Measuring Donor Loyalty

Loyalty runs deep, and metrics such as the Net Promoter Score can be too simplistic.
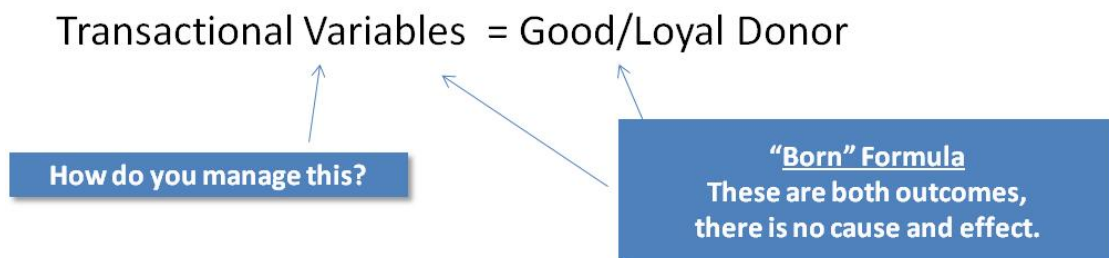
January 2013 By Adrian Sargeant and Kevin Schulman

For decades, business has pursued an effective formula for customer loyalty. But, despite rigor and expense, the secret to enduring relationships remains elusive. Over the years the concepts of "satisfaction," "value" and "quality" have all taken their turns as the key to customer profitability. However, one by one each has proven to be an insufficient indicator of future customer behavior. Although these costly frameworks improve survey scores, they often have limited impact on the bottom line.

The current in-vogue concept is the Net Promoter Score (NPS) that, like many of its predecessors, uses an "attitudinal" framework to measure loyalty.

More about NPS in a moment, but first this question: Why measure attitudes at all? The reason attitudinal frameworks exist is simple: Capturing how the customer (or donor) thinks or feels provides different insights from what we can learn looking at past behavior — i.e., transactional data. If measured properly, attitudinal insights can be additive, providing a multidimensional and more accurate view of the donor, customer or constituent.

For example, are there some past behavior patterns that viewed through the transactional lens look like "loyalty" but are in fact spurious? Similarly, isn't it possible that many constituents are very committed to an organization yet those feelings have not manifested in "good" behavior, as measured by past transactional conduct (i.e., what we call "latent loyalty")?

Transactional Variables = Good/Loyal Donor

How do you manage this?

"Born" Formula
These are both outcomes,
there is no cause and effect.

Conventional practice seeks to measure and understand loyalty through transactional analysis. Fundraisers working from this perspective equate loyalty with a particular pattern of purchases, contributions, advocacy actions, etc., and seek to build it by pushing enough of the "stuff" that seems to generate these behaviors (appeals, catalogs, e-mails, videos, petitions, etc.). The goal isn't to create loyal donors through communications. Rather this approach assumes that some donors are innately loyal and simply need to be prodded to give. Put another way, sufficient volume increases the likelihood of "good" donors raising their hands, responding and thus keeping themselves in the "good" bucket.

It seems silly and overly simplified to state it this way, but this is precisely how much of the nonprofit sector tends to operate — the group who pushes out the most stuff through the most channels wins. This is little more than a race to the bottom.

If not that, then what?
What is the alternative? A better approach is to separate cause and effect, and develop a more accurate representation of what actually is occurring in the marketplace. It is borderline heretical to say this, but nonprofits do not directly impact donor behavior, only indirectly. What an organization directly impacts through the experiences it serves up across marketing, fundraising and donor service are donor perceptions of the organization and its approach. This in turn shapes how donors view their relationship with the nonprofit and determines their behavior.

Thus, organizations can more efficiently and effectively improve donor behavior by getting a handle on what organizational actions they take today that improve or detract from the donor relationship. It is the quality of the donor relationship that dictates whether donors stay or go.

If you accept this conceptual "creation" formula, then two high-level requirements are mandatory for any framework claiming to measure attitudinal loyalty. It must be:

1. predictive of outcomes (i.e., right side of the formula above) and
2. able to identify (with modeling) the organizational levers (i.e., experiences) that matter most to increasing loyalty and value.

Of course, life isn't that neat. Identifying the most appropriate organizational levers and then finding adequate ways to measure them have proved difficult, and over the past three decades a number of different approaches have fallen in and out of favor. In the 1980s, Parasuraman, Zeithaml and Berry developed the SERVQUAL scale to measure the quality of service provided

to customers. This measured what they regarded as the five underlying components of any service, namely:

- Tangibles: the appearance of facilities, staff, premises and communication materials
- Reliability: the company's ability to perform the desired service dependably and accurately
- Responsiveness: the company's willingness to help customers and provide prompt service
- Assurance: the knowledge and courtesy of employees and their ability to convey trust and confidence
- Empathy: the degree to which the company offered individualized and caring attention

While a plausible approach, it is now generally accepted that SERVQUAL failed to provide the results the authors had originally envisaged. The dimensions were very general, making it difficult to highlight specific areas where actions might be taken to improve the quality of service. The scores on each dimension reflected the aggregate approach of the organization as a whole rather than one department or individual, and it proved impossible to make concrete recommendations for change. It also proved ill-suited to the arena of direct response, where customers rarely had the level of direct contact necessary to answer the full suite of questions posed by the authors.

Despite its weaknesses, the SERVQUAL approach gained much traction because of a mounting body of evidence of a link between customer satisfaction, loyalty and, ultimately, profitability. As researchers began to understand more of the dynamic, we learned that although this was the case, the relationship between satisfaction and loyalty was nonlinear and that behavior tended to be impacted by extremes of experience. Customers who were "very satisfied or delighted" were substantively more loyal, while customers who were dissatisfied were very unlikely to repurchase and substantively more likely to engage in negative behaviors such as bad-mouthing the organization to others.

More recently, the Net Promoter Score developed by Frederick Reichheld has targeted specifically the notion of the "buzz" generated by an organization and in particular the willingness on the part of consumers to engage in positive and negative word-of-mouth. In his approach, customers are asked, "How likely is it that you would recommend us to a friend or colleague?" Then they provide a rating from 0 ("Not at all likely") to 10 ("Extremely likely").

The measure is called the "net promoter" score because detractors are subtracted from promoters. Detractors are defined as respondents rating their likelihood to recommend 6 or less, with promoters only those who rated their likelihood a 9 or 10. Respondents who selected 7 or 8 are considered neutral. The NPS measure can run from -100 (0 percent promoters, 100 percent detractors) to 100 (0 percent detractors, 100 percent promoters), with typical measures in the 30 percent to 40 percent range.

NetPromoter Score®

Promoters minus Detractors = NPS
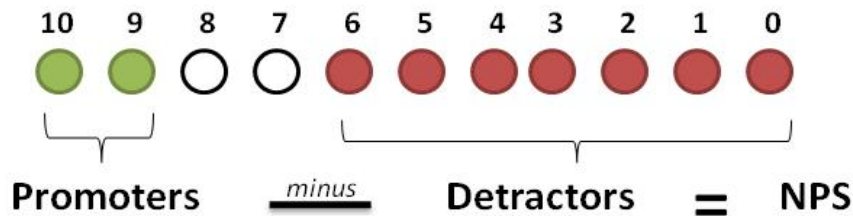
What is wrong with NPS?
As we have indicated above, the approach is beautifully straightforward. You only need to concern yourself with a single key metric and maximizing the net number of recommends garnered. Superior performance in fundraising then is linked to the size of the number obtained.

Unfortunately, there is increasing evidence that the approach is flawed and doesn't deliver the silver bullet necessary for managing retention in our sector. Here's why.

1. NPS assumes low scores are active "detractors" of brand. Reichheld and other proponents of NPS have taken what is clearly a unipolar question of willingness to do something or not (i.e., will or will not recommend) and turned it into a bipolar one with willingness to recommend on one end and willingness to detract on the low end. In other words, we are to accept or believe that those who give a low score on the "willingness to recommend" question are not going to recommend your brand and also will actively say bad things about it — hence the "detractor" term.

From a management standpoint, if nonprofits are to treat low NPS scores as mission-critical, it is likely they will devote more effort and resources to improving the scores of the detractor segment. They would be mistaken to do so since low scores are in themselves indicative of nothing. Critically, a low score may not be an indicator of a negative sentiment, merely that the individual in question does not engage in offering recommendations. He may well have a favorable view of the organization and may indeed be passionate about the work undertaken, but he just doesn't like talking about his experiences to others.

In fundraising, the measure is particularly problematic because it is deemed culturally inappropriate in many countries and contexts to discuss one's charitable giving. People simply don't discuss their philanthropic choices in the same way they do their cars, holiday destinations or computing choices. As a consequence, using the Reichheld method in the nonprofit sector results in a disproportionately large segment of apparent detractors, and the net score is therefore a meaningless amalgam of different perceptions.

2. NPS throws away data. Throwing away data is an odd description, but in essence, that is what NPS does by collapsing the 9s and 10s and the 0 to 6s and ignoring the 7s and 8s. There is ample statistical and empirical evidence for this being wrongheaded with 0s being behavior-wise nothing like a 6. And this says nothing of the 7s and 8s who are ignored completely in this methodology.

In aggregate, the approach has a very arbitrary feel with the rich pattern of attitudes originally articulated by respondents almost completely ignored. If the desire was simply to create a binary variable (will recommend, will not recommend), one can only wonder as to why that was not the option presented to consumers/donors in the first place.

3. NPS does not identify the full set of organizational experiences that matter. The system of NPS consists of only a single question: willingness to recommend. That's it! And while simplicity is an important goal, NPS takes it to the extreme. Reichheld argues that NPS is the ultimate measure and that everything you need to know to predict growth can be explained with NPS. He goes so far as to assert that other survey-based metrics such as customer satisfaction have no link to growth at all.

Current academic thinking and research, by sharp contrast, have highlighted the importance of a wide range of factors that drive customer and donor loyalty, with the most successful predictive models based on a broad range of different dimensions. We know, for example, that donor loyalty is driven by an amalgam of satisfaction with the service provided by the fundraising team (i.e., the donor experience), commitment to the organization's mission and trust in the organization to have the impacts it has promised with its beneficiaries. Models embracing a range of these different dimensions typically have substantively more predictive power than any one measure.

Employing the NPS system also provides zero indication of why people score the way they do. There is no guidance — specific, general or otherwise — on how to do root cause analysis and understand the "why" of responses and determine the specific levers under the organization's control to drive up NPS. By contrast, if one measures the different dimensions of the donor experience and, critically, how important they are to donors, then one begins to generate managerially useful data. Those aspects rated as high in terms of importance and low in terms of satisfaction are obvious candidates for management intervention.

4. Recommendation is not the primary goal. In response to the criticism above, the NPS creators have suggested conducting a key driver analysis with NPS as the dependent variable to identify the organizational activities that impact it – i.e., to look at what drives the ratings obtained. We see no rationale for adopting such an approach because it isn't recommendation per se that is of interest to most fundraisers. Are we really interested in spending time and effort isolating the factors that cause people to recommend us, or are we more interested in isolating the factors that drive up donor satisfaction with experience and lifetime value to the organization?

NPS rose to prominence off the back of the assertion that it was a good predictor of loyalty. But what do we mean by loyalty? Continuing to be a donor is not the same thing as increasing (or decreasing) the amount of one's giving, or spending a bigger (or smaller) proportion of one's charitable pot on a focal organization. All of these behaviors, in turn, are quite different from being willing to recommend the organization to a friend. Each one of these dimensions is associated with loyalty, but loyal donors need not exhibit all these behaviors — and most don't.

In simple terms, looking only at a willingness to recommend is too narrow an approach to capture the richness of donor behavior, particularly when it never occurs to many individuals to recommend a favored charity to someone in the first place.

5. NPS is not as predictive of giving as other measures. The purpose of attitudinal frameworks like NPS is to help organizations increase donor loyalty by nurturing those who love the organization (to get a greater share of wallet and actual recommend behavior) and properly identify those who don't. Then, where financially worthwhile, to repair what is broken and grow the relationship.

Unfortunately, NPS does not do a very good job of discriminating key behaviors. Put another way, the "promoters" are not all that different from the "detractors" when you look at how they behave. The high, low, % incresse chart (at right), from our recent [Donor Commitment Study](#) ([DonorVoice](#) U.S. Donor Commitment study, November, 2011) affirms what many others have found — NPS (the last column on the right) is not as good as Donor Satisfaction (or a model based on commitment) at identifying differences in behavior as evidenced by the last row showing the percentage difference in giving among those "high" and "low" on the various frameworks. Perhaps the ultimate indictment of "willingness to recommend" comes from a study by [Schneider, Berent, Thomas & Krosnick](#) (2007), who found willingness to recommend is not as good as satisfaction in predicting actual recommend behavior.

6. Confusion over what NPS is really designed to do. In June 2011, NPS creator Reichheld wrote: "[The reason that so many researchers hate NPS is that so many senior line executives love it.](#)"

He continued to defend NPS by saying that while it was less accurate for predicting individual customer behavior than other measures, it was better at predicting business growth. But a few weeks later he wrote that predicting individual behavior was the basis of NPS — rather than a correlation to growth. Recent responses to criticism on the part of those responsible for NPS are characterized by caution, caveats and more than a bit of confusion. Is it designed to predict loyalty? Or growth?

The emergent academic evidence on NPS is damning in both respects. [Keiningham et al (2007), in a study](#) published in the [*Sloan Management Review*](#), found no evidence that NPS was the best predictor across customers' future loyalty intentions. The authors also attempted to substantiate the assertion of a link between NPS and growth, a facet of the measure that has proved highly attractive to managers.

Keiningham et al examined data from more than 15,000 consumers and 21 companies over multiple years. They then added in the growth rates for those companies under investigation. None of the range of metrics they examined, including NPS, was found to be a good predictor of growth. As the authors note, "even when ignoring statistical significance (the likelihood that the correlations occurred by chance), Net Promoter was the best predictor in only two out of 19 cases." They conclude that "based on our research it is difficult to imagine a scenario in which Net Promoter could be called the superior metric."

Conclusions

Loyal donors are those who perceive they have a strong relationship with the organization. In measuring loyalty, one must therefore unpack this relationship in a meaningful, not simplistic, way and understand what genuinely drives the perceptions of donors.

We know from research that multiple factors are at work, notably how satisfied donors might be with the quality of service provided by the fundraising team. Organizations must therefore unpack the dimensions of their services and ascertain the extent to which donors are satisfied with each. To obtain managerially useful information, however, data must also be gathered in respect to perceptions of importance. Then, those aspects of the service scoring high on importance and low in terms of satisfaction would be clear candidates for investment.

Equally, extant research also tells us that satisfaction is not in itself enough. Learning from the commercial world has taught us that sometimes even very satisfied customers will defect, doing so because while they may be very satisfied with their treatment they lack commitment to the organization. In our world, donors who are committed to the organization, cause and/or brand are substantively more loyal than those who are not.

We also understand a lot about the implications of trust on giving. In this context, most donors (unless they are major donors) are not able to see for themselves exactly how their gifts of $20 or $50 were applied to the cause. Instead they must trust the organization to do what it promised to do in its communications. Donors with higher levels of trust in a focal organization donate higher proportions of their philanthropic "pots" than those with lower levels of trust. They also have longer lifetimes of support and consequent lifetime value.

Finally, to take other learning from the commercial world, attitudes are one of the best predictors of subsequent loyalty. Specifically, if I indicate in a survey that I will continue to be a loyal donor, by and large I will continue to be a loyal donor. Equally, if I indicate that I intend to give again next year, I very likely will. Repurchase intention, as it would be labeled in the commercial world, is a very good indicator of subsequent behavior.

So what is a fundraiser to do?

We recommend developing a composite measure of donor attitudes and opinions that captures data on two or more of the constructs that we know are good indicators of loyalty: satisfaction, trust, commitment and/or repurchase intention. An amalgam of all four produces the strongest measure of subsequent loyalty, although there are obvious trade-offs with how cumbersome a measurement instrument might become.

We also urge managers to pick an instrument that includes a diagnostic dimension. Knowing that a donor is satisfied, committed, trusting and predisposed to giving again (or not) is conceptually interesting and might feed into a balanced scorecard of performance, but it isn't helpful in guiding actions to improve the status quo. One must also understand why someone achieves a given level of satisfaction, commitment, etc., in order to take action.

In developing this understanding, one begins to access the levers that might be pulled to engineer loyalty and perhaps even more importantly enhance the donor's overall experience of giving.

In that way, we make giving enjoyable, enhance the "warm glow" that derives from giving and develop the philanthropy of our society as a consequence.

References

• *Schneider, D., Berent, M. K., Thomas, R., & Krosnick, J. A. (2007). Measuring customer satisfaction and loyalty: Improving the 'net promoter' score. Paper presented at the World Association for Public Opinion Research annual meeting, Berlin, Germany.*
• *Timothy L. Keiningham, Bruce Cooil, Tor Wallin Andreassen, and Lerzan Aksoy (2007). A Longitudinal Examination of Net Promoter and Firm Revenue Growth. Journal of Marketing: Vol. 71, No. 3, pp. 39-51.*

[Adrian Sargeant](#) is the [Robert F. Hartsook Professor of Fundraising](#) at [Indiana University](#). Reach him at [asargean@iupui.edu](#). Kevin Schulman is CEO at [DonorVoice](#). Reach him at [kschulman@thedonorvoice.com](#)