Ehrenberg-Bass Institute **Working Paper**:

The Net Promoter Score: What Should Managers Know?

This working paper, dated 12 July 2023, is forthcoming in the ***International Journal of Market Research.***

Author:

Professor John Dawes - Ehrenberg-Bass Institute

University of South Australia | Ehrenberg-Bass Institute for Marketing Science

The Net Promoter Score: What Should Managers Know?

John G. Dawes
Ehrenberg-Bass Institute for Marketing Science
University of South Australia
Australia
Email John.Dawes@marketingscience.info

Conflict of Interest
The author declares that there is no conflict of interest.

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

The Net Promoter Score: What Should Managers Know?

Abstract

The Net Promoter Score is a popular management tool that is used in a variety of ways by firms, not-for-profits, and government. This study firstly provides an overview of the various ways in which the NPS is used. It then canvasses four concerns raised by researchers, authors and commentators about the NPS. These relate to (1) its presumed link to business growth, (2) the assumption that low NPS scores indicate negative word of mouth, (3) the weak association between stated likelihood to recommend and actual recommending, and (4) the claim that NPS is a superior metric to customer satisfaction. The evidence pertaining to those concerns is examined. The study then discusses another problem with the NPS that many practitioners are aware of, but has not been studied. The problem is that the counting method used to calculate the NPS introduces additional variation in scores compared to mean average likelihood-to-recommend scores. This additional variation occurs both across brands in a study, as well as for the same brand over survey waves. This variation is likely to be difficult for market research providers, or those who commission NPS work, to explain. The study concludes with alternative courses of action for NPS users.

Keywords: Net Promoter, Customer Satisfaction, Surveys.

The Net Promoter Score: What Should Managers Know?

Introduction

The Net Promoter Score (NPS) is widely used, generally as a word-of-mouth indicator or as a proxy for customer satisfaction. Thousands of businesses and brands use the metric (Colvin 2020), aided by survey providers and consultants. However, the NPS has received considerable criticism. The aim of this paper is to firstly, provide an overview of how the NPS is used in industry, then to discuss and critique a range of viewpoints on the NPS. The paper also identifies a previously undocumented shortcoming of the NPS, namely that the counting method to derive 'Net' scores (i.e. 'promoters minus 'detractors') increases the variation in scores between brands, and from survey to survey for the same brand.

The uses of the NPS

As a 'headline' performance metric

Businesses use the NPS in a variety of ways. First, it is used as a headline-level performance metric to complement financial results. For example, the US airline Jet Blue began reporting the NPS in its annual reports in 2008 (Jet Blue 2008), along with scores for staff confidence in management, customer complaints and departure times. Another US airline, Southwest, began tracking NPS in approximately 2014 and included it in annual reports alongside non-financial metrics such as safety performance from 2015 (Southwest Airlines 2015). Top management bonuses in these corporations were linked to performance on NPS and those other metrics. In other countries such as the UK, for example, several of the largest banks have also commenced reporting NPS in annual reports. Barclays started reporting NPS in 2013, however, only reports its ranking relative to peers, not the actual score (Barclays 2013). A similar approach is taken by

HSBC (HSBC Holdings plc 2022). Natwest (2021) reported NPS scores in its annual report for 2020 - including the fact they had declined over several years. Overall, companies such as these appear to have included the NPS in their annual reports to (a) signal to stakeholders the business is committed to being customer-oriented, (b) to show that the company enjoys positive sentiment among customers, or has a plan to improve that sentiment; and (c) demonstrate its performance or improvement in a customer-based metric, often as part of a managerial bonus plan.

Frontline staff monitoring

Another use of the NPS by business is as a tool to monitor and motivate frontline staff performance. For example, customers of a business are invited to complete an NPS survey, often following a service encounter. Research providers refer to this as transactional NPS (Owen & Brooks 2008) and distinguish it from long-term or relationship NPS.

It is common practice to issue these transactional survey invitations via email, text or a mobile app, with the potential benefit of quick feedback (Alismail *et al.* 2020). This feedback can allow timely management intervention if results depart from the firm's service standards. The rationale is that quick remediation of customer problems, and the origin of those problems, should improve customer sentiment. Moreover, if frontline staff know their customers will be surveyed, they have an incentive to perform well. In addition, giving frontline staff an easily understood metric that is relevant to their job is said to be motivating (Seifert & Yukl 2010).

As a component of customer management systems

NPS results (along with related customer service metrics) are input into reporting systems, so that managers can track performance at frequent intervals or even in real-time along with other metrics such as the proportion of customer problems resolved in one contact. Moreover, NPS scores can be split out by product, staff member or team, and by touchpoint, along with the reasons for high or low scores (Qualtrics 2022). Therefore, NPS has evolved from a survey question to be a lead component in customer management systems. In relation to using NPS in routinised surveys relating to the performance of frontline staff, some commentators express concern they can result in staff "begging for good scores" (e.g. Ramshaw 2014). Of course, this problem is not unique to NPS, it can also occur for satisfaction or other types of customer experience surveys.

Related to the use of NPS as a management metric is the arguably dramatic nature of NPS scores. The NPS ranges from -100 to 100, while satisfaction scores are often reported as mean scores out of 7 or 10, or as the proportion of 'satisfied / highly satisfied' which is typically a figure of 50% or more. However, NPS scores can be far lower, and can therefore be attention-grabbing. Owen (2019) illustrates this facet of the NPS with an anecdote: "Everyone in the management meeting wakes up when they are told they scored 10 out of a possible 100" (p.58). Therefore, a reason to adopt or use NPS appears to be that it can produce scores that are provocative for the business, therefore more likely to get managerial attention and prompt behaviour change in organisations. Related to this point is that the NPS is seen as presenting a balance between positive and negative (promoters, minus detractors) that is more striking than using mean average scores. However, as we shall see, the attention-getting nature of the NPS comes at the cost of high variation in scores.

NPS is also used in reporting systems to try to diagnose what aspects of a firm's operations 'drive' positive or negative likelihood to recommend. As has been the case for many years with customer satisfaction research, firms attempt to model the reasons for high or low scores using variables such as staff knowledge, staff attitudes, time taken to resolve problems, or communications from the provider (e.g. Ho 2018). The rationale is to help managers know what aspects of the product or service are most important to obtaining high NPS scores (or avoiding low scores).

As an indicator of advocacy, satisfaction, and loyalty.

The NPS question asks about recommendation, but NPS scores are used to indicate other related outcomes for firms, such as satisfaction. To start off with, though, NPS scores are certainly used to indicate the level of advocacy for a firm, in line with Reichheld's prescription (Reichheld 2003). For example, the largest UK retail bank Barclays noted in its 2013 annual report,

> "[NPS reflects] our objective of moving from satisfied customers to customers who are strong advocates of our business" (Barclays 2013).

Similarly, adidas management stated

> "NPS has become an important part of the adidas brand's advocacy program … understand … key drivers which motivate them [buyers] to recommend the brand to their friends (adidas 2019).

The act of advocacy, or willingness to engage in it, is widely interpreted as a facet of customer *loyalty*. Indeed, Reichheld (2003) said that referrals were "the strongest sign of customer loyalty" (p. 48). As an example of this interpretation by firms, Jet Blue stated

> "NPS is a non-financial measure that assesses brand loyalty based on a customer's subjective survey responses to a customer experience" (Jet Blue 2015 p. 42).

It is problematic to interpret two concepts, such as loyalty and recommendation, from the same measure. We cannot assume a customer is loyal from their recommendation score, or that they will recommend given their level of loyalty. For instance, a low NPS score does not necessarily indicate low loyalty in terms of repeat business or tenure. A client of a firm might choose not to recommend it, even though they are otherwise highly loyal, because the client might consider the firm's services are not appropriate for friends or colleagues. Similarly, a client might only have just commenced dealing with a firm, therefore could not be thought of as necessarily loyal in the behavioral sense. However, they could give a high likelihood to recommend because they are, so far, very pleased with the provider. Lastly, loyalty is influenced by factors such as switching costs (Jones *et al.* 2002) whereas recommendation is not. Therefore, firms should distinguish between NPS scores, and metrics more specifically developed for customer loyalty (Farris *et al.* 2016) such as share of requirements and tenure.

Next, while Reichheld (2003) described the NPS as a distinct and superior metric to customer satisfaction, many organisations use NPS as an indicator or proxy for satisfaction. In the UK, satisfaction levels with banks are publicly reported by a government agency, yet respondents are actually asked the Net Promoter question (Competition and Markets Authority 2020). Another example from the UK is health care. In response to an enquiry that identified poor standards of care, the government introduced patient satisfaction surveys (Department of Health and Social Care 2012). However, as was the case with banks, the question asked pertains to recommendation, not satisfaction, and uses the NPS scoring method (NHS 2020). Other examples of the NPS in healthcare are discussed in Stirling and Jenkins (2019) and Hamilton et al (2014). Despite its use in healthcare, critics note that the NPS question suffers shortcomings in such settings. One is that many medical patients have limited choice, therefore the concept of providing recommendations about a particular provider to another person could be confusing (Customer Faithful 2022). Adams et al (2022) found in a systematic review that there were several conditions that made the NPS less useful in healthcare, including score variation across different medical procedures and age groups.

Benchmarking

Several companies provide NPS benchmarks to allow one to compare their own NPS performance to competitors in their industry. For example, one major provider publishes excerpts of its own NPS tracking to show top performing firms in sectors such as banking, airlines, insurance and retail. Detailed results are made available for a fee. An important rationale for benchmarking is that NPS scores do appear to systematically differ across industry sectors. For example, US department stores score on average 62 on the NPS, hotels 40, and health insurance providers only 13 (Satmetrix 2018). Therefore, benchmarking NPS may help companies understand what level of scores are realistic or achievable in their industry, as well as to set informed targets for improvement.

Modelling future behavior

Some firms monitor the future behavior of customers based on their stated NPS score. For example, AirBnB surveys customers to obtain feedback on factors such as property cleanliness and accuracy of property description. It added the NPS in 2013 (Qian 2015). The goal was to determine if NPS added further insight into customers' likelihood of re-booking. Interestingly, it found "higher NPS does in general correspond to more referrals and re-bookings. But we find that controlling for other factors, it does not significantly improve our ability to predict if a guest will book on Airbnb again in the next year" (2015, no page number). In terms of actual referral behavior, the study found people who gave a high NPS score had a *slightly* higher rate of actually referring: promoters did so at a rate only 4% higher than detractors. AirBnB's experience was also noteworthy in that it found the conventional approach of classifying customers as either promoters or detractors was sub-optimal. It concluded that combining the 9 and 10 scores together (the 'promoter' classification as per Reichheld) resulted in information loss (Qian 2015), and that greater insight was gained from preserving the original scores rather than aggregating. This result mirrors prior criticism of the NPS counting method (e.g. Grisaffe 2007).

AirBnB provides an example of endeavoring to link respondent's NPS to their future behavior. Relatedly, some industry providers suggest that NPS can be used to identify at-risk customers, meaning that their NPS will provide an indication of their likelihood of defecting (e.g. Perceptive 2022). Customers who give low NPS scores can therefore be the subject of interventions such as apologies, remediation, or incentives to keep their business with the firm. That said, there are alternative questions more specifically designed to measure the probability of customer defection, such as the Juster Scale or the related Verbal Probability Scale (Parackal & Brennan 1998; Hoek & Gendall 2010).

As a mindset metric?

Baehre and co-authors (Baehre *et al.* 2022; Baehre *et al.* 2021) report that the NPS can be used as a 'mindset' metric, in other words, can be used to ascertain sentiment towards a brand among non-customers as well as customers. However, the principal commercial survey providers that offer NPS still emphasise that NPS is primarily a mechanism to obtain feedback from customers, not from non-customers. Researchers should be cautious of asking non-customers their likelihood to recommend a brand. The reason is because if respondents are not familiar with the brand, they will probably give it a low score, not because they have negative sentiment about it, but because they do not feel informed enough about the brand to recommend it. One could incorrectly conclude the brand has many 'detractors,' but the results could merely reflect low brand awareness.

In summary, the NPS is used in a wide variety of ways. It is used as a corporate Key Performance Indicator and as a signal to stakeholders that the firm takes customer sentiment seriously. It is also a component of managerial dashboards, is interpreted as a metric not only for advocacy but also satisfaction and loyalty. The NPS is used to manage or motivate service staff, and some firms use the metric in attempts to estimate individual customer's future behavior. It is used by businesses and government, and in diverse settings including education (Kara *et al.* 2022) and healthcare. We now discuss a series of criticisms levelled at the NPS.

Criticisms of the NPS

NPS linked to revenue growth?

NPS was introduced and promoted by Reichheld (2003) as 'the one number you need to grow'. That is, achieving high NPS scores will underpin future business growth. However, examples used to support this claim often show that the growth occurred before, or partly before the high NPS, not afterwards. Keiningham et al (2007b) cites a study by the Listening Company (the study itself is no longer publicly available) that reported NPS was positively correlated with firm growth, but the growth figures were for the period preceding the NPS scores, not after them. Keiningham et al (2007b) also note that Reichheld's 2003 work, and a Satmetrix report in 2004 in support of the NPS, used past growth figures, concluding "This means that Net Promoter was tied to past growth rates (as opposed to future growth rates)" (p. 42). They also examined the NPS' correlation to business growth compared to customer satisfaction and found it was not a superior measure to satisfaction. Pingitore, Morgan, Rego, Gigliotti and Meyers (2007) analysed NPS against revenue growth for rental cars and airlines, but their revenue figures also preceded the NPS.

Many other studies have examined how NPS scores relate to revenue growth, with mixed results. Morgan and Rego (2006) used an NPS-like metric based on word of mouth and complaints, and found it did not correlate positively with sales growth figures derived from COMPUSTAT (a large-scale database of financial information for US firms). Keiningham, Cooil, Andreasen and

Aksoy (2007b) examined the link between NPS and revenue across five industries. The overall average correlation from the study is r=0.15, which is considered weak by social science standards (Cohen & Cohen 1983). Soon after the Keiningham et al study, a book by Owen and Brooks (2008) reported much higher correlations between NPS and subsequent revenue growth in two industries. The authors were associated with Satmetrix, which has NPS data for dozens of industries, therefore the rationale for showing only two is unclear. A broader study by Sauro (2018, 2019) reported an average correlation of approximately 0.60 between NPS and revenue growth over a subsequent two years, across multiple industries. At face value these two studies do support Reichheld's claim that NPS is related to future growth. It is, however, possible that revenue growth (or decline) for firms tends to persist over time, that is, high growth in one year may mean a firm is more likely to grow again the next year, or that decline in one year means it is more likely that the next year will be similar. There is some evidence that firm growth does follow such a pattern (Coad *et al.* 2018 p. 67; Chandra & Ro 2008). Therefore, it could be the case that high *current* growth is associated with high *current* NPS scores, and then an element of persistence in revenue growth makes it appear that the NPS score 'drives' future growth. Consistent with this idea is a finding by Mecredy, Wright and Feetham (2018) that examined NPS and growth in a B2B context in New Zealand. They found that the NPS positively correlated with past, present *and* future revenue.

At a similar time to the Mecredy et al study, Fiserova, Pugh, Dimos and Stephenson (2018) found evidence that NPS scores were positively associated with sales growth among the branches of a UK furniture retailer. This study is more supportive of a positive NPS-revenue link. However, several other studies present negative evidence. Zaki et al (2016) analyzed a large B2B firm's clientele and found NPS performed very poorly at predicting clients who would be loyal or not in a future year. Korneta (2018) in a study of Polish B2B firms found a positive link between NPS and profit, but not between NPS and revenue growth. Another study, by Farooq et al (2019) found a negative association between NPS and growth in three of four Pakistan telecommunications firms. Jahnert & Schmeiser (2021) found a positive association between NPS scores and the profitability of insurers, but not between NPS and revenue growth, which is the main claim made by Reichheld.

The final study reported on here concluded that NPS scores predicted sales one quarter into the future (Baehre *et al.* 2021) for athletic wear in the US. This study is notable in that it used NPS data obtained from both brand buyers and non-buyers. It seems logical that non-buyers would only give a high recommendation score for brands they are aware of. Therefore, the link between NPS and sales in this study could reflect an effect of brand awareness or familiarity rather than recommendation. To summarise this section, there is mixed evidence pertaining to the link between NPS and future revenue growth.

Low NPS scores and negative word of mouth?

According to Reichheld, respondents who give low likelihood to recommend scores will give the brand negative word of mouth and harm its growth prospects. Indeed, respondents who score between 0 and 6 on the likelihood to recommend scale are labelled as 'detractors'. However, low likelihood to recommend is not the same thing as giving negative word of mouth (East 2008). A consumer may give a low likelihood to recommend score because they do not think friends or colleagues would be interested in hearing a recommendation. For example, a consumer might buy a new vacuum cleaner, be quite satisfied with it, but correctly say they are not likely to recommend it to other people because (a) the topic of vacuum cleaners is unlikely to be a conversation topic; and relatedly (b) the consumer does not know of friends or colleagues who, at present would be interested in hearing a recommendation about such a product. In the non-profit sector, Schulman and Sergeant (2013) point out that charity donors are very unlikely to say they will recommend that charity, but it is because they simply do not wish to discuss their donation behavior with others. Seth et al (2016) reported that Asian respondents, even if they are highly satisfied with a provider, tend to give low NPS scores. The reason is that they fear their friendships or reputation might be damaged by recommending a provider that does not provide satisfactory service. Romaniuk, Nguyen and East (2011) found that 95% of those who said they would not recommend indeed did not; but also found they did not give negative word of mouth. Therefore, these three studies suggest low scores reflect a low probability of giving positive word of mouth, not a high probability of giving negative word of mouth. Further evidence on the usefulness of the NPS question in measuring negative word of mouth comes from Schneider et al (2008) who conducted a series of experiments using different types of questions pertaining to recommendation. The study found that variation in negative word of mouth (past and current, not future) was far better explained by a question specifically asking about likelihood of giving a negative recommendation, compared to the NPS question. Overall, therefore, there is evidence that a low likelihood to recommend, or a low NPS score, does not necessarily equate to negative word of mouth.

Stated likelihood to recommend is not actual recommendation behavior

NPS is a measure of stated likelihood to recommend, not actual recommendation. The reason for asking respondents about their likelihood to recommend is that it allows marketers to ascertain the effect of current customer experience on future recommendation. The basic premise is that pleased customers will say they are more likely to recommend, displeased customers will say they are unlikely to do so. In theory, questions relating to respondent's future likelihood of certain behavior can produce quite accurate aggregate-level estimates. For example, Juster developed a zero-to-ten purchase likelihood scale that has been used to estimate brand penetration (Nenycz-Thiel *et al.* 2013) with reasonable accuracy, as well as demand for various product categories and brands (Wright *et al.* 2002). However, the NPS arguably suffers from a tendency by respondents to over-estimate their likelihood of recommending. This is because they may fail to take account of the rather low chance they will engage in a conversation about their providers of goods and services to friends and colleagues. Rather, NPS responses may reflect that respondents imagine their likelihood of recommending *if* they were engaging in such a conversation.

In terms of empirical evidence, Keiningham (2007a) did find reasonably strong associations between likelihood to recommend and later recommendations of approximately r=0.40. However, other variables such as simple brand preference performed approximately as well (refer Table 3 in that study). Another study, that followed up NPS survey respondents, found that most people who gave high likelihood to recommend scores did not later recommend (Kumar *et al.* 2007). Third, as mentioned earlier, AirBnB did find a higher level of referral among those who have very high NPS scores, but only 4% higher than those who gave very low scores. More evidence is needed to verify the extent to which stated likelihood to recommend does translate into future recommendations.

NPS is promoted as a superior measure to customer satisfaction

The original work on NPS stated clearly that it was a superior metric to customer satisfaction in predicting business growth. One of the bases of this claim was that many apparently satisfied customers defect (e.g. Reichheld 1996, 2003). That observation, however, preceded work that showed a very large proportion of supposed defectors do not defect from a business, rather, their loss is primarily due to unavoidable changes (Bogomolova & Romaniuk 2009). While likelihood to recommend is a different concept to customer satisfaction, firms' NPS scores or likelihood-to-recommend scores are generally highly correlated with satisfaction. Keiningham et al (2007b) found remarkable similarity between NPS scores and ACSI (American Customer Satisfaction Index) scores. Van Doorn, Leeflang and Tijs (2013) reported the correlation between firm's NPS and customer satisfaction scores was r=0.92 (refer Appendix A of that study). At the level of individual respondents, the correlation between NPS scores and satisfaction scores is generally also high: Hosany and Martin (2012) reported a correlation of r=0.80, De Haan et al (2015) r=0.34; Da Silva Santos (2019) r=0.82, and Devlin (2020) reported an r of 0.69. A reason why NPS (or likelihood to recommend) often correlates highly with satisfaction is probably because most word of mouth about a brand comes from brand users (East *et al.* 2007) and most brand users are quite satisfied (Peterson & Wilson 1992; Dawes *et al.* 2020). This high correlation could indicate that NPS provides little information over and above what satisfaction scores do. It therefore seems difficult see how NPS could be a markedly superior metric to satisfaction when the two metrics are usually highly correlated. It is also noteworthy that in a study conducted to determine which of a series of metrics explained the variation in actual past and current recommendation behavior, liking outperformed likelihood to recommend (Schneider *et al.* 2008). Therefore, a reasonable conclusion is that NPS functions more as a proxy for customer satisfaction or possibly an overall favorable attitude to a firm or brand, rather than a loyalty measure or an actual word of mouth measure.

Despite these shortcomings, businesses still use the NPS. It is a simple metric, which has been heavily promoted. But are there other issues with it than the ones discussed above? This study identifies another, as-yet undocumented problem specifically with the counting method used to derive the NPS. The problem is that by subtracting low scores from high scores, the NPS method induces significant unwanted volatility in wave-to-wave results. In the next section we show a simple example of this effect from a brand survey. We then expand the explanation using a hypothetical set of likelihood to recommend scores, then run a simulation study that shows NPS scores fluctuate five times more than what occurs for a simple average of zero-to-ten likelihood to scores.

The NPS method – and the score variation problem

NPS scores are first obtained by asking about likelihood to recommend on a zero to ten scale. Scores of 7 and 8 are discarded on the basis that they represent ‘passive’ customers, however, at face value 7 or 8 out of 10 is a positive score, and indeed Lewis and Mehmet (2020) found scores of 7 and 8 reflected positive customer sentiment. Next, the 0-6 scores are deducted from the 9 and 10 scores. The 9 and 10 scores are combined, as are the 0-6 scores. However, the original scores out of 10 have information value. A score of 10 is different to a score of 9, therefore combining them results in information loss. Moreover, a score of 6 is far better than a score of 1, 2, or 3. But the NPS discards this fine-grained information by aggregating all high scores into one group, and all low scores into another group.

Another aspect of subtracting the 0-6 scores from the 9-10 scores is that small differences in the original likelihood to recommend score are inflated into much larger differences in the NPS. In other words, the NPS double-penalises low scores[1]. To illustrate, suppose we have 10 responses to an NPS survey. Of course, one would not survey ten respondents, this example is merely to illustrate a point. The scores are 10,9,8,8,8,7,6,5,4,4. The average score is 6.9 out of 10. Now let us calculate the NPS. We have two scores of nine and above (20% of scores). We have four scores of six and below – remember that even six out of ten is a negative score according to Reichheld (2003). This means 40% of the scores are negative. Therefore, the NPS is (20% positive – 40% negative) = minus 20. This is 40% of the highest possible score since the NPS ranges from -100 to 100. It is therefore very low in comparison to the mean score (6.9/10), which is nearly 70% of the highest possible score.

Next, suppose we make one minor change, namely we alter the one score of 9/10, and make it 8/10 instead. This makes little difference to the average (now 6.8) but now, because we only have one score above 9, the NPS dramatically plunges to -30! A small 0.1 decimal point change in the average has translated to a 10-point drop in the NPS. Such a large drop in NPS seems quite serious. But here, it has co-incided with a very small change in the average score. This issue can be illustrated further with real data from an NPS survey conducted in the US on financial services brands. The sample comprised business respondents with responsibility for

[1] Sauro (2015) comments that the confidence intervals for NPS scores are often twice as wide as those of average likelihood to recommend scores, depending on the specific proportions of ‘promoters’ and ‘detractors’. Similar points were made by Kristensen and Eskildsen (2011), Eskildsen & Kristensen (2011) and Rocks (2016). Wide confidence intervals is a related, but different point to the issue raised here, which is simply that NPS scores induce more variation across brands or across surveys compared to the original likelihood to recommend scores upon which they are based. Fraser (2017) provides instructions on how to calculate the NPS confidence interval, and notes that NPS requires much larger sample sizes to obtain accurate results compared to alternatives.

commercial banking decisions. The survey was conducted via telephone. The names of the brands are masked for commercial confidentiality. The NPS score, as well as the mean likelihood to recommend score were reported. The results are shown in Table 1.

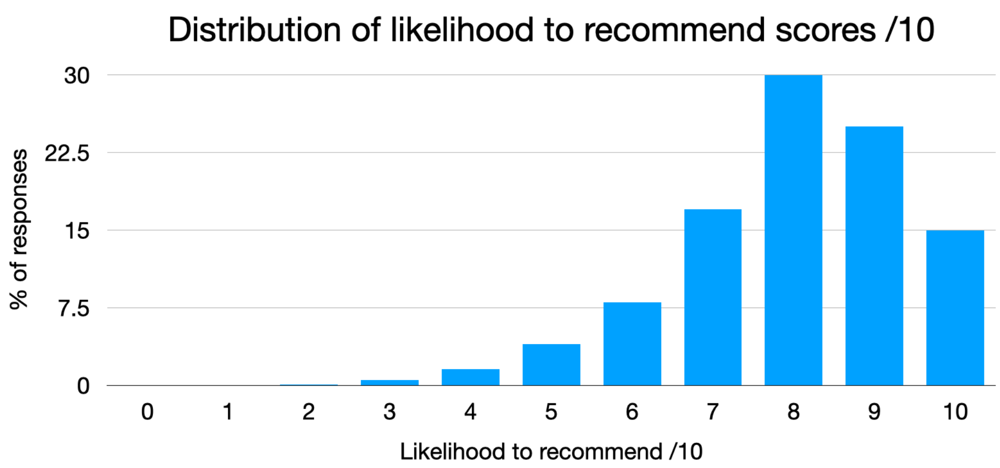
Table 1 NPS scores compared to average likelihood to recommend scores.

Financial Services Brands (US)	NPS score	Mean Average likelihood to recommend score /10
Brand 1	55	8.7
Brand 2	47	8.5
Brand 3	40	8.4
Brand 4	42	8.4
Brand 5	47	8.4
Brand 6	34	8.3

As we see in Table 1, the brands do vary considerably in their NPS scores, with a twenty-one point difference between the highest and lowest score (55 v 34) or in proportional terms, a 38% lower score for brand 6 compared to brand 1. This low score would arguably be concerning for brand 6's management. However, when we examine the average likelihood to recommend scores, the difference between brand 6 and brand 1 is very small: 0.4 scale points, or in proportional terms 5%. This inflated difference in the NPS score is due to minor differences in the proportion of 9-10 scores and 0-6 scores between the brands.

These examples highlight a problem with the NPS scoring system in that it appears to induce heightened variation in scores. To investigate further, we ran a simulation to see how much more period-to-period variation the NPS creates, compared to simply using an average likelihood to recommend score out of ten. Our simulation comprised a population with an average likelihood to recommend score of 8.0 out of 10. We used this figure based on an overall average score over multiple studies. The figure is also consistent with the fact that average satisfaction scores for businesses are usually quite positive (Peterson & Wilson 1992); and as stated, satisfaction and recommendation scores are highly correlated. The population scores for individual 'respondents' in the simulation had a distribution, or spread closely matching real scores that we obtained from our own likelihood to recommend surveys. The distribution of scores is shown in Figure 1.

Figure 1. Distribution of Likelihood to Recommend scores



We checked if the shape, or spread of the scores in our simulated population made any difference to the results. We tested different distributions such as a truncated normal distribution, and the results did not materially alter. We also tested different average likelihood to recommend scores, these did not change the results either.

Also note that the *population* likelihood-to-recommend score stayed the same over the simulation. We took randomly selected samples of this population, using sample sizes of 200 each time. We repeated this process ten times, noting the average likelihood-to-recommend score each time. In effect, this approach is just like doing ten waves of marketplace surveys where the population likelihood to recommend is stable. While the 'sample size' is small, it is the comparison of variation between the mean score and the NPS that is of interest, and this comparison is not contingent on sample size.

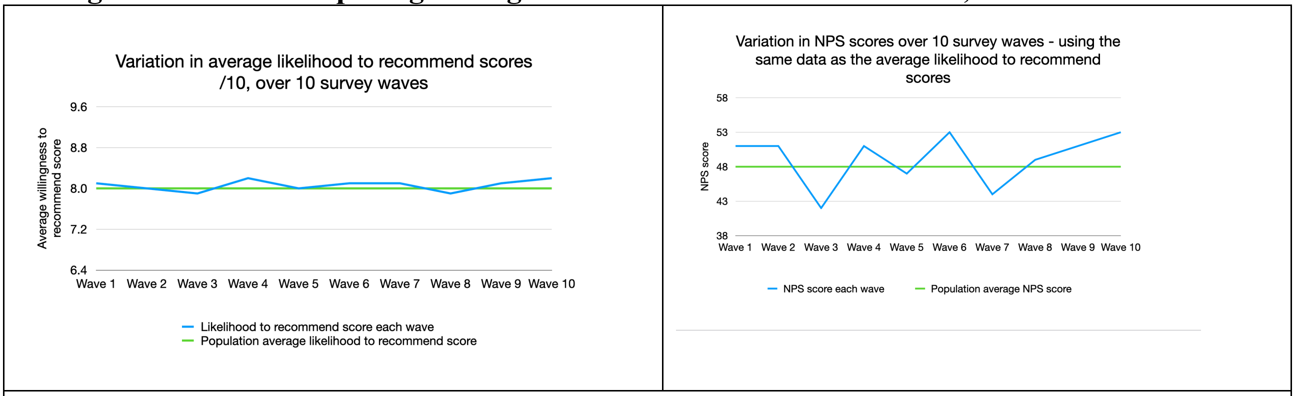
The sample scores we obtained are, of course, not always exactly 8.0 because of random sampling variation. We calculated the mean average score for each sample, or 'wave', and calculated the NPS each time. To compare the volatility of the NPS compared to the average likelihood to recommend score, we calculated the average percentage change in scores from wave to wave, (for example, (wave 2 score – wave 1 score) / wave 1 score * 100) for both the metrics. This metric summarises the wave-to-wave variation in both metrics, controlling for the fact one has a mean of 8.0 and one has a mean just under 50.

The results are shown in Figures 2 and 3 as well as in a numerical table. To make the graphs comparable, we set the range of the Y axis to be plus or minus 20% of the population mean score in both cases.

The likelihood to recommend scores, as per Figure 2, vary from wave to wave by approximately 0.2 of a scale point. They are between around 7.9 out of 10 to 8.2 out of 10. In contrast, the

NPS as seen in Figure 3 fluctuate considerably more, ranging from 53 down to 42. It is noteworthy that in some cases the exact same mean likelihood to recommend score of say, 8.0 translates into different NPS scores of 51 and 47! Plainly this must be undesirable from a measurement viewpoint. The reason for this score fluctuation is that slight differences in the total number of 9 or 10 scores, or zero to six scores, dramatically change the NPS score.

Figures 2 and 3 Comparing average likelihood to recommend scores, to NPS



The variation in NPS is far greater than it is for average likelihood to recommend scores

Table 2. Comparing NPS to average likelihood to recommend scores

Wave:	1	2	3	4	5	6	7	8	9	10	**Average % wave-to-wave change**
Average **likelihood to recommend** score	8.1	8.0	7.9	8.2	8.0	8.1	8.1	7.9	8.1	8.2	**2%**
Corresponding **Net Promoter Score** from the same wave	51	51	42	51	47	53	44	49	51	53	**10%**

In the rightmost column in Table 1 we see that the NPS shows five times as much random variation (10% vs 2%) from wave to wave compared to a simple average likelihood to recommend score. This demonstrates that businesses using the Net Promoter Score will likely be facing heightened variation in scores from wave to wave. Such variation will make it more difficult for market research providers or insights managers to explain what is going on with the NPS results, or to identify if they are meaningfully related to marketing interventions designed to boost the business's NPS level. NPS clients should request (a) mean likelihood to recommend scores as well as the NPS score, to determine if changes in the NPS are reflected in any change in the former, (b) percentage of 'promoters' and 'detractors' to help unpick reasons for change, as well as (c) the confidence interval for the NPS score, which will help communicate to the business that the NPS tool has a very wide margin of error.

Summary

This study outlined the various uses of the NPS in industry. It then identified a series of criticisms of the NPS, and evaluated the evidence pertaining to them. That review concluded (1) there is mixed evidence concerning the presumed link between NPS and firm growth; (2) evidence suggests low NPS scores do not necessarily have the outcome of negative word of mouth; (3) more evidence is needed relating to how NPS scores translate into actual recommendation behavior; and (4) NPS does not appear to be a superior metric to satisfaction in predicting business growth, in fact evidence suggests NPS functions more as a proxy of customer satisfaction than being an indicator of loyalty or word of mouth.

The NPS arguably has a managerial appeal that the scores, since they can be low or even negative, provide significant motivation for managers to improve them. However, balancing this benefit is a disadvantage is that the counting method to produce them produces large differences between brands, and also increases the survey-to-survey score variation for a given brand. This variation could mean that managers might falsely perceive their firm is performing far better, or far worse than competitors when a large part of that apparent performance difference is driven simply by the fact that NPS 'double penalises' low scores. The same set of firms may differ from each other far less if the average likelihood to recommend score is used instead of the NPS. The heightened NPS variation will also make it more difficult for a firm to understand the reasons or 'drivers' of its NPS.

Next, the NPS is based on asking respondents about their likelihood to recommend. Unfortunately, the evidence to date between likelihood to recommend and actual recommendation is mixed. If a business believes that recommendation is important, an alternative is to measure actual recommendations, both given and received by respondents, rather that stated likelihood to recommend. While actual recommendation is more sector-specific, being linked to purchase cycles and product interest, it is potentially far more informative in this regard than NPS.

NPS is also used to ascertain customer satisfaction. However, if a business wants to measure customer satisfaction, it should use questions that specifically ask about satisfaction, not use the proxy of likelihood to recommend. Lastly, if a business is already paying to track customer satisfaction, the high correlation between satisfaction and NPS means the former is already providing most of the information the business would obtain from doing NPS research. While many market researchers are mandated by senior managers to use the NPS, they and the executives who make decisions on key metrics should know that tracking NPS as well as customer satisfaction may provide little incremental information.

References

Adams, C., Walpola, R., Schembri, A.M., and Harrison, R. (2022) The Ultimate Question? Evaluating the Use of Net Promoter Score in Healthcare: A Systematic Review. *Health Expectations*, **ahead-of-print**.

adidas (2019) Adidas Annual Report 2019. Herzogenaurach, Germany: adidas, pp.1-230.https://report.adidas-group.com/2019/en/

Alismail, A., Schaeffer, B., Oh, A., Hamiduzzaman, S., Daher, N., Song, H.-Y., Furukawa, B., and Tan, L.D. (2020) The Use of the Net Promoter Score (Nps) in an Outpatient Allergy and Pulmonary Clinic: An Innovative Look into Using Tablet-Based Tool Vs Traditional Survey Method. *Patient Related Outcome Measures*, **11**, pp.137-42.

Baehre, S., O'Dwyer, M., O'Malley, L., and Story, V.M. (2022) Customer Mindset Metrics: A Systematic Evaluation of the Net Promoter Score (Nps) Vs. Alternative Calculation Methods. *Journal of Business Research*, **149**, pp.353-62.

Baehre, S., O'Dwyer, M., O'Malley, L., and Lee, N. (2021) The Use of Net Promoter Score (Nps) to Predict Sales Growth: Insights from an Empirical Investigation. *Journal of the Academy of Marketing Science*, pp.1-18.

Barclays (2013) Building the 'Go-to' Bank. In *Annual Report*, London, UK: Barclays Bank PLC, pp.1-345.https://home.barclays/content/dam/home-barclays/documents/investor-relations/annualreports/ar2013/barclays-bank-plc-annual-report-2013.pdf

Bogomolova, S. and Romaniuk, J. (2009) Brand Defection in a Business-to-Business Financial Service. *Journal of Business Research*, **62**, 3, pp.291-96.

Chandra, U. and Ro, B.T. (2008) The Role of Revenue in Firm Valuation. *Accounting Horizons*, **22**, 2, pp.199-222.

Coad, A., Daunfeldt, S.-O., and Halvarsson, D. (2018) Bursting into Life: Firm Growth and Growth Persistence by Age. *Small Business Economics*, **50**, 1, pp.55-75.

Cohen, J. and Cohen, P. (1983) *Applied Multiple Regression/Correlation Analysis for the Behavioural Sciences.* New York, United States: Lawrence Erlbaum Associates.

Colvin, G. (2020) The Simple Metric That's Taking over Big Business. In *Fortune*, New York, USA: Fortune.https://fortune.com/longform/net-promoter-score-fortune-500-customer-satisfaction-metric/

Competition and Markets Authority (2020) Latest Banking Customer Satisfaction Results Published. Available at: https://www.gov.uk/government/news/latest-banking-customer-satisfaction-results-published.

Customer Faithful (2022) Why Net Promoter Score Failed Its First Test in Nhs. Available at: https://www.customerfaithful.com/what-we-say/net-promoter-score-fails-its-first-test-in-nhs.

Dawes, J., Stocchi, L., and Dall'Olmo-Riley, F. (2020) Over-Time Variation in Individual's Customer Satisfaction Scores. *International Journal of Market Research*, **62**, 3, pp.262-71.

De Haan, E., Verhoef, P.C., and Wiesel, T. (2015) The Predictive Ability of Different Customer Feedback Metrics for Retention. *International Journal of Research in Marketing*, **32**, pp.195-206.

Department of Health and Social Care (2012) Friends and Family Test: What It Means for Nhs. Available at: https://www.gov.uk/government/news/friends-and-family-test-what-it-means-for-nhs.

Devlin, M.A. (2020) Comparing Net Promoter Score and Jd Power Customer Satisfaction Index: Retail and Small Business Banking. In *Business Administration*, Vol. Doctor of Business Adminstration, Ann Arbor, USA: Wilmington University, pp.99.

East, R. (2008) Measurement Deficiencies in the Net Promoter Score. In *ANZMAC*, Sydney, Australia.

East, R., Hammond, K., and Wright, M. (2007) The Relative Incidence of Positive and Negative Word of Mouth: A Multi-Category Study. *International Journal of Research in Marketing*, **24**, 2, pp.175-84.

Eskildsen, J.K. and Kristensen, K. (2011) The Accuracy of the Net Promoter Score under Different Distributional Assumptions. In *International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, Xi'an, China: IEEE, pp.964-69.

Farooq, M., Raju, V., Khalil-Ur-Rehman, F., Younas, W., Ahmed, Q.M., and Ali, M. (2019) Investigating Relationship between Net Promoter Score and Company Performance: A Longitudinal Study. *Global Journal of Emerging Sciences*, **1**, 1, pp.1-10.

Farris, P., Bendle, N., Pfeifer, P.E., and Reibstein, D.J. (2016) *Marketing Metrics: The Manager's Guide to Measuring Marketing Performance.* New Jersey, United States: Pearson.

Fiserova, J., Pugh, G., Dimos, C., and Stephenson, A. (2018) The Effect of the Net Promoter Score on Sales: A Study of a Retail Firm Using Customer and Store-Level Data. In *British Academy of Management 2018*, Bristol, UK, pp.1-18.

Fraser, C. (2017) How to Find (the Extremely Wide) Confidence Intervals for Net Promoter Scores. Available at: https://medium.com/@colin.fraser/how-to-find-the-extremely-wide-confidence-intervals-for-net-promoter-d1baa8124654.

Grisaffe, D.B. (2007) Questions About the Ultimate Question: Conceptual Considerations in Evaluating Reichheld's Net Promoter Score (Nps). *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, **20**, pp.36-53.

Hamilton, D.F., Lane, J.V., Gaston, P., Patton, J.T., Macdonald, D.J., Simpson, A., and Howie, C.R. (2014) Assessing Treatment Outcomes Using a Single Question: The Net Promoter Score. *The Bone & Joint Journal*, **96b**, 5, pp.622-28.

Ho, E. (2018) Stop Just Looking at Your Net Promoter Score! A Better Way to Utilize Nps. San Francisco, USA: Medium.https://medium.com/userinterviews/stop-just-looking-at-your-net-promoter-score-a-better-way-to-utilize-nps-1ed97df4bbc3

Hoek, J. and Gendall, P. (2010) Colors, Brands, and Trademarks: The Marketing (and Legal) Problem of Establishing Distinctiveness. *Journal of Advertising Research*, **50**, 3, pp.316-22.

Hosany, S. and Martin, D. (2012) Self-Image Congruence in Consumer Behavior. *Journal of Business Research*, **65**, 5, pp.685-91.

HSBC Holdings plc (2022) Annual Report and Accounts 2021. In *Annual Report and Accounts*: HSBC, pp.1-408.https://www.hsbc.com/investors/results-and-announcements/annual-report

Jahnert, J.R. and Schmeiser, H. (2021) The Relationship between Net Promoter Score and Insurers' Profitability: An Empirical Analysis at the Customer Level. *The Geneva Papers on Risk and Insurance-Issues and Practice*, pp.1-29.

Jet Blue (2008) Proxy Statement 2008 Annual Meeting of Stockholders. In *Proxy Statements*, Forest Hills, New York, USA, pp.1-72.http://investor.jetblue.com/~/media/Files/J/Jetblue-IR-V2/Annual-Reports/2008jetblueproxy.pdf

Jet Blue (2015) Morningstar Document Research

Form Def 14a. Long Island City, New York, USA: Jet Blue.

Jones, M.A., Mothersbaugh, D.L., and Beatty, S.E. (2002) Why Customers Stay: Measuring the Underlying Dimensions of Services Switching Costs and Managing Their Differential Strategic Outcomes. *Journal of Business Research*, **55**, 6, pp.441-50.

Kara, A., Mintu-Wimsatt, A., and Spillan, J.E. (2022) An Application of the Net Promoter Score in Higher Education. *Journal of Marketing for Higher Education*, pp.1-24.

Keiningham, T.L., Cooil, B., Aksoy, L., Andreassen, T.W., and Weiner, J. (2007a) The Value of Different Customer Satisfaction and Loyalty Metrics in Predicting Customer Retention, Recommendation, and Share-of-Wallet. *Managing Service Quality*, **17**, 4, pp.361-84.

Keiningham, T.L., Cooil, B., Andreassen, T.W., and Aksoy, L. (2007b) A Longitudinal Examination of Net Promoter and Firm Revenue Growth. *Journal of Marketing*, **71**, pp.39-51.

Korneta, P. (2018) Net Promoter Score, Growth, and Profitability of Transportation Companies. *International Journal of Management and Economics*, **54**, 2, pp.136-48.

Kristensen, K. and Eskildsen, J. (2011) Is the Net Promoter Score a Reliable Performance Measure? In *IEEE International Conference on Quality and Reliability*, Bangkok, Thailand: IEEE, pp.249-53.

Kumar, V., Petersen, J.A., and Leone, R.P. (2007) How Valuable Is Word of Mouth? *Harvard Business Review*, **October**, pp.139-46.

Lewis, C. and Mehmet, M. (2020) Does the Nps® Reflect Consumer Sentiment? A Qualitative Examination of the Nps Using a Sentiment Analysis Approach. *International Journal of Market Research*, **62**, 1, pp.9-17.

Mecredy, P., Wright, M.J., and Feetham, P. (2018) Are Promoters Valuable Customers? An Application of the Net Promoter Scale to Predict Future Customer Spend. *Australasian Marketing Journal*, **26**, 1, pp.3-9.

Morgan, N.A. and Rego, L.L. (2006) The Value of Different Customer Satisfaction and Loyalty Metrics in Predicting Business Performance. *Marketing Science*, **25**, 5, pp.426-39.

NatWest Group (2021) Annual Report and Accounts 2020. In *Annual Report and Accounts*, London, UK: NatWest Group plc, pp.1-

367.https://investors.natwestgroup.com/~/media/Files/R/RBS-IR-V2/results-center/19022021/natwest-group-annual-report-accounts-2020-v1.pdf

Nenycz-Thiel, M., Beal, V., Ludwichowska, G., and Romaniuk, J. (2013) Investigating the Accuracy of Self-Reports of Brand Usage Behavior. *Journal of Business Research*, **66**, 2, pp.224-32.

NHS (2020) Friends and Family Test Data. Available at: https://www.england.nhs.uk/fft/friends-and-family-test-data/.

Owen, R. (2019) Net Promoter Score and Its Successful Application. *Marketing Wisdom* Springer, pp. 17-29.

Owen, R. and Brooks, L.L. (2008) *Answering the Ultimate Question: How Net Promoter Can Transform Your Business.* Hoboken, New Jersey, USA: John Wiley & Sons.

Parackal, M. and Brennan, M. (1998) Obtaining Purchase Probabilities Via a Web Based Survey: The Juster Scale and the Verbal Probability Scale. *Marketing Bulletin*, **9**, May, pp.67-75.

Perceptive (2022) Grow Your Business with Nps. Available at: https://www.customermonitor.com/grow-your-business-with-nps.

Peterson, R.A. and Wilson, W.R. (1992) Measuring Customer Satisfaction: Fact and Artifact. *Journal of the Academy of Marketing Science*, **20**, No. 1, pp.61-71.

Pingitore, G., Morgan, N.A., Rego, L.L., Gigliotti, A., and Meyers, J. (2007) The Single-Question Trap. *Marketing Research*, **19**, 2, pp.8-14.

Qian, L. (2015) How Well Does Nps Predict Rebooking? In *The Airbnb Tech Blog*, Vol. 2022, ed. Airbnb, California, USA: Airbnb.https://medium.com/airbnb-engineering/how-well-does-nps-predict-rebooking-9c84641a79a7

Qualtrics (2022) What Is Nps? Your Ultimate Guide to Net Promoter Score. Available at: https://www.qualtrics.com/au/experience-management/customer/net-promoter-score/.

Ramshaw, A. (2014) The Practicalities of Giving Frontline Staff Net Promoter Targets. In *Customer Think*, Vol. 2022, Coronado, United States.https://customerthink.com/the-practicalities-of-giving-frontline-staff-net-promoter-targets/

Reichheld, F.F. (1996) Learning from Customer Defections. *Harvard Business Review*, **74**, 2, pp.56-69.

Reichheld, F.F. (2003) The One Number You Need to Grow. *Harvard Business Review*, December, pp.46-54.

Rocks, B. (2016) Interval Estimation for the “Net Promoter Score”. *The American Statistician*, **70**, 4, pp.365-72.

Romaniuk, J., East, R., and Nguyen, C. (2011) The Accuracy of Self-Reported Probabilities of Giving Recommendations. *International Journal of Market Research*, **53**, 4, pp.507-21.

Santos, D.A.d.S. (2019) Assessing the Pattern between Customer Metrics and Net Promoter Score: A Portuguese Insurance Sector Case Study. Lisbon, Portugal: Universidade Nova de Lisboa, pp.1-42.http://hdl.handle.net/10362/71864

Satmetrix (2018) Us Consumer 2018 Net Promoter Benchmarks at a Glance. Redwood City, CA.NICE-Satmetrix-infographic-2018-b2c-nps-benchmarks-050418.pdf

Sauro, J. (2015) *Customer Analytics for Dummies.* John Wiley & Sons.

Sauro, J. (2018) The One Number You Need to Grow (a Replication). In *Measuring U*: Measuring U.https://measuringu.com/nps-replication/

Sauro, J. (2019) Assessing the Predictive Ability of the Net Promoter Score in 14 Industries. Denver, CA.https://measuringu.com/nps-predictive-ability/

Schneider, D., Berent, M., Thomas, R., and Krosnick, J. (2008) Measuring Customer Satisfaction and Loyalty: Improving the 'Net-Promoter'score. In *Annual Meeting of the American Association for Public Opinion Research*, New Orleans, Louisiana, USA, pp.1-94.

Schulman, K. and Sargeant, A. (2013) Measuring Donor Loyalty: Key Reasons Why Net Promoter Score (Nps) Is Not the Way. *International Journal of Nonprofit and Voluntary Sector Marketing*, **18**, 1, pp.1-6.

Seifert, C.F. and Yukl, G. (2010) Effects of Repeated Multi-Source Feedback on the Influence Behavior and Effectiveness of Managers: A Field Experiment. *The Leadership Quarterly*, **21**, 5, pp.856-66.

Seth, S., Scott, D., Svihel, C., and Murphy-Shigematsu, S. (2016) Solving the Mystery of Consistent Negative/Low Net Promoter Score (Nps) in Cross-Cultural Marketing Research. *Asia Marketing Journal*, **17**, 4, pp.43-61.

Southwest Airlines (2015) Proxy Statement for Annual Meeting of Shareholders. Dallas, Texas, USA, pp.1-94.http://investors.southwest.com/~/media/Files/S/Southwest-IR/documents/2015-proxy-statement.pdf

Stirling, P., Jenkins, P.J., Clement, N.D., Duckworth, A.D., and McEachan, J.E. (2019) The Net Promoter Scores with Friends and Family Test after Four Hand Surgery Procedures. *Journal of Hand Surgery (European Volume)*, **44**, 3, pp.290-95.

Van Doorn, J., Leeflang, P.S., and Tijs, M. (2013) Satisfaction as a Predictor of Future Performance: A Replication. *International Journal of Research in Marketing*, **30**, 3, pp.314-18.

Wright, M., Sharp, A., and Sharp, B. (2002) Market Statistics for the Dirichlet Model: Using the Juster Scale to Replace Panel Data. *International Journal of Research in Marketing*, **19**, 1, pp.81 - 90.

Zaki, M., Kandeil, D., Neely, A., and McColl-Kennedy, J.R. (2016) The Fallacy of the Net Promoter Score: Customer Loyalty Predictive Model. Cambridge, UK: University of Cambridge, pp.1-26.