selling models. The pressure they put on salespeople to boost scores often results in post-sale pleading with customers to provide top ratings—even if they must offer something like free floor mats or oil changes in return. Dealers are usually complicit with salespeople in this process, a circumstance that further degrades the integrity of these scores. Indeed, some savvy customers negotiate a low price— and then offer to sell the dealer a set of top satisfaction survey ratings for another $500 off the price."*
>
> *(REICHHELD F. F., 2004)*

In reality, one of the biggest risks of implementing this type of metrics and linking them to benefits (compensation, product availability…) is that the individuals, if not entire teams, will easily find ways to *cheat* on the scores that will be very difficult to identify by the company. On the interviewed companies with mature processes both reported it as an area of concern. In fact, both said that it was strictly forbidden to guide the customer on any way but accepted that it was almost impossible to identify if an individual was doing it. In fact 62.4% of the surveyed people reported that they have been guided and 38.7% that they have had to answer a satisfaction survey to the same person they were evaluating (fig. 14).

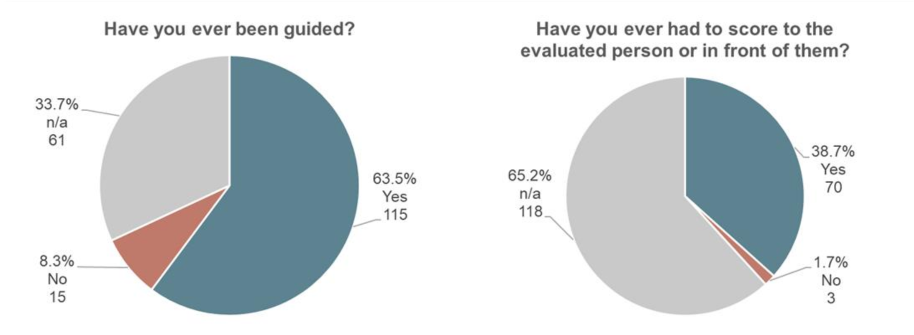

Fig. 14: Proportion of responses to each option on the questions 'Have you ever been guided?' and 'Have you ever had to score to the evaluated person or in front of them?'.
Source: Self-generated from conducted survey.

At the same time, in both cases they implemented the metrics from the bottom up, which made that different regions and divisions had different strategies and communications for the same metric. During the 2nd half of the 2010's both companies consolidated and centralised the approach to customer satisfaction and were able to identify not only the different strategies but also differences on execution:

> *"Before the creation of the team there was no standard: each team had its own approach, rules, methodologies, measurements… The team did an audit to identify and standardize the approach globally."*
>
> *B2C / B2B Company with a mature customer satisfaction process*
>
> *"We identified a difference of more than 20 points between [Region A] and [Region B] in 2018. After investigating, we found out that the [Region A] branch was guiding its customers. Since then is strictly forbidden and the process is standardized globally."*
>
> *B2B Company with a mature customer satisfaction process*

This is not only highlighted by the companies but it is also mentioned by some respondents on the survey comments which ultimately damages the image that the customers have on the survey:

> *"After a phone call to any company, the employee always asks you to score them well on the survey that they are going to transfer you."*
>
> *"I think that [the surveys] must be listened."*

In this research one of the objectives was to try to identify whether there is an impact by introducing a soft-guide in the survey user interface. To evaluate the impact 2 surveys were distributed at random: one presenting a neutral scoring with the only mentioning that 0 is the lowest score and 10 the highest in the question, and another with a coloured scale right above the scoring section as can be seen in figure 15.

Fig. 15: User view while answering the survey for the guided survey type.
Source: Self-generated from conducted survey.

The respondents had to evaluate 3 fictional scenarios:

1. The attention has been beyond your expectation and you are positively surprised.
2. The attention has been what you expected, but with no surprises (positive nor negative).
3. The attention received has been far from what you expected and you are dissatisfied.

The analysis performed on each one of the scenarios delivered the following outcomes.

SATISFIED

When compared, the distributions of guided vs neutral surveys on the satisfied scenario are almost exactly the same (fig. 16).

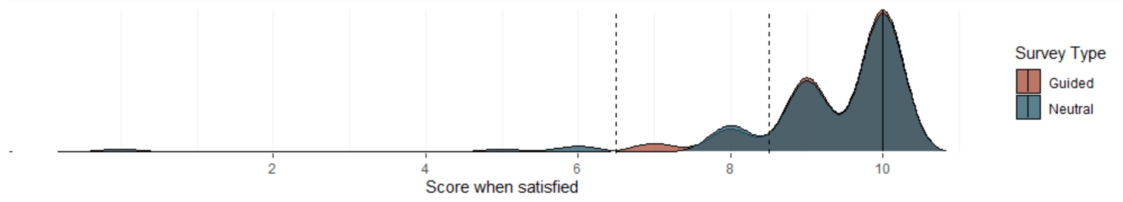

Fig. 16: Distribution of satisfied scores depending on the survey type including median.
Source: Self-generated from conducted survey.

It is interesting to identify that in the neutral survey there are a few surveys scoring with the lowest possible score. It may be an error from the customer when not paying enough attention.

The output of the ANOVA analysis identifies no dependency between the survey and the output as can be seen below:

```
> rawData_SatScores %>% aov(Score_RateSatisfied ~ SurveyType, data = .) %>%
summary()
             Df Sum Sq Mean Sq F value Pr(>F)
SurveyType    1    1.5   1.504   1.146  0.286
Residuals   169  221.8   1.312
```

When analysed considering the NPS logic, it results on a total NPS score of 83.6. And when tested using the NPS package, that includes a built in Wald Chi-Squared test for the score, it offers the below output:

```
> nps.test(NPS_SType$Neutral$Score_RateSatisfied,
NPS_SType$Guided$Score_RateSatisfied, test = 'wald', conf = 0.95, breaks =
npsBreaks)
Two sample Net Promoter Score Z test

NPS of x: 0.81 (n = 90)
NPS of y: 0.86 (n = 81)
Difference: 0.05

Standard error of difference: 0.064
Confidence level: 0.95
p value: 0.4092764
Confidence interval: -0.07300611 0.1791789
```

Which is aligned with the previous findings.

Similarly, when we do an ANOVA test on the CSat scores (considering values 8 or higher as satisfied and any value below 8 as non-satisfied), the output rejects the null hypothesis of dependency: there's no statistical influence on the result when the customer is guided with a colour scale:

```
> rawData_SatScores %>% aov(cSat_RateSatisfiedNum ~ SurveyType, data = .) %>%
summary()
             Df Sum Sq Mean Sq F value Pr(>F)
SurveyType    1  0.007 0.00737   0.186  0.667
Residuals   169  6.706 0.03968
```

NEUTRAL

In the score distribution for the neutral scenario can be appreciated a slightly bigger difference, although the quartiles are at the same point (fig. 17).

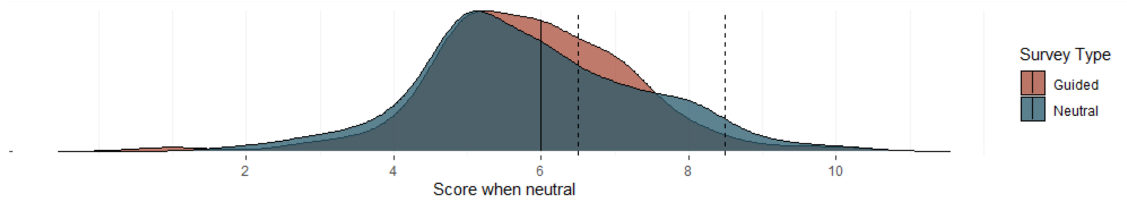

Fig. 17: Distribution of neutral scores depending on the survey type including median.
Source: Self-generated from conducted survey.

The output of the ANOVA analysis, however, does also fails to reject the null hypothesis (result is independent from the variable):

```
> rawData_SatScores %>% aov(Score_RateNeutral ~ SurveyType, data = .) %>% summary()
             Df Sum Sq Mean Sq F value Pr(>F)
SurveyType    1    0.0   0.001   0.001  0.982
Residuals   169  330.6   1.956
```

The result is the same when analysed considering the NPS logic (first R console output) or the CSat logic (second R console output):

```
> nps.test(NPS_SType$Neutral$Score_RateNeutral,
NPS_SType$Guided$Score_RateNeutral,   test = 'wald', conf = 0.95, breaks =
npsBreaks)
Two sample Net Promoter Score Z test

NPS of x: -0.66 (n = 90)
NPS of y: -0.69 (n = 81)
Difference: 0.04

Standard error of difference: 0.081
Confidence level: 0.95
p value: 0.6568319
Confidence interval: -0.1221379 0.1937428

> rawData_SatScores %>% aov(cSat_RateNeutralNum ~ SurveyType, data = .) %>%
summary()
             Df Sum Sq Mean Sq F value Pr(>F)
SurveyType    1  0.236  0.2359   2.288  0.132
Residuals   169 17.425  0.1031
```

DISSATISFIED

In the dissatisfied scenario can be appreciated the same difference than in the satisfied one: there is an even more noticeable amount of high scored dissatisfied customers (fig. 18).

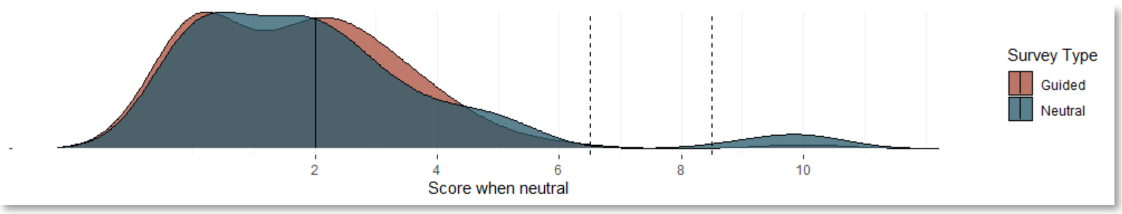

Fig. 18: Distribution of dissatisfied scores depending on the survey type including median
Source: Self-generated from conducted survey.

As with the previous analysis, the ANOVA test finds the differences inconclusive not being able to confirm any dependency:

```
> rawData_SatScores %>% aov(Score_RateDissat ~ SurveyType, data = .) %>% summary()
             Df Sum Sq Mean Sq F value Pr(>F)
SurveyType    1    3.6   3.600   0.868  0.353
Residuals   169  700.7   4.146
```

As with the previous scenarios, the NPS and CSat logics present no dependency between the survey type and the score (although the p-value is below 0.20 on both tests):

```
> nps.test(NPS_SType$Neutral$Score_RateDissat,    NPS_SType$Guided$Score_RateDissat,
test = 'wald', conf = 0.95, breaks = npsBreaks)
Two sample Net Promoter Score Z test

NPS of x: -0.91 (n = 90)
NPS of y: -0.98 (n = 81)
Difference: 0.06

Standard error of difference: 0.05
Confidence level: 0.95
p value: 0.1982283
Confidence interval: -0.03359767 0.1619927

> rawData_SatScores %>% aov(cSat_RateDissatNum ~ SurveyType, data = .) %>% summary()
             Df Sum Sq Mean Sq F value Pr(>F)
```

```
SurveyType    1  0.056 0.05577   1.964  0.163
Residuals   169  4.798 0.02839
```

ANALYSIS CONCLUSION

The results were analysed using the three first methodologies explained (average score, NPS and CSat), and there were no significant differences on the results comparing guided vs non-guided surveys. However, in the non-guided surveys there were some responses on both the satisfied and dissatisfied scenarios that resulted in scores on the opposite side of the scale (dissatisfied customers rating with a 10 and satisfied customers rating with a 0).

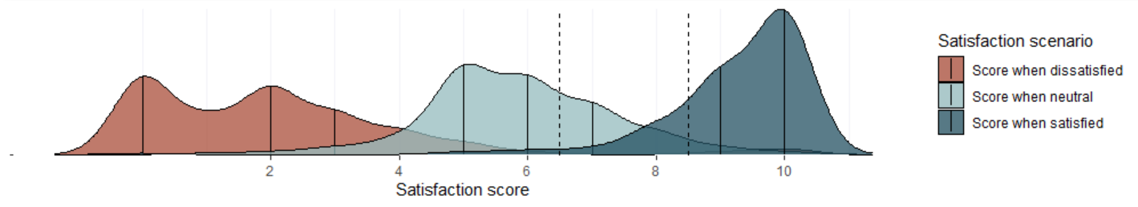

Fig. 19: Distribution of the satisfaction scores on the different scenarios the 3 quartiles (0.25, 0.50 and 0.75).
Source: Self-generated from conducted survey.

From this results, we can say that soft-guiding has no negative influence on the satisfaction scores and that it might help avoiding errors on interpreting the scale (although it would be a subject to investigate further).

THE IMPORTANCE OF CONTEXT

Even with all the mentioned research, customer satisfaction as a metric is just an indicator based on the difference between the perceived quality (a subjective perception) and expected quality (a subjective expectation) on a given moment in time (therefore influenced by the respondent environment and personal situation) and on a scale (influenced by cultural traditions and considerations). It means that the number as itself does not provide any value. Even the same metric on different companies, if the measurement is not conducted following the same guidelines (process, timelines, channels, structure…) and ensuring a similar sample or a normalized result (similar proportion of respondents from the variables that have a meaningful impact on the metric), will not be a comparable value to benchmark.

Therefore it should be used as an indicator of evolution as well as an indicator of opportunities to improve, although it would be important to normalize the results to ensure that changes on the survey drivers (for instance channel mix) do not drive the results without a change in actual performance.

The opportunities to improve will never be identified by the metric itself but by the additional information from the survey or that can be connected to it:

- Interaction information (for transactional surveys): product, type of touchpoint, type of interaction, length of interaction, channel, issue type…
- Customer history: purchases, previous satisfaction, previous complaints…
- Sociodemographic information: age, gender, studies, type of job…

The importance of not using Net Promoter Score (or any satisfaction measure) as a standalone metric is perfectly described with the Oldsmobile case, one of the oldest car manufacturing companies, that had a consistently increasing NPS score but that went out of business in 2004 as at the same time their satisfaction rate increased, sales were plunging. The reason was that they were not converting detractors into promoters, they just lost their detractors for the competitors keeping a constantly shrinking customer base.

From the research done, we have identified that the behaviour when evaluating a human touchpoint is different than when evaluating an automatic touchpoint (fig. 20).

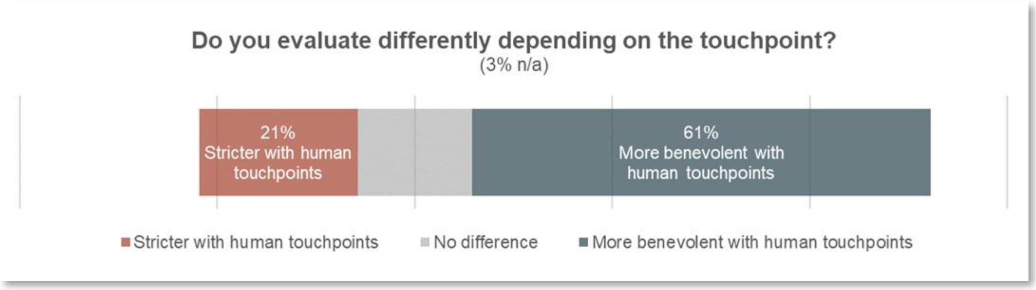

Fig. 20: Personal perspective on rating different types of touchpoints.
Source: Self-generated from conducted survey.

When evaluating factors that generate different satisfaction the analysis the analysis delivered the following outcomes.

SATISFIED

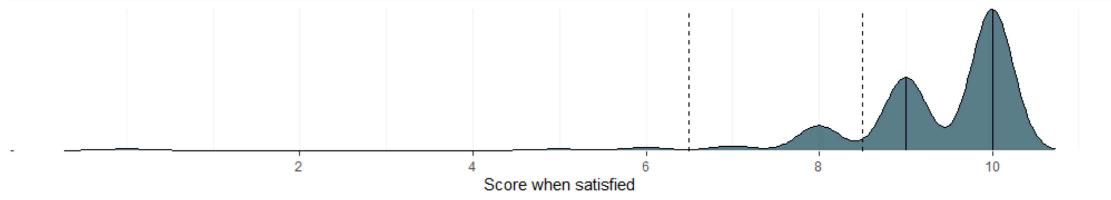

Fig. 21: Distribution of satisfied scores including the 3 quartiles.
Source: Self-generated from conducted survey.

When analysing the factors in general and also for respondents with recent job experience, there was only one factor identified as significant for both: the response rate when dissatisfied:

```
> ### 05a. SATISFIED  ####
> ### AVG SCORE
> rawData_SatScores %>% aov(Score_RateSatisfied ~ RespRate_Std + RespRate_Sat +
RespRate_Dissat + Gender_G + Education_G + AgeBand + Price_Top3 + Quality_Top3 +
Convenience_Top3 + PrevSatisf_Top3 + BuyingExp_Top3, data = .) %>% summary()
                  Df Sum Sq Mean Sq F value  Pr(>F)
RespRate_Std       4   3.42   0.856   0.687 0.60197
RespRate_Sat       2   3.57   1.786   1.434 0.24176
RespRate_Dissat    2  16.37   8.184   6.569 0.00185 **
Gender_G           2   2.38   1.192   0.956 0.38670
Education_G        4   3.08   0.770   0.618 0.65050
AgeBand            5   6.12   1.225   0.983 0.43019
Price_Top3         1   1.19   1.187   0.953 0.33058
Quality_Top3       1   0.46   0.463   0.372 0.54300
Convenience_Top3   1   2.30   2.298   1.844 0.17655
PrevSatisf_Top3    1   2.06   2.058   1.652 0.20078
BuyingExp_Top3     1   0.44   0.444   0.356 0.55151
Residuals        146 181.91   1.246
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> rawData_SatScores %>% aov(Score_RateSatisfied ~ RespRate_Std + RespRate_Sat +
RespRate_Dissat + Gender_G + Education_G + AgeBand + Job_Customer + Job_Manager +
Job_SatSurveysCons + Price_Top3 + Quality_Top3 + Convenience_Top3 + PrevSatisf_Top3
+ BuyingExp_Top3, data = .) %>% summary()
                    Df Sum Sq Mean Sq F value     Pr(>F)
RespRate_Std         3   3.63   1.211   0.962     0.4135
RespRate_Sat         2   4.16   2.081   1.654     0.1961
RespRate_Dissat      2  33.18  16.588  13.183 0.00000762 ***
Gender_G             2   1.59   0.794   0.631     0.5339
Education_G          4   4.13   1.033   0.821     0.5144
AgeBand              5   4.97   0.995   0.791     0.5586
```

```
Job_Customer          4   5.83   1.458   1.159     0.3333
Job_Manager           2   0.21   0.104   0.083     0.9208
Job_SatSurveysCons    2   1.47   0.736   0.585     0.5591
Price_Top3            1   5.12   5.117   4.066     0.0463 *
Quality_Top3          1   1.09   1.091   0.867     0.3540
Convenience_Top3      1   1.81   1.814   1.442     0.2325
PrevSatisf_Top3       1   4.06   4.055   3.223     0.0754 .
BuyingExp_Top3        1   0.60   0.598   0.475     0.4921
Residuals           107 134.64   1.258
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
32 observations deleted due to missingness
```

However, it is a factor with no influence on the satisfied score. When repeated for the NPS and CSat scores, the ANOVA model identifies price as relevant for both NPS and CSat and previous satisfaction is almost significant for the group of respondents that have a recent work experience. However, if we visualize the scores of the 2 factors (figs. 22 and 23) we'll see that there are a few responses on each that scores as satisfied with low scores. Assuming that scores below 6 are errors interpreting the scale and re-running the ANOVA model excluding those, we get the below output confirming that the significance was coming from errors on the responses.

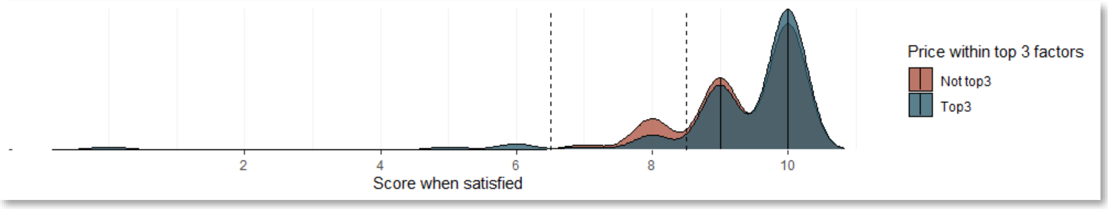

Fig. 22: Distribution of satisfied scores comparing customers that put price within the top 3 future purchase decision making and customers that did not.
Source: Self-generated from conducted survey.

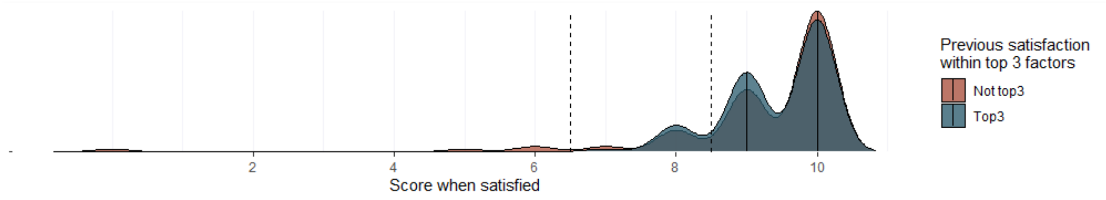

Fig. 23: Distribution of satisfied scores comparing customers that put the previous satisfaction within the top 3 future purchase decision making and customers that did not.
Source: Self-generated from conducted survey.

```
> rawData_SatScores %>% filter(Score_RateSatisfied > 6) %>% aov(Score_RateSatisfied
~ RespRate_Std + RespRate_Sat + RespRate_Dissat + Gender_G + Education_G + AgeBand +
Job_Customer + Job_Manager + Job_SatSurveysCons + Price_Top3 + Quality_Top3 +
Convenience_Top3 + PrevSatisf_Top3 + BuyingExp_Top3, data = .) %>% summary()
                    Df Sum Sq Mean Sq F value    Pr(>F)
RespRate_Std         3   1.20   0.401   0.807    0.4925
RespRate_Sat         2   0.40   0.199   0.400    0.6713
RespRate_Dissat      2  12.07   6.037  12.144 0.0000184 ***
Gender_G             2   1.10   0.548   1.103    0.3357
Education_G          4   1.11   0.276   0.556    0.6953
AgeBand              5   2.77   0.555   1.116    0.3566
Job_Customer         4   4.71   1.177   2.368    0.0575 .
Job_Manager          2   0.17   0.086   0.172    0.8421
Job_SatSurveysCons   2   0.30   0.152   0.305    0.7378
Price_Top3           1   0.18   0.183   0.369    0.5448
Quality_Top3         1   0.11   0.112   0.226    0.6358
Convenience_Top3     1   0.89   0.887   1.785    0.1845
PrevSatisf_Top3      1   0.98   0.981   1.974    0.1630
BuyingExp_Top3       1   0.01   0.008   0.017    0.8963
Residuals          103  51.21   0.497
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
32 observations deleted due to missingness
```

NEUTRAL

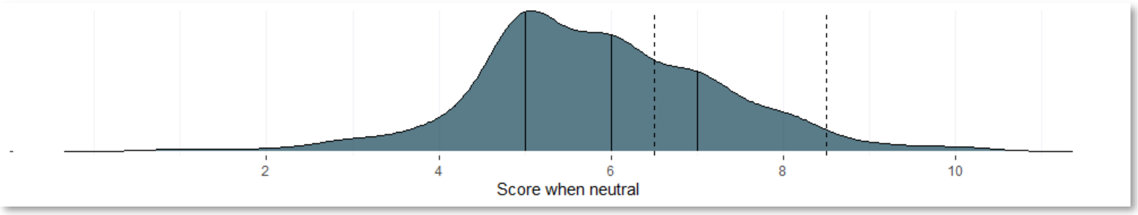

Fig. 24: Distribution of neutral scores including the 3 quartiles.
Source: Self-generated from conducted survey.

In the neutral scores, following the same analysis structure than in the satisfied, there are many factors that generate a different score. And that makes sense: neutral scores (satisfied but not surprised) do not have a pre-defined “standard” score as it may happen with the satisfied and the dissatisfied, although on a smaller measure as we will see later. This is also a critical area, as the neutral customers can move easily towards satisfied or dissatisfied with very little changes as well as going for alternative solutions.

The fact that the score will be naturally around the centre of the scale will create a bigger spread which will make easier to identify different factors influencing the score that, due to the combination of smaller variance and small sample size may not be identified on the satisfied or dissatisfied areas.

From the distribution, it is interesting to see that the median score is 6, which falls below the neutral/passive customer denomination of NPS (and below the satisfied in CSat) and the mode is scoring with a 5. With that we can say that from the analysed data, the tendency when the customer expectations are met but without nothing beyond their expectations is to score as what traditionally would mean “passing an exam” (between a 5 and a 7) but that means being in the frontier between neutral/passive and dissatisfied on an NPS scoring system.

Analysing the factors for neutral scores for all variables non-job related for everyone and then subsetting the dataset so it is analysed for people with recent job experience we can identify several of them having an influence on the neutral score:

```
> ### NEUTRAL SCORE
> rawData_SatScores %>% aov(Score_RateNeutral ~ RespRate_Std + RespRate_Sat +
RespRate_Dissat + Gender_G + Education_G + AgeBand + Price_Top3 + Quality_Top3 +
Convenience_Top3 + PrevSatisf_Top3 + BuyingExp_Top3, data = .) %>% summary()
                  Df Sum Sq Mean Sq F value  Pr(>F)
RespRate_Std       4   4.83   1.209   0.678 0.60832
RespRate_Sat       2  13.78   6.891   3.865 0.02313 *
RespRate_Dissat    2  12.05   6.023   3.379 0.03679 *
Gender_G           2   0.60   0.299   0.168 0.84579
Education_G        4   3.92   0.981   0.550 0.69912
AgeBand            5  14.93   2.985   1.675 0.14427
Price_Top3         1  14.19  14.194   7.962 0.00544 **
Quality_Top3       1   0.61   0.614   0.344 0.55835
Convenience_Top3   1   1.23   1.229   0.689 0.40778
PrevSatisf_Top3    1   1.10   1.101   0.618 0.43320
BuyingExp_Top3     1   3.10   3.102   1.740 0.18922
Residuals        146 260.28   1.783
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> rawData_SatScores %>% aov(Score_RateNeutral ~ RespRate_Std + RespRate_Sat +
RespRate_Dissat + Gender_G + Education_G + AgeBand + Job_Customer + Job_Manager +
```

```
Job_SatSurveysCons + Price_Top3 + Quality_Top3 + Convenience_Top3 + PrevSatisf_Top3
+ BuyingExp_Top3, data = .) %>% summary()
                  Df Sum Sq Mean Sq F value  Pr(>F)
RespRate_Std       3   6.38   2.127   1.300 0.27830
RespRate_Sat       2  13.17   6.586   4.026 0.02062 *
RespRate_Dissat    2  12.96   6.481   3.961 0.02189 *
Gender_G           2   0.78   0.389   0.238 0.78874
Education_G        4   6.25   1.563   0.955 0.43521
AgeBand            5  20.74   4.149   2.536 0.03283 *
Job_Customer       4   6.27   1.567   0.958 0.43396
Job_Manager        2   0.05   0.027   0.017 0.98353
Job_SatSurveysCons 2  13.11   6.557   4.007 0.02097 *
Price_Top3         1  15.64  15.644   9.562 0.00253 **
Quality_Top3       1   0.08   0.082   0.050 0.82322
Convenience_Top3   1   0.88   0.876   0.535 0.46601
PrevSatisf_Top3    1   1.58   1.579   0.965 0.32817
BuyingExp_Top3     1   3.62   3.622   2.214 0.13971
Residuals        107 175.06   1.636
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
32 observations deleted due to missingness
```

Although, when transforming a full scale from 0 to 10 into a scale with 3 levels (NPS) the significance of the different factors is reduced being significant only the response rate when satisfied and whether the respondent has a work related to satisfaction surveys. The significance of all factors disappear when translated into CSat.

Specially interesting is the difference that generates the fact of working with satisfaction surveys. If visualized, it can be seen that the distribution of scores is significantly different: the respondents that work with satisfaction surveys on any way tend to score significantly higher than those who does not (fig. 25).

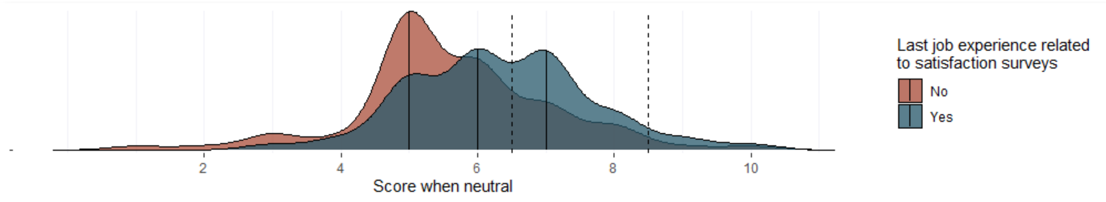

Fig. 25: Distribution of neutral scores comparing respondents whose last job experience is somehow related to satisfaction surveys and those that did not.
Source: Self-generated from conducted survey.

Similarly, if plotted the price and aftersales we will be able to see that those who put price (fig. 26) or aftersales (fig. 27) on their top3 factors tend to score lower than those who does not.

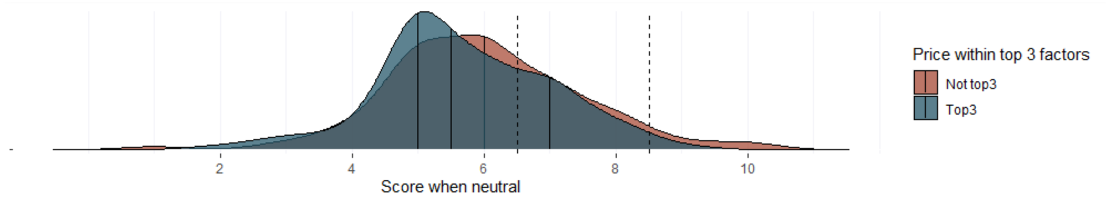

Fig. 26: Distribution of neutral scores comparing respondents that put price within the top 3 future purchase decision making and customers that did not.
Source: Self-generated from conducted survey.

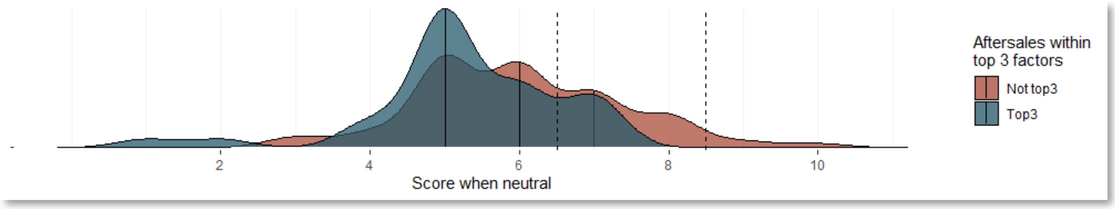

Fig. 27: Distribution of neutral scores comparing customers that put aftersales within the top 3 future purchase decision making and customers that did not.
Source: Self-generated from conducted survey.

With the age happens something similar, although in this case we can see a trend where the older the respondent, the lower the neutral score (fig. 28).

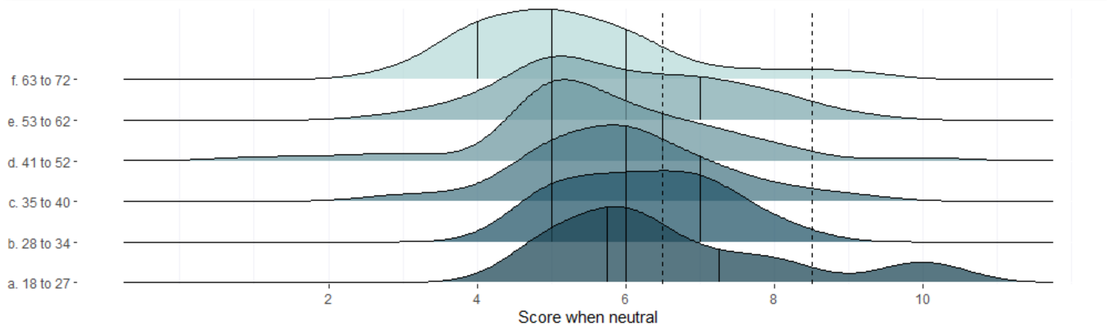

Fig. 28: Distribution of neutral scores depending on respondent age band.
Source: Self-generated from conducted survey.

If we analyse the correlation, we can find a weak although significant correlation between the two factors:

```
> Age_Neutral <- rawData_SatScores %>%
+   mutate(AgeBand = case_when(
+     AgeBand == 'a. 18 to 27' ~ 1,
+     AgeBand == 'b. 28 to 34' ~ 2,
+     AgeBand == 'c. 35 to 40' ~ 3,
+     AgeBand == 'd. 41 to 52' ~ 4,
+     AgeBand == 'e. 53 to 62' ~ 5,
+     AgeBand == 'f. 63 to 72' ~ 6,
+     TRUE                     ~ 0
+   )) %>%
+   select(AgeBand, Score_RateNeutral)
> cor.test(Age_Neutral$AgeBand, Age_Neutral$Score_RateNeutral, method = 'pearson')

Pearson's product-moment correlation

data:  Age_Neutral$AgeBand and Age_Neutral$Score_RateNeutral
t = -2.6953, df = 169, p-value = 0.007744
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.34264472 -0.05460263
sample estimates:
       cor
-0.2030115
```

DISSATISFIED

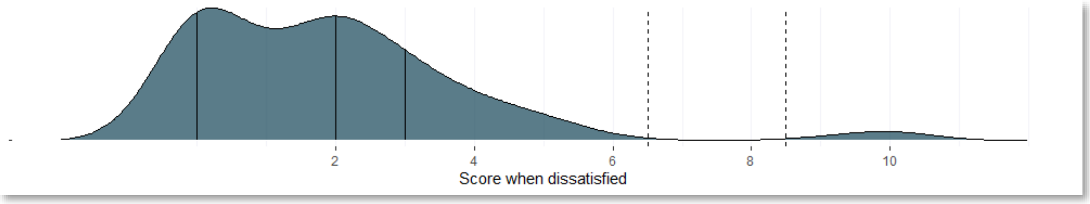

Fig. 29: Distribution of dissatisfied scores including the 3 quartiles.
Source: Self-generated from conducted survey.

The dissatisfied scenario is something like a middle-point between the satisfied, where the distribution was extremely skewed to the left, and the neutral score, where the distribution was quite wide. In this case, it is skewed to the right.

Doing the previously performed ANOVA but directly trimming the results scored above 8 to eliminate the high-scored dissatisfied evaluations, the only significant value is the response rate when dissatisfied for respondents with recent job experience, which is almost significant when analysed for the entire dataset:

```
> ### DISSAT SCORE CLEANED
> rawData_SatScores %>% filter(Score_RateDissat < 8) %>% aov(Score_RateDissat ~
RespRate_Std + RespRate_Sat + RespRate_Dissat + Gender_G + Education_G + AgeBand +
Price_Top3 + Quality_Top3 + Convenience_Top3 + PrevSatisf_Top3 + BuyingExp_Top3,
data = .) %>% summary()
                  Df Sum Sq Mean Sq F value Pr(>F)
RespRate_Std       4    7.0   1.752   0.780 0.5397
RespRate_Sat       2    7.7   3.859   1.719 0.1830
RespRate_Dissat    2   13.0   6.515   2.902 0.0582 .
Gender_G           2    9.1   4.537   2.021 0.1364
Education_G        4    8.4   2.103   0.937 0.4447
AgeBand            5   13.1   2.619   1.166 0.3288
Price_Top3         1    6.1   6.085   2.710 0.1020
Quality_Top3       1    1.5   1.540   0.686 0.4090
Convenience_Top3   1    0.1   0.147   0.065 0.7986
PrevSatisf_Top3    1    1.8   1.804   0.804 0.3715
BuyingExp_Top3     1    0.4   0.443   0.197 0.6577
Residuals        141  316.6   2.245
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> rawData_SatScores %>% filter(Score_RateDissat < 8) %>% aov(Score_RateDissat ~
RespRate_Std + RespRate_Sat + RespRate_Dissat + Gender_G + Education_G + AgeBand +
Job_Customer + Job_Manager + Job_SatSurveysCons + Price_Top3 + Quality_Top3 +
Convenience_Top3 + PrevSatisf_Top3 + BuyingExp_Top3, data = .) %>% summary()
                    Df Sum Sq Mean Sq F value Pr(>F)
RespRate_Std         3   4.84   1.612   0.701 0.5533
RespRate_Sat         2   6.21   3.106   1.352 0.2633
RespRate_Dissat      2  19.51   9.757   4.246 0.0169 *
Gender_G             2   8.29   4.144   1.804 0.1698
Education_G          4   7.05   1.761   0.767 0.5494
AgeBand              5  19.39   3.878   1.688 0.1439
Job_Customer         4  14.69   3.673   1.598 0.1802
Job_Manager          2   0.89   0.445   0.194 0.8243
Job_SatSurveysCons   2   3.85   1.925   0.838 0.4356
Price_Top3           1   4.72   4.722   2.055 0.1547
Quality_Top3         1   0.34   0.343   0.149 0.7001
Convenience_Top3     1   0.51   0.510   0.222 0.6386
PrevSatisf_Top3      1   2.06   2.062   0.897 0.3456
BuyingExp_Top3       1   1.50   1.495   0.651 0.4217
Residuals          105 241.27   2.298
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
29 observations deleted due to missingness
```

Analysing the different scores depending on the response rate, we can see that the ones that will have higher response rate is the group that will score the lowest (fig. 30). However, that would only affect when the satisfaction measure is based on the average.

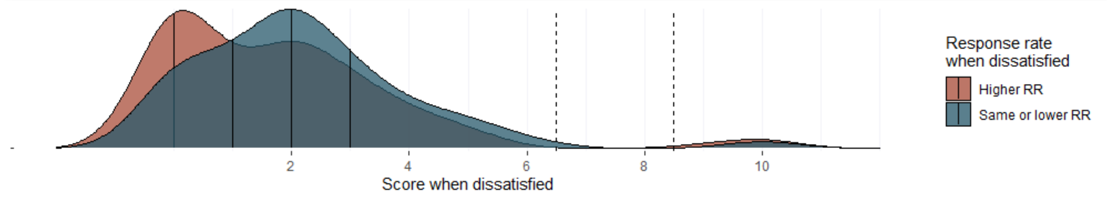

Fig. 30: Distribution of dissatisfied scores including the 3 quartiles depending on their response rate when dissatisfied.
Source: Self-generated from conducted survey.

The results for NPS and CSat when eliminated the highly scored responses (assuming they are errors interpreting the scale) are that there is no significant difference: all the customers are dissatisfied.

ANALYSIS CONCLUSION

The most critical measurement is the neutral score, which is where the "swing" customers are going to be. They are close to be satisfied, but they have not been surprised. From the analysis it has been identified that for the customers that consider price or the aftersales service important factors, not being enough satisfied will generate a lower score. Similarly, we've also identified a correlation: younger customers tend to give higher scores and the older the customer is, the lower the score. That can be connected with changes on the scoring of exams (from scoring exams from 0 to 10 to scoring exams from *F* to *A* or other scoring scales), but that is just an assumption.

Another very important factor is that when the survey is submitted by someone that works around satisfaction surveys, the score tends to be higher. That can be due to 2 factors:

1. Even without guiding (it has no influence in the analysed dataset), their background makes them score naturally considering the NPS logic, *polluting* the results or scoring based on different logic.
2. They tend to be less strict when they are "not enough satisfied" as some kind of *kindness* towards the individual they are scoring.

A final important finding is that the customers that report having a higher response rate when dissatisfied also tend to give lower scores. However it will have an impact on companies that work with average score as on NPS or CSat the "dissatisfied score" will be the same for any value below 6 (NPS) or 7 (for CSat, although the value may be different depending on the company).

The touchpoint type does not generate any difference in the results as it has not been considered for the scenarios, but as seen on page 23, the big difference towards the scoring mindset on the mentioned interaction types shows that it is likely to generate a difference.

From the comments left by the respondents on the survey, we can identify some that highlight the different mindsets towards the surveys that are coming from people with different backgrounds:

> *"Since retiring, I have more time to answer satisfaction surveys"*

RESPONSE RATE DIFFERENCES

Another of the issues that this type of metric can generate is the discrepancy between your actual customer satisfaction and the real customer satisfaction.

Although the current research has been focused on socio-demographic differences, it is to be expected that different verticals within companies may trigger different response rates depending on the impact the underperformance has in the usage of the product (for instance, in the telecom industry it may not trigger the same response rate a dissatisfaction because there is a typo on the name on an invoice than a dissatisfaction coming from 3 days without service). Still, at high level and before further analysing the data, a pattern can be identified in the response rate behaviour depending on the satisfaction level (fig. 31).

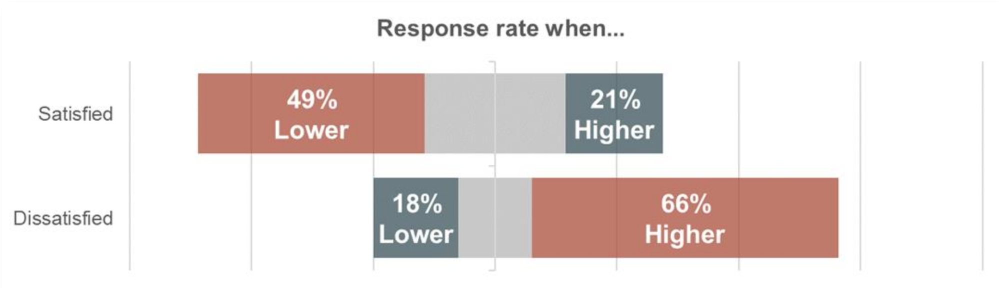

Fig. 31: Different response rate tendencies depending on the level of satisfaction.
Source: Self-generated from conducted survey.

STANDARD RESPONSE RATE

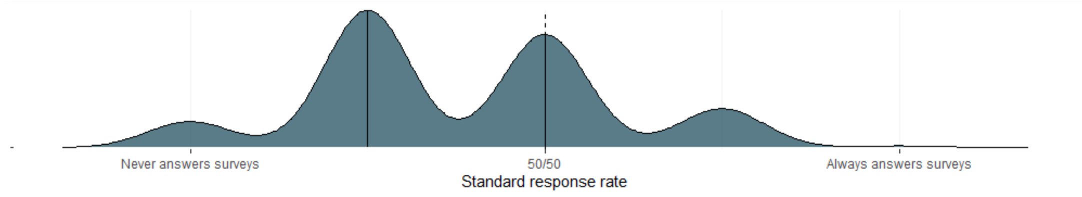

Fig. 32: Distribution of standard response rate.
Source: Self-generated from conducted survey.

Transforming the response rate into a numerical variable (-2 being never answer the surveys and +2 being always answer the surveys) we can see that the average is on -0.46, which means that the tendency is to not answer the surveys:

```
> ### RR standard
> mean(rawData_SatScores$RespRate_Std_Num, na.rm = TRUE)
[1] -0.4619883
> rawData_SatScores %>%
+   aov(RespRate_Std_Num ~ RespRate_Sat + RespRate_Dissat + Gender_G + Education_G +
AgeBand + Price_Top3 + Quality_Top3 + Convenience_Top3 + PrevSatisf_Top3 +
BuyingExp_Top3, data = .) %>% summary()
                  Df Sum Sq Mean Sq F value Pr(>F)
RespRate_Sat       2   3.53  1.7633   2.768 0.0660 .
RespRate_Dissat    2   1.47  0.7341   1.152 0.3187
Gender_G           2   2.34  1.1681   1.833 0.1634
Education_G        4   6.91  1.7276   2.712 0.0322 *
AgeBand            5   5.84  1.1676   1.833 0.1098
Price_Top3         1   0.35  0.3549   0.557 0.4567
Quality_Top3       1   1.27  1.2705   1.994 0.1600
Convenience_Top3   1   0.34  0.3374   0.530 0.4679
PrevSatisf_Top3    1   0.72  0.7196   1.129 0.2896
BuyingExp_Top3     1   0.17  0.1736   0.272 0.6025
Residuals        150  95.57  0.6371
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> rawData_SatScores %>%
+   aov(RespRate_Std_Num ~ RespRate_Sat + RespRate_Dissat + Gender_G + Education_G +
AgeBand + Price_Top3 + Quality_Top3 + Convenience_Top3 + PrevSatisf_Top3 +
BuyingExp_Top3 + Job_Customer + Job_Manager + Job_SatSurveysCons, data = .) %>%
summary()
                    Df Sum Sq Mean Sq F value Pr(>F)
RespRate_Sat         2   2.84  1.4207   2.281  0.107
```

```
RespRate_Dissat      2   2.23  1.1137   1.788  0.172
Gender_G             2   0.91  0.4574   0.734  0.482
Education_G          4   4.67  1.1673   1.874  0.120
AgeBand              5   2.52  0.5039   0.809  0.546
Price_Top3           1   1.05  1.0545   1.693  0.196
Quality_Top3         1   0.31  0.3070   0.493  0.484
Convenience_Top3     1   1.60  1.6007   2.569  0.112
PrevSatisf_Top3      1   1.39  1.3884   2.229  0.138
BuyingExp_Top3       1   0.00  0.0027   0.004  0.948
Job_Customer         4   4.18  1.0453   1.678  0.160
Job_Manager          2   0.92  0.4589   0.737  0.481
Job_SatSurveysCons   2   1.30  0.6478   1.040  0.357
Residuals          110  68.53  0.6230
32 observations deleted due to missingness
```

So we can find that the education level is the only significant variable. Generating the below distribution of responses for each value, we can find that there is a negative weak (although significant) correlation between response rate and education: the higher the education, the lower the standard response rate (fig. 33).

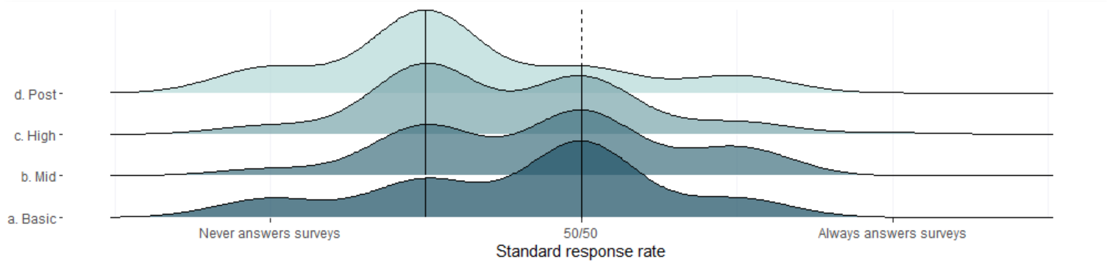

Fig. 33: Distribution of standard response rate depending on the education level.
Source: Self-generated from conducted survey.

```
> Edu_RR <- rawData_SatScores %>%
+   filter(Education_G != 'z. n/a') %>%
+   mutate(Education_G = case_when(
+     Education_G == 'a. Basic' ~ 1,
+     Education_G == 'b. Mid'   ~ 2,
+     Education_G == 'c. High'  ~ 3,
+     Education_G == 'd. Post'  ~ 4,
+     TRUE                      ~ 0
+   )) %>%
+   select(Education_G, RespRate_Std_Num)
> cor.test(Edu_RR$Education_G, Edu_RR$RespRate_Std_Num, method = 'pearson')

Pearson's product-moment correlation

data:  Edu_RR$Education_G and Edu_RR$RespRate_Std_Num
t = -2.6021, df = 167, p-value = 0.0101
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.33827605 -0.04786252
sample estimates:
       cor
-0.1973963
```

RESPONSE RATE WHEN SATISFIED

Fig. 34: Distribution of response rate when satisfied.
Source: Self-generated from conducted survey.

Analysing the variable considering -1 lower response rate and +1 higher response rate, we can find that still the trend is to answer less frequently the surveys when satisfied (the average is -0.3)

```
> ### RR when satisfied
> mean(rawData_SatScores$RespRate_Sat_Num, na.rm = TRUE)
[1] -0.2982456
> rawData_SatScores %>% aov(RespRate_Sat_Num ~ RespRate_Std + RespRate_Dissat +
Gender_G + Education_G + AgeBand + Price_Top3 + Quality_Top3 + Convenience_Top3 +
PrevSatisf_Top3 + BuyingExp_Top3, data = .) %>% summary()
                 Df Sum Sq Mean Sq F value                Pr(>F)
RespRate_Std      4   1.75   0.438   1.042               0.38749
RespRate_Dissat   2  37.19  18.597  44.248 0.000000000000000864 ***
Gender_G          2   0.37   0.186   0.443               0.64318
Education_G       4   2.07   0.518   1.232               0.29975
AgeBand           5   6.98   1.397   3.323               0.00712 **
Price_Top3        1   0.33   0.332   0.790               0.37563
Quality_Top3      1   0.16   0.160   0.380               0.53830
Convenience_Top3  1   0.19   0.185   0.441               0.50758
PrevSatisf_Top3   1   0.31   0.314   0.746               0.38906
BuyingExp_Top3    1   0.23   0.225   0.535               0.46546
Residuals       148  62.20   0.420
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> rawData_SatScores %>% aov(RespRate_Sat_Num ~ RespRate_Std + RespRate_Dissat +
Gender_G + Education_G + AgeBand + Price_Top3 + Quality_Top3 + Convenience_Top3 +
PrevSatisf_Top3 + BuyingExp_Top3 + Job_Customer + Job_Manager + Job_SatSurveysCons,
data = .) %>% summary()
                    Df Sum Sq Mean Sq F value       Pr(>F)
RespRate_Std         3   2.78   0.925   2.095       0.1051
RespRate_Dissat      2  17.11   8.556  19.377 0.0000000631 ***
Gender_G             2   0.65   0.325   0.737       0.4809
Education_G          4   2.06   0.515   1.167       0.3294
AgeBand              5   5.25   1.051   2.380       0.0432 *
Price_Top3           1   0.38   0.384   0.870       0.3530
Quality_Top3         1   0.09   0.086   0.195       0.6595
Convenience_Top3     1   0.39   0.391   0.886       0.3488
PrevSatisf_Top3      1   0.01   0.006   0.014       0.9056
BuyingExp_Top3       1   0.06   0.060   0.136       0.7134
Job_Customer         4   0.69   0.174   0.393       0.8129
Job_Manager          2   0.09   0.043   0.098       0.9070
Job_SatSurveysCons   2   1.93   0.966   2.188       0.1171
Residuals          109  48.13   0.442
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
32 observations deleted due to missingness
```

The data also shows that there is some relationship between response rate when satisfied and when dissatisfied as well as with the age band. If we analyse the different response rates for each age band, we'll get the following output:

```
> rawData_SatScores %>% aov(RespRate_Sat_Num ~ AgeBand, data = .) %>% TukeyHSD()
  Tukey multiple comparisons of means
    95% family-wise confidence level
```

```
Fit: aov(formula = RespRate_Sat_Num ~ AgeBand, data = .)

$AgeBand
                                diff        lwr       upr     p adj
b. 28 to 34-a. 18 to 27 -0.209677419 -1.0286684 0.6093136 0.9768591
c. 35 to 40-a. 18 to 27 -0.214285714 -1.0422277 0.6136563 0.9757264
d. 41 to 52-a. 18 to 27  0.421568627 -0.3637910 1.2069283 0.6339014
e. 53 to 62-a. 18 to 27  0.112903226 -0.7060878 0.9318942 0.9987007
f. 63 to 72-a. 18 to 27  1.000000000  0.1473379 1.8526621 0.0114135
c. 35 to 40-b. 28 to 34 -0.004608295 -0.5430501 0.5338335 1.0000000
d. 41 to 52-b. 28 to 34  0.631246047  0.1609042 1.1015879 0.0021362
e. 53 to 62-b. 28 to 34  0.322580645 -0.2019934 0.8471547 0.4859050
f. 63 to 72-b. 28 to 34  1.209677419  0.6339480 1.7854068 0.0000001
d. 41 to 52-c. 35 to 40  0.635854342  0.1500939 1.1216147 0.0030139
e. 53 to 62-c. 35 to 40  0.327188940 -0.2112529 0.8656308 0.4995480
f. 63 to 72-c. 35 to 40  1.214285714  0.6258930 1.8026785 0.0000002
e. 53 to 62-d. 41 to 52 -0.308665402 -0.7790073 0.1616765 0.4105467
f. 63 to 72-d. 41 to 52  0.578431373  0.0516412 1.1052215 0.0223022
f. 63 to 72-e. 53 to 62  0.887096774  0.3113674 1.4628262 0.0002321

> rawData_SatScores %>%
+   group_by(AgeBand) %>%
+   summarise(
+     'Avg_RR' = mean(RespRate_Sat_Num, na.rm = TRUE)
+     ,'surveys' = n_distinct(Id)
+     ,'Percentage of surveys' = percent(n_distinct(Id) / TotalSurveys_SatScores,
accuracy = 0.1)
+   )
# A tibble: 6 x 4
  AgeBand      Avg_RR surveys `Percentage of surveys`
* <chr>         <dbl>   <int> <chr>
1 a. 18 to 27 -0.5          8 4.7%
2 b. 28 to 34 -0.710       31 18.1%
3 c. 35 to 40 -0.714       28 16.4%
4 d. 41 to 52 -0.0784      51 29.8%
5 e. 53 to 62 -0.387       31 18.1%
6 f. 63 to 72  0.5         22 12.9%
```

In which we can see that there significant differences between multiple of the age bands (especially the respondents with 63 years or more vs the rest of respondents), although there is no gradual evolution, fact that can be also visualized in figure 35.

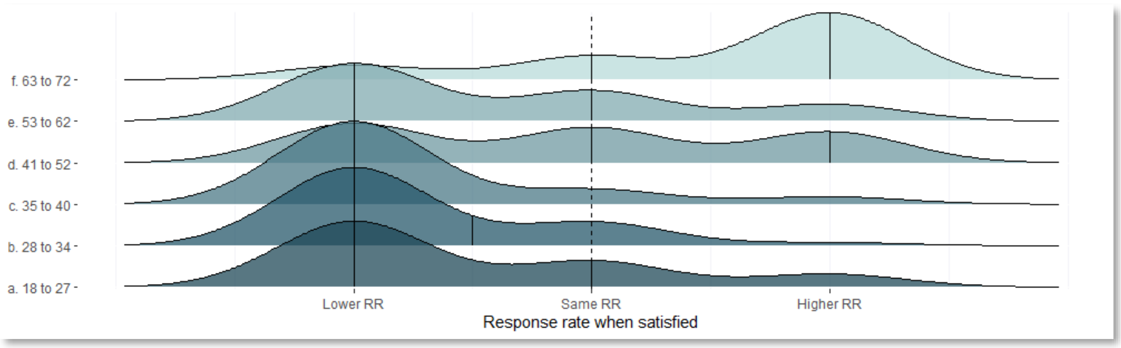

Fig. 35: Distribution of response rate when satisfied by respondent age.
Source: Self-generated from conducted survey.

RESPONSE RATE WHEN DISSATISFIED

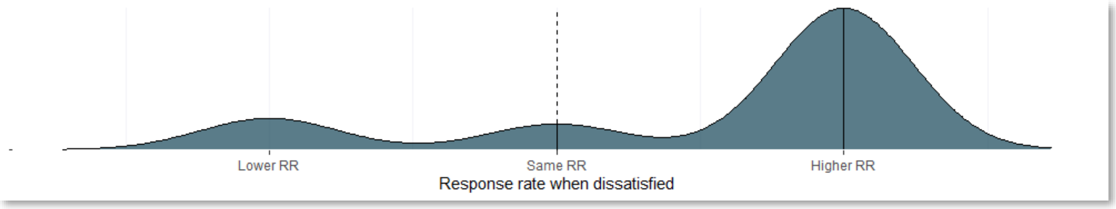

Fig. 36: Distribution of response rate when dissatisfied.
Source: Self-generated from conducted survey.

From the analysed dataset there are 3 variables that generate a different response rate when dissatisfied: response rate when satisfied, age band and education, although in the three cases their significance change when compared between all data or only the data for respondents with recent job experiences:

```
> ### RR when dissatisfied
> mean(rawData_SatScores$RespRate_Dissat_Num, na.rm = TRUE)
[1] 0.5555556
> rawData_SatScores %>% aov(RespRate_Dissat_Num ~ RespRate_Std + RespRate_Sat +
Gender_G + Education_G + AgeBand + Price_Top3 + Quality_Top3 + Convenience_Top3 +
PrevSatisf_Top3 + BuyingExp_Top3, data = .) %>% summary()
                  Df Sum Sq Mean Sq F value              Pr(>F)
RespRate_Std       4   0.93   0.232   0.667             0.61611
RespRate_Sat       2  30.67  15.334  44.061 0.00000000000000971 ***
Gender_G           2   0.95   0.473   1.359             0.26019
Education_G        4   2.94   0.736   2.114             0.08192 .
AgeBand            5   7.48   1.495   4.297             0.00112 **
Price_Top3         1   0.14   0.136   0.391             0.53262
Quality_Top3       1   0.03   0.026   0.076             0.78346
Convenience_Top3   1   0.06   0.064   0.185             0.66793
PrevSatisf_Top3    1   0.85   0.846   2.431             0.12111
BuyingExp_Top3     1   0.68   0.680   1.953             0.16436
Residuals        148  51.51   0.348
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> rawData_SatScores %>% aov(RespRate_Dissat_Num ~ RespRate_Std + RespRate_Sat +
Gender_G + Education_G + AgeBand + Price_Top3 + Quality_Top3 + Convenience_Top3 +
PrevSatisf_Top3 + BuyingExp_Top3 + Job_Customer + Job_Manager + Job_SatSurveysCons,
data = .) %>% summary()
                    Df Sum Sq Mean Sq F value      Pr(>F)
RespRate_Std         3   0.16   0.053   0.168      0.9175
RespRate_Sat         2  10.79   5.397  17.187 0.000000325 ***
Gender_G             2   0.66   0.329   1.049      0.3538
Education_G          4   3.79   0.947   3.016      0.0211 *
AgeBand              5   3.34   0.669   2.129      0.0673 .
Price_Top3           1   0.00   0.001   0.004      0.9516
Quality_Top3         1   0.01   0.010   0.033      0.8559
Convenience_Top3     1   0.15   0.148   0.470      0.4944
PrevSatisf_Top3      1   0.66   0.660   2.101      0.1501
BuyingExp_Top3       1   0.41   0.410   1.307      0.2555
Job_Customer         4   0.73   0.183   0.582      0.6760
Job_Manager          2   0.19   0.094   0.299      0.7420
Job_SatSurveysCons   2   0.19   0.097   0.310      0.7342
Residuals          109  34.23   0.314
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
32 observations deleted due to missingness
```

The case of education is particular because it has also been identified as a factor for the standard response rate. Combined, we can see how the standard response rate evolves as explained, while on the dissatisfied response rate, the higher the education the higher the response rate that can be confirmed with the following Pearson’s correlation results (fig. 37).

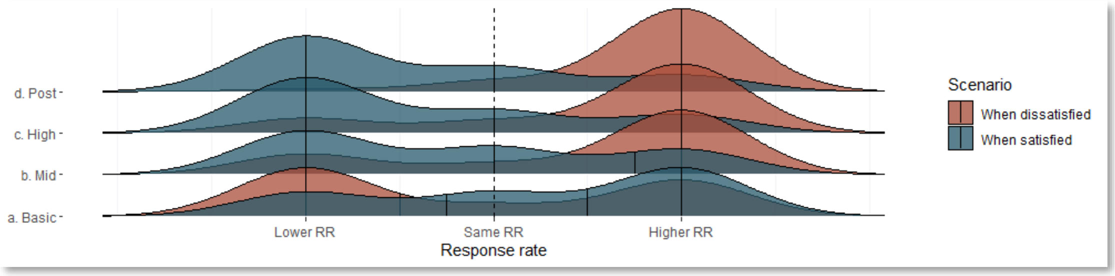

Fig. 37: Distribution of response rate by education level depending on the scenario: satisfied and dissatisfied.
Source: Self-generated from conducted survey.

```
> Edu_RR_D <- rawData_SatScores %>%
+   filter(Education_G != 'z. n/a') %>%
+   mutate(Education_G = case_when(
+     Education_G == 'a. Basic' ~ 1,
+     Education_G == 'b. Mid'   ~ 2,
+     Education_G == 'c. High'  ~ 3,
+     Education_G == 'd. Post'  ~ 4,
+     TRUE                      ~ 0
+   )) %>%
+   select(Education_G, RespRate_Dissat_Num)
> cor.test(Edu_RR_D$Education_G, Edu_RR_D$RespRate_Dissat_Num, method = 'pearson')

Pearson's product-moment correlation

data:  Edu_RR_D$Education_G and Edu_RR_D$RespRate_Dissat_Num
t = 3.2839, df = 167, p-value = 0.001247
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.09901254 0.38301090
sample estimates:
      cor
0.2462913
```

In regards to age band, we can see when we plot the results that there is one age band with a completely different behaviour when dissatisfied (fig. 38)

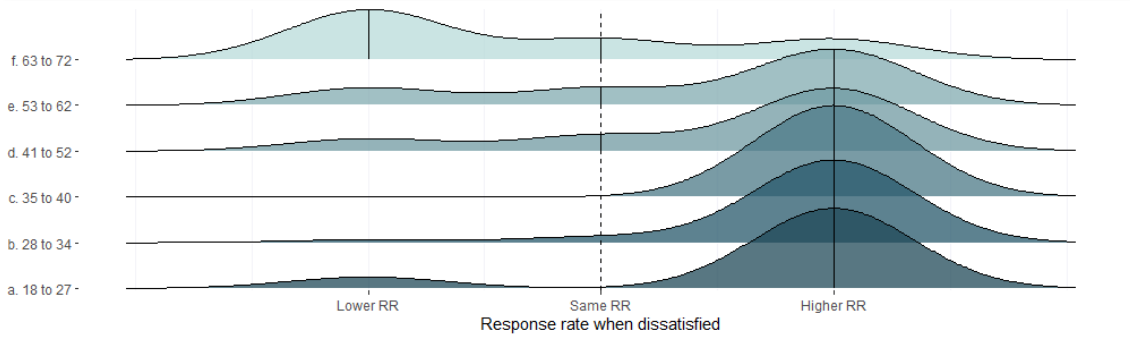

Fig. 38: Distribution of response rate when dissatisfied depending on respondent age.
Source: Self-generated from conducted survey.

In fact after consolidating the data into just 2 bands: 62 years or younger and 63 years or older the p-value of the factor gains significance:

```
> rawData_SatScores %>%
+   mutate(AgeBand = case_when(
+     AgeBand == 'f. 63 to 72' ~ 'b. 63 to 72',
+     TRUE                     ~ 'a. 18 to 62'
+   )) %>%
+   aov(RespRate_Dissat_Num ~ AgeBand, data = .) %>%
+   summary()
```

```
             Df Sum Sq Mean Sq F value         Pr(>F)
AgeBand       1  19.27  19.275   42.33 0.000000000838 ***
Residuals   169  76.95   0.455
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

ANALYSIS CONCLUSION

So we have identified 3 factors:

1. Education: the higher the education the lower the standard response rate but higher response rate when dissatisfied.
2. Age: customers of 63 years or more will have lower response rate when dissatisfied, all younger customers will have higher response rate when dissatisfied, and all will present the opposite behaviour when satisfied as can be seen in figure 39.

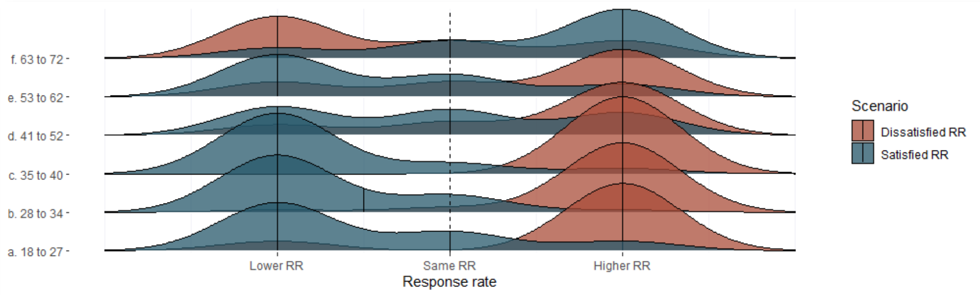

Fig. 39: Distribution of response rate by age depending on the scenario: satisfied and dissatisfied.
Source: Self-generated from conducted survey.

3. Interaction between response rates when satisfied or dissatisfied: A very interesting outcome from the analysis comes from comparing the response rates depending on the satisfaction as they have a very significant difference. A Pearson correlation analysis presents a strong negative correlation (-0.56 with a p-value of nearly 0) between them. Plotting the interaction heatmap, we get the following chart where we can clearly see that in general, people tend to have lower response rates when satisfied but higher response rates when they are dissatisfied: almost 50% of the respondents report this and the proportion reaches almost 65% when we compare higher response rate when dissatisfied and lower or similar response rate when satisfied (fig. 40).

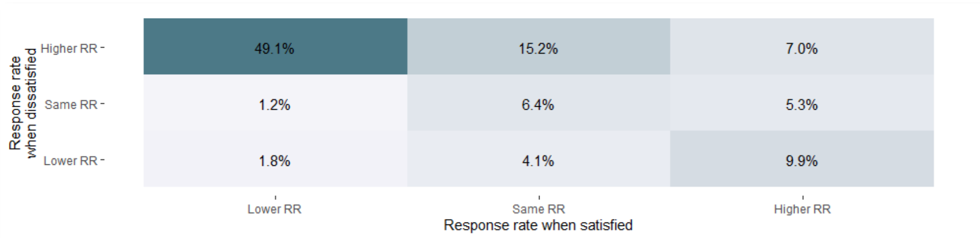

Fig. 40: Percentage of respondents on depending on their response rate when dissatisfied and satisfied.
Source: Self-generated from conducted survey.

From the conducted analysis we can then establish that any satisfaction measured will be lower than the actual customer satisfaction as the sample across the satisfaction levels is not going to be similar to the actual population of satisfied and dissatisfied customers.

Besides the response rate analysis, and reconnecting with the touchpoint type differences on scoring mindset identified before, the different touchpoints could be affecting the overall satisfaction because the automatic touchpoints may have even more extreme response rate (for instance, only answering when there is an issue, but not when it works as expected or above expectations).

“I use to score an after a human interaction, but never after an automatic interaction”

RETENTION MEASURE VS ACQUISITION MEASURES

Up to this point, it is important to be aware that in general terms all customer satisfaction metrics are reactive measurements that helps identifying the areas where there is a mismatch between expectations and performance according to the described theoretical definition of customer satisfaction. But at no point it gives any insight on the perceived worth of the products and/or services compared to the competition.

However, the customer satisfaction measures can provide 2 very interesting insights or high level indications:

1. Future purchase decision making: as mentioned earlier, previous experience is one of the key factors selected by the respondents for future decision making. Therefore, ensuring a high customer satisfaction will help ensuring that the customer considers the company for future purchases.
2. Quality of sales process: one of the key metrics to monitor to ensure a healthy sales-driven growth is through customer churn. At the end of the day, a high churn rate early in the customer journey in many occasions comes from aggressive sales or just inaccurate communications setting wrong expectations that are not met during the first periods of use. After that, the customer will probably leave. A complete customer satisfaction survey similar to the survey processes identified in the mature companies will help identifying, before the customer leaves, the areas where the dissatisfaction happens and lead the implementation of contingency measures.

MONITORING STRATEGIES

(IMPLEMENTING A CUSTOMER SATISFACTION MONITORING STRATEGY)

VERY EARLY STAGES: STARTUPS

A start-up is a very specific type of company, and each one has its own particularities. The timelines between the 3 stages identified in the following infographic may vary massively, but still the general concept would be similar and, therefore, the metrics to focus and to evaluate their customers' satisfaction would be the same. Customer experience and satisfaction is not a growth metric by itself (the quality of your service can help on the brand positioning, but customer satisfaction will not generate direct revenue) but an optimization metric helping to identify customer leakage (churn) even before it happens as well as collaborate with the different teams to ensure that the product-market fit is optimised.

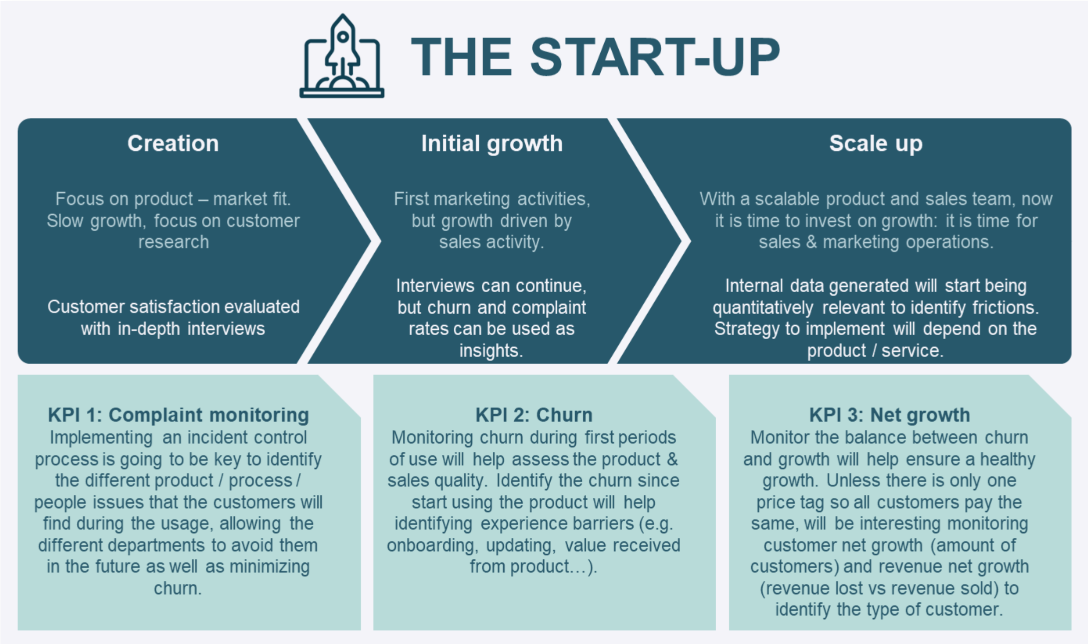

Fig. 41: Summary of approach to customer satisfaction and strategies identified on startups.
Source: Self-generated from interview feedback.

MATURE COMPANIES: THE TRANSACTION VS THE RELATIONSHIP

Companies with mature customer satisfaction monitoring can move forward to becoming a data driven customer centric organization. This type of companies usually already have transactional data monitoring and in many cases, as the two interviewed companies that can fit under this category do, also have some kind of relational data with lower frequency but higher detail.

Merging the two data sources into a consolidated CRM will allow transforming the transactional sales teams into account management teams that together with the support of a customer intelligence team will be able to support the customer-focused team (sales, support…) with predictive analytics.

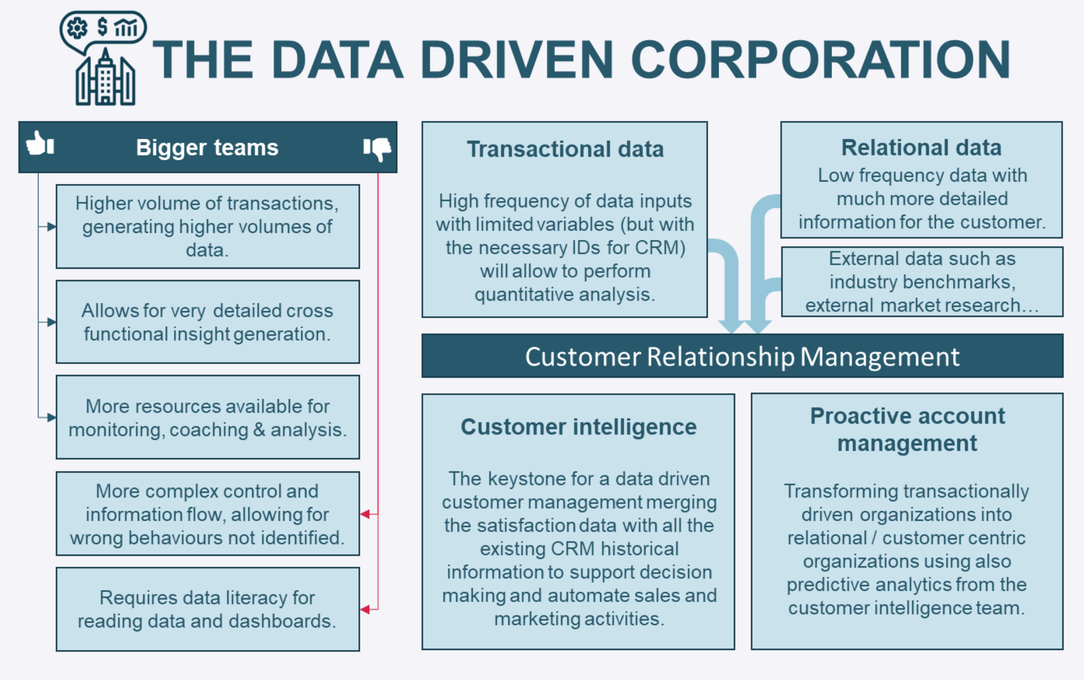

Fig. 42: Summary of approach to customer satisfaction and strategies identified on companies with mature customer monitoring systems.
Source: Self-generated from interview feedback.

THE OUTLIER: WHEN CUSTOMER SATISFACTION IS NOT MONITORED

Even though it may not seem the standard in the data world we are, there are still companies thriving that do not have any customer satisfaction monitoring in place. However, even in this type of organizations there is a proxy to measure similar purposes. From the interviews with one company falling under this category the main reason for not having direct customer satisfaction is a combination of 2 factors:

- Inefficient information flow: the area of business where they operate is defined by having several stakeholders that act as nodes but, at the same time, filter the information flow going back to the manufacturer.
- Extremely low volume of transaction per customer: the products served have very long life, therefore any transactional analysis that can deliver enough data volume will stop on the intermediaries, and the customers usually don't need (nor expect) this information flow.

> *"The communication followed always the same pattern than sales. In the case of issues, it went backwards. But except for big issues, they did not reach us and was solved (satisfactory or not) by the intermediaries."*
>
> *B2B Company without customer satisfaction process*

The infrastructure for complaint management exists, although it is can be absorbed by the account management team depending on the organization. New technologies are acting as an enabler of new information flows without forcing a disconnection from the stakeholders, key in their goods distribution process:

> *"There's a plan to launch an online channel that will act as a marketplace where the end user will make the order and it will be sent to the manufacturer and sold by them so they don't lose the commission."*
>
> *B2B Company without customer satisfaction process*

RISKS ON IMPLEMENTING METRICS

SPAMMING THE CUSTOMER TO EXHAUSTION

Currently there is a survey offered almost after any business transaction except for small retailers. That may tire the customers and create a negative effect on response rate (customers will only answer the survey when dissatisfied).

In fact, from the interviewed companies with monitoring systems in place, all of them had policies to limit the amount of surveys done by customer.

SAMPLE SIZE, COMPENSATION, PERFORMANCE MANAGEMENT & OUTLIERS

It is critical to be very careful using the data, specially when there is a low volume of surveys. On big companies that may not be an issue on top level, but can jeopardize the intention of the metric if used for compensation or strict performance management with very low sample rates. The interviewed companies reported having that in mind and either not using customer satisfaction data on the individuals compensation or considering it only for the annual compensation or calculated at team / region level.

Another common issue occurs when visualizing data to leadership without context and presenting, for instance, a drop of 20 points on NPS but with 10 surveys each one of the periods. On this instance, besides adding quantifiable information measuring already known customer satisfaction levers, there are 2 best practices that can help mitigate the concerns:

1. Work with longer periods or higher level of aggregation
2. When visualizing the satisfaction score, include the volume of surveys and the significance of the change.

METRIC AS A GOAL

At the end of the day, the metric is and needs to be a representation of the customers' satisfaction and the goal is and needs to be to have the best possible customer experience. The metric is the indicator that helps visualizing an insight that otherwise may be extremely difficult to assess.

However, on many occasions and during the day to day operations, the goal of focusing on customer satisfaction is replaced by the metric, which generates friction between teams, increases misalignments…

CONCLUSIONS

GOAL RECAP

Goal #1: Identify if there is any impact on the satisfaction scores that a soft guide on the survey (colour scale).

Based on the analysis performed, and considering the limitations, we can see that there seems to be no impact on the scoring. However, the limitations of the dataset may not allow us to confirm whether it is due to the data used for the analysis or if there is no difference between the two survey models.

Goal #2: Identify if there are differences on response rate depending on the satisfaction or any other factor.

There are 2 factors that generate alterations on the response rate: education and age.

In general terms, the higher the education the lower the response rate. This pattern is inverted when the customer is dissatisfied. On that instance, the higher educated respondents tend to have higher response rate when dissatisfied than the lower educated ones.

The age is also an important factor that has opposite impacts on satisfied and dissatisfied response rates. Customers with 63 years or more will have higher response rates when satisfied and lower when dissatisfied, while customers up to 62 years will behave opposite: lower response rates when satisfied and higher when dissatisfied.

Besides this, there is a general pattern that will always skew the satisfaction score towards lower results than the real satisfaction: a big majority of the customers tend to answer with lower frequency when satisfied but increase their response rate when dissatisfied.

Goal #3: Identify if customer satisfaction (previous satisfaction) has an influence on future purchase decision making.

Previous satisfaction is put as top factor on 16% of the surveys (the 2^{nd} factor with highest frequency on top) and in 49% is within the top 3 factors. That correlates with previous research findings on modelling future purchase decision making based on customer satisfaction metrics.

Goal #4: Identify if the different decision making factors generate different satisfaction scoring

When the customer is satisfied there seems to be no factor modifying the results.

When the customer is neither satisfied or dissatisfied, there are 4 factors influencing the score:

- When the customer works on a job connected to satisfaction surveys, tends to give higher scores.
- When the customer considers price or the aftersales an important factor for future purchase decision making tends to give a lower score.
- There is a weak although significant negative correlation between age and satisfaction score.

When the customer is dissatisfied there is one factor influencing on the score, which is the response rate when dissatisfied: the customers that have a higher response rate when dissatisfied also tend to give lower scores.

Goal #5: Identify if the mindset towards scoring after a human interaction is the same than when scoring an automatic interaction (expectations and response rate)

When asked about how do they evaluate human interactions in compared to automatic interactions, the vast majority of respondents tend to score more benevolently human interactions.

Goal #6: Identify what would be the approach to monitor customer satisfaction on 3 scenarios: start-up, mature companies with already a monitoring system and mature companies with no monitoring customer satisfaction.

Start-ups are focused on product and growth. On this instance the best proxy would be to monitor the amount of complaints, customer churn and customer and revenue net growth. Once a critical volume of business is reached, then a team of customer success and/or customer intelligence (depending on the acquisition and sales process) can be created to optimize customer experience and lifetime value.

For already scaled up companies, there are two possible customer satisfaction data sources: transactional surveys (high frequency and very focused on specific areas) and relational surveys (lower frequency but broader scope). The two consolidated with a strong customer intelligence team can generate synergies across teams as well as implement predictive analytics.

For mature companies without customer satisfaction monitoring system in place, the best approach would be monitoring, if possible and relevant, the same areas than the start-ups and implement very efficient and fluent communications channels with the different stakeholders as well as with their customers.

GENERAL CONCLUSIONS

In general terms, customer satisfaction monitoring is a mean to an end, which is ensure that the customers are happy with the overall experience with the company and the product/service so they will want to re-purchase and will recommend the brand, product or service to their friends. But it is not an end on itself.

Usually when you define a goal on a company, even when you also create a well defined process to achieve it, there will be people working outside of the process trying to find a better way to achieve the goal and, as part of human behaviour, there will also some staff members that will try to “hack” the system to get to the goal at the expense of other results.

This behaviour is one of the top dangers on implementing a customer satisfaction metric: any shortcut (well-meaning or not) to get the highest possible score has the potential to end up damaging the experience of the customer (in the short or long run) and, specially, will mean that the leadership team will have lower quality data to help them make the best possible decision.

As seen in the research, customer satisfaction is an approximation of satisfaction but is skewed naturally towards lower scores due to the differences in the response rate and different priorities (price, experience, quality…) will result in different satisfactions.

Any customer satisfaction metric based on the customer being surveyed needs to be monitored with a wide range of data that can be internal such as customer profile, CRM data or purchase history but also external, if possible, such as socio-demographic, previous purchase, navigation information… (information available not only on ecommerce but also on street retailers with tools like Kelvin Retail, a customer insights tool from Banc Sabadell). With all this data, the information gathered with the customer satisfaction monitoring will be put in context and only then can be transformed by a customer intelligence team (a team with wide knowledge of the business, processes and customer base that is able to apply advanced statistical methodologies) into actionable information.

Focusing exclusively on customer satisfaction without adding contextual information can be as dangerous as just focusing on customer satisfaction as an independent metric, and will end up transforming the metric on an end instead of a mean.

LIMITATIONS AND FURTHER ANALYSIS OPPORTUNITIES IDENTIFIED

LIMITED SAMPLE SIZE

The time limitations and evolution of the research has derived in a relatively small sample size limited to ~85 responses on each survey type (to compare the results between survey type) and ~170 in total.

On top of that, the data has forced to trim the oldest age band due to very limited data and has slightly underrepresented the youngest age band (although in this case with no observed statistical impact on the results).

A/B TESTING BASED ON FICTIONAL SCENARIOS

Trying to assess the influence of guiding the customer on the survey, and added to the previous point, presents a weak point: it is based on fictional scenarios, for each satisfaction level the respondents had to imagine what would have they scored given the description.

Considering this my recommendation would be to repeat the research having a bigger response rate with real interactions (probably best in collaboration with as many companies as possible so it can be done across multiple industries and countries).

ILLOGIC RESPONSES (satisfied & low score, dissatisfied & high score)

On the non-guided surveys there are a few responses that score the opposite from what the scenario would describe. In this case a follow up cannot be done as the survey was completely anonymous, but it would be interesting to identify the reason for the score.

If an extended research (like the one suggested in the previous point) confirms that a soft-guide (like a colour scale) does not influence on the outcome of the survey and, with wider figures, it is confirmed that the illogic responses only occur in the non-guided models, the recommendation would be to make small changes on the user interface so it helps avoiding this type of errors.

PREDICTIVE & PRESCRIPTIVE ANALYSIS

The current dataset has been unable to find any model that would help predict:

- Factors of repurchasing: identifying the factors of repurchasing for each customer would help tailoring the sales process and customer experience with the product and/or service to what is relevant to them. That could help increasing the repurchase behaviour and increasing the CLV by increasing the purchase value and / or increasing the amount of repurchases.
- Neutral satisfaction scoring: identifying in advance the customers that may have neutral satisfaction scores is critical from a performance indicator standpoint but also from a loyalty point of view. May times, a neutral customer can be turned into a satisfied customer by just slightly adapting the message and experience towards what is relevant to them, converting a neutral or passive customer into a promoter. But also has been mentioned earlier that as promoters tend to be loyal, neutral customers have lower loyalty, so identifying the customers that are not promoters (but close) will help increase customer loyalty.

IMPACT OF AUTOMATION

One of the areas mentioned on the in-depth interviews is that while the company is increasing the amount of automatic touchpoints, they are finding a negative correlation between the proportion of automated touchpoints and the satisfaction score.

On the surveys the respondents were asked about how do they score human touchpoints vs automated touchpoints and, in general, the tendency is to be more benevolent (less harsher) when evaluating a human interaction than when evaluating an automatic interaction. However, with the increase of chatbots and what the advances on AI and deep learning may bring, being able to assess and identify the expectations from each type of transactions will help on improving the customer's experience.

EPILOGUE

WORK ASSESSMENT

I choose the area of research because I have seen many malpractices on customer satisfaction surveys and data usage as an employee, as a manager, as an analyst and as a customer that made me guess what was the point on measuring it. This exercise allowed me not only to investigate and deepen my knowledge on the different customer satisfaction methodologies and implications for the business, but also to apply multiple of the learnings from different subjects studied throughout the degree.

Completing this exercise has been a big challenge especially given the time constraints from combining it with other subjects and a demanding full time job, I have also reached at some points some level of *“analysis paralysis”* that made me re-write more than twice the more than one thousand lines of code used for the quantitative analysis or entire sections of this document.

Still, even though the delays on launching the quantitative survey and the limited sample gathered made the quantitative output not as strong as I wanted, I am still proud of the work done, analysis performed and hours invested into it.

ACKNOWLEDGEMENTS

I would not like to end the document without thanking all the companies that agreed to collaborate and the interviewees for making the process and the conversations so easy as well as everyone who collaborated on spreading the survey reaching the 200 surveys answered in just a handful of days.

I would like to thank my tutor, Santi Ariste, for the guidance, support, patience and quick response throughout the development of the work.

And I would like to specially thank my wife and my family (the ones that are here and the one that is *there*) for supporting me not only during this tiresome semester (*gràcies!*) but always and unconditionally, helping me become who I am now.

BIBLIOGRAPHY

Airbnb Locations. (n.d.). airbnb.com. Retrieved June 2021, from https://www.airbnb.com/s/homes?ne_lat=46.6859437858168&ne_lng=14.205853843921147&sw_lat=33.96926220984778&sw_lng=-8.272173499828853&room_types%5B%5D=Entire%20home

ALMOHAIMEED, B. (2019). Pillars of customer retention: an empirical study on the influence of customer satisfaction, customer loyalty, customer profitability on customer retention. *Serbian Journal of Management*, 421-435.

Amazon Store screenshot. (n.d.). amazon.es. Retrieved June 2021, from https://www.amazon.es/kindle-ahora-con-luz-frontal-integrada-negro/dp/B07FQ4XCR1/

ANDERSON, J. C., & NARUS, J. A. (1998). *Business marketing: understand what customers value.* Retrieved from Harvard Business Review: https://hbr.org/1998/11/business-marketing-understand-what-customers-value

ANDERSON, N. (2020, Jan 20). *Is it time we move beyond the NPS to measure user experience?* Retrieved from UX Collective (Medium): https://uxdesign.cc/the-nps-sucks-11d33b748774

BEST, R. J. (2007). *Marketing estratégico.* Madrid: Pearson Educación.

DAY, R. L. (1984). Modelling choices among alternative responses to dissatisfaction. *NA - Advances in Consumer Research Volume 11*, 496-499.

Delta Blade Consultants. (2019, Apr 05). *Customer satisfaction measurement: crucial, not comprehensive.* Retrieved from Delta Blade Consultants: https://www.deltablade.co.uk/customer-satisfaction/customer-satisfaction-measurement-crucial-not-comprehensive/

DIEBNER, R., MALFARA, D., NEHER, K., THOMPSON, M., & VANCAUWENBERGHE, M. (2021, Feb 24). *Prediction: The future of CX.* Retrieved from McKinsey: https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/prediction-the-future-of-cx

DIXON, M., FREEMAN, K., & TOMAN, N. (2010). Stop trying to delight your customers. *Harvard business review (88)*. Retrieved from https://hbr.org/2010/07/stop-trying-to-delight-your-customers

DUNN, B. (2003, 08). *Loyalty myth #8: It costs five times more to acquire a customer than to retain a customer.* Retrieved from Ipsos Loyalty: https://www.ipsos.com/sites/default/files/publication/2003-08/Ipsos_Loyalty_Myth_8_Excerpt.pdf

FISHER, N. I., & KORDUPLESKI, R. E. (2019). Good and bad market research: a critical review of Net promoter score. *Applied Stochastic Models in Business & Industry, 35 (1)*, 138-151.

FORNELL, C., JOHNSON, M. D., ANDERSON, E. W., CHA, J., & BRYANT, B. E. (1996). The American Customer Satisfaction Index: Nature, purpose and findings. *Journal of Marketing, Vol 60 (4)*, 7-18.

Google Maps Screenshot. (n.d.). Google. Retrieved June 2021, from https://www.google.com/maps

GURAU, C., & RANCHHOD, A. (2002). Measuring customer satisfaction: a platform for calculating predicting and increasing customer profitability. *Journal of targeting, measurement and analysis of marketing, 10 (3)*, 203-219.

GUSTAFSSON, A., JOHNSON, M. D., & ROOS, I. (2005). The Effects of Customer Satisfaction, Relationship Commitment Dimensions, and Triggers on Customer Retention. *Journal of Marketing, 69 (4)*, 210-218.

HAUMANN, T., QUAISER, B., WISEKE, J., & RESE, M. (2014). Footprints in the Sands of Time: A Comparative Analysis of the Effectiveness of Customer Satisfaction and Customer–Company Identification over Time. *Journal of Marketing, 78 (6)*, 78-102.

HESKETT, J. L., SASSER, W. E., & SCHLESINGER, L. A. (1997). The Service Profit Chain: How Leading Companies Link Profit and Growth to Loyalty, Satisfaction, and Value. New York: Free Press.

HO, E. (2018, Jun 16). *Stop looking at your Net Promoter Score! A better way to utilize NPS.* Retrieved from User Interviews (Medium): https://medium.com/userinterviews/stop-just-looking-at-your-net-promoter-score-a-better-way-to-utilize-nps-1ed97df4bbc3

HUANG, P.-L., LEE, B. C., & CHEN, C.-C. (2019). he influence of service quality on customer satisfaction and loyalty in B2B technology service industry. *Total Quality Management, 20 (13)*, 1449-1469.

KANNING, U. P., & BERGMANN, N. (2009). Predictors of customer satisfaction: Testing the classical paradigms. *Managing Service Quality, 19(4)*, 377-390.

KATSIKEAS, C. S., MORGAN, N. A., LEONDIOU, L. C., & HULT, G. M. (2016). Assessing Performance Outcomes in Marketing. *Journal of Marketing, 80 (2)*, 1-20.

LIM, L. G., TULI, K. R., & GREWAL, R. (2020). Customer satisfaction and its impact on the future costs of selling. *American Marketing Association, Vol 84 (4)*, 23-44.

LOCKE, E. A. (1969). What is job satisfaction? *Oganizational Behavior and Human Performance, vol 4*, 309-336.

MARSHALL, N. W. (2010). Commitment, loyalty and customer lifetime value: investigating the relationships among key determinants. *Journal of Business & Economics Research, 8 (8)*, 67-84.

QI, J.-Y., ZHOU, Y.-P., CHEN, W.-J., & QU, Q.-X. (2012). Are customer satisfaction and customer loyalty drivers of customer lifetime value in mobile data services: a comparative cross-country study. *Information Technology and Management, 13*, 281-296. Retrieved from Information Technology and Management.

REICHHELD, F. F. (2004). The one number you need to grow. *Harvard business review*, 46-54. Retrieved from https://hbr.org/2003/12/the-one-number-you-need-to-grow

REICHHELD, F. F., & SASSER, W. E. (1990). Zero defections: Quality comes to service. *Harvard Business Review, 68 (5)*, 105-111.

REICHHELD, F. F., MARKEY, R. G., & HOPTON, C. (2000). E-customer loyalty - applying the traditional rules of business for online success. *European business journal, Vol 12, No 4*, 173-179.

REICHHELD, F. F., MARKEY, R. G., & HOPTON, C. (2000). The loyalty effect - the relationship between loyalty and profits. *European Business Journal*, 134-139.

TSE, D. K., & WILTON, P. C. (1984). History and future of consumer satisfaction research. *Historical Perspective in Consumer Research: National and International Perspectives*, 251-256.

VAN DOORN, J., & VERHOEF, P. C. (2008). Critical Incidents and the Impact of Satisfaction on Customer Share. *Journal of Marketing, 72 (July)*, 123-142.

VOSS, G. B., GODFREY, A., & SEIDERS, K. (2010). How Complementarity and Substitution Alter the Customer Satisfaction–Repurchase Link. *Journal of Marketing, 74 (6)*, 111-127.

WIESEKE, J., ALAVI, S., & HABEL, J. (2014). Willing to Pay More, Eager to Pay Less: The Role of Customer Loyalty in Price Negotiations. *Journal of Marketing, 78 (6)*, 17-37.

Zwift NPS email survey. (n.d.). zwift.com. Retrieved June 2021, from https://gozwift.typeform.com/to/OMC6R9