

Understanding Business Drivers for Software Products from Net Promoter[®] Score Surveys

Dr. Thomas M. Fehlmann

Euro Project Office AG, Zurich, Switzerland
thomas.fehlmann@e-p-o.com

Abstract. Net Promoter[®] Score (NPS) has become the de facto standard for customer surveys; however, how can you get specific information about what drives customer referrals out of a “one question” survey? The goal of this paper is to explain how NPS survey results can be leveraged for achieving software business success, using Six Sigma concepts. The approach includes detection, selection, and validation of business drivers for software development, and uncovers importance and customer satisfaction with their implementation in the software product.

Keywords: Software Product Management, Net Promoter[®] Score, Six Sigma, Quality Function Deployment, Business Drivers.

1 Introduction

1.1 The Net Promoter[®] Score as a Measurement Method for Loyalty

Net Promoter[®] Score¹ (NPS) surveys have become very popular since they allow asking customers for their preferences with only two questions. The first question is: “How likely is it that you would recommend <our product/service/other> to a friend or colleague?” Respondents are expected to rate likeliness of such a referral with a score between 0 and 10; ten being extremely likely, five neutrals, and zero not at all likely. The second question asks for a short comment explaining why this score was selected. The first question evokes emotional factors that might be reflected in the free-text comments.

The NPS is calculated by subtracting the percentage of respondents that gave a score of 6 or below, called *Detractors*, from the percentage that gave a 9 or 10, called *Promoters*. Scores of 7 or 8 are considered *Passives*; they don’t add to the calculation of the net difference. Research performed by Fred Reichheld of Bain & Company, Inc. and by Satmetrix Systems, Inc. [12] claims that NPS correlates with the following behavior of customers:

¹ Net Promoter, NPS, and Net Promoter Score are trademarks of Satmetrix Systems, Inc., Bain & Company, and Fred Reichheld.

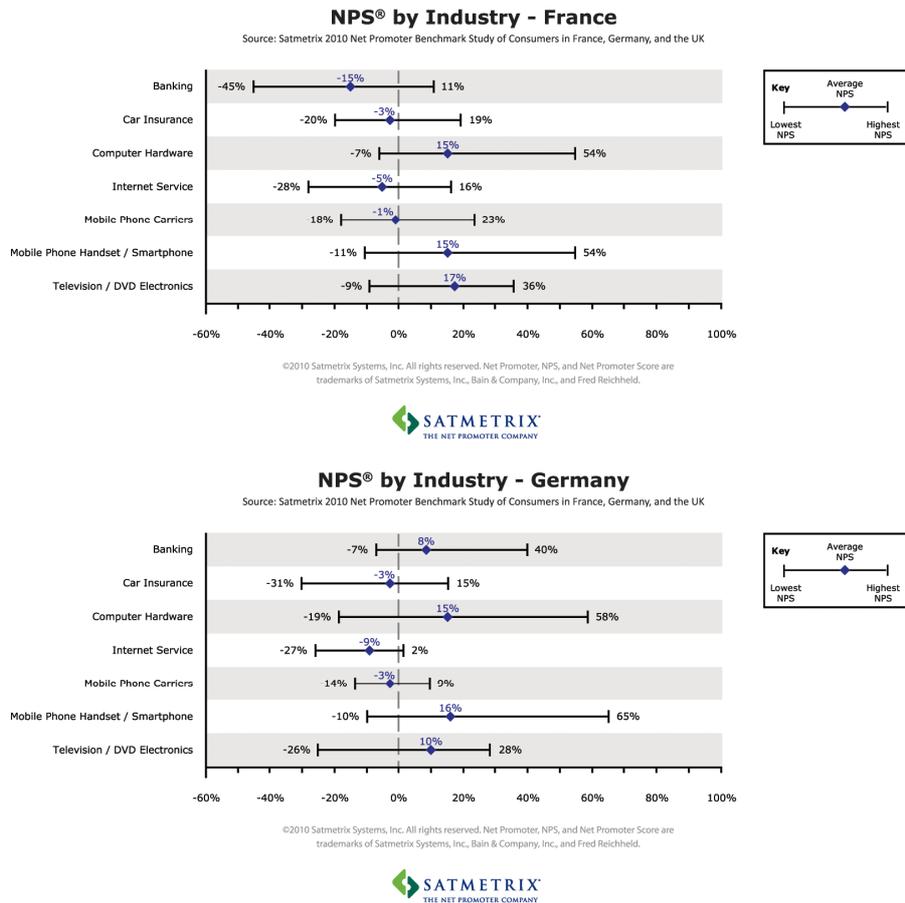
- Continue buying additional features and upgrades
- Buy other software products
- Refer software to other prospects
- Give feedback

These characteristics pertain to behavior which is typical for *Customer Loyalty* [13]. Consequently, a high NPS indicates that management of a software company is well positioned to invest into the future and into new software products; a low NPS suggests that even if current sales look satisfactory, overall turnover will decline as soon as alternatives are available, and investing into new products is rather risky.

1.2 NPS Benchmarks

Satmetrix publishes benchmarks that compare industries in different countries.

Fig. 1. NPS® by Industry Samples – France and Germany



The NPS scale between 0 and 10 doubles the Likert [9] scale of 5 points between 1 and 5; it offers the highest intercultural applicability, according to Reichheld and Satmetrix. However, nationality, language, and role status of the respondents might influence the score. If absolute values differ, it is possible to investigate the reasons why they do. In the sample cases presented in **Fig. 1**, they might differ for many reasons, including a legal environment that facilitates or deteriorates business activities like banking, or car insurance. The customer experience depends from internal processes of each bank or insurance company. Thus, NPS may uncover more information, when combined with information about the respondents and the environment.

1.3 Satisfaction does not mean Loyalty!

Customer satisfaction is not necessarily correlated to loyalty. Customers, even when satisfied with a product, do not intend to become loyal. This would be against their business interest; they need to remain free to select whatever they consider best for them and for their business success in a *Business-to-Business* (B2B) setting, or in personal life (*Business-to-Consumer*, B2C). Next time, the satisfied customer probably buys a competing product, if for nothing else than to remain independent.

Loyalty arises from different sources than customer satisfaction. It could be that the product is unique, unrivalled. Such situations occurred in software product business frequently in the past but seldom lasted long. Products could appeal for social or status splendor, but not so much in B2B as in consumer business. Thus, looking at the value for the customer's business is probably the best candidate for driving loyalty in B2B – a software supplier that has repeatedly excelled in providing value is probably the most likely to keep customers loyal. Therefore we should try to understand what drives value for business when analyzing NPS surveys.

2 Analyzing NPS Surveys

When analyzing an NPS survey, only one kind of numerical data is available from the first question of the survey. The second question is open and needs special analysis methods. In *Quality Function Deployment* (QFD) [1], essential element of *Design for Six Sigma* (DfSS) [2], suitable analysis techniques for NPS surveys are available. Originally, QFD matrices evolved as a combination of Ishikawa (fishbone) diagrams. Specifics of QFD, when applied to software development, are explained in [5].

2.1 What Drives High NPS?

NPS surveys would be useless if one assumes that emotions are driven by irrational factors. Research shows that the contrary is true. Business decisions taken by “*Gut Feeling*” [7] are often better than decisions based on rational measurement methods; “better” in the sense of yielding better business results at the end. If that's the case, emotions must depend from those business factors that drive business success.

Thus, our hypothesis is that NPS for B2B software products is driven by *Business Drivers*; a concept introduced by Denney [3] to deliver quality in software, and used as completion criteria in Lean Six Sigma software development [6]. Sample business drivers include performance, ease-of-use, integration with other systems, workflow automation, integrated business intelligence, also security, safety, and privacy, and other similar criteria that typically appear in acceptance criteria for software projects as so-called *Non-Functional Requirements*. Meeting requirements for business drivers in B2B helps software users to make their business targets and might be paramount to success. If an NPS survey helps detecting and selecting the right business drivers for software product development, such surveys add significant value to an organization.

2.2 A Sample Case – Segmentation of Respondents

NPS is a single score measurement. However, your customers are not homogeneous. You can segment customers into industries or categories, and respondents into categories such as *Deciders, Influencers, and Users* [12].

For instance, take a software company that segments its customers into two categories: *Enterprises, and Factories*. Enterprises are vertically integrated, with marketing, ICT, and product management. They use the software product in end-to-end business processes, they care themselves for everything that resulted out of the software product, covering all consequential follow-up actions.

Factories only use the software for delivering one single step of value added in some business process that is not under their full control. They get an order; deliver it as a service, and their customer cares for all the following business processes. Thus, it might be an outsourced service, requiring the same software features, but perception in factories is considerably different. While integration with existing ICT applications is not so important for the factory, for enterprises it is crucial. Factories have to look at software pricing more closely than enterprises; enterprises can better afford looking for value rather than price cuts. It is therefore to be expected that the NPS of factory or enterprise representatives differ for the same software product. Since the product must serve different market segments, it must meet different expectations at once.

Table 1. Sample NPS Profile According Customer Segments

| | | |
|--------------------------|---------|-----|
| Enterprise – Decider: | NPS-E.1 | 50% |
| Enterprise – Influencer: | NPS-E.2 | 31% |
| Enterprise – User: | NPS-E.3 | 50% |
| Factory – Decider: | NPS-F.1 | 39% |
| Factory – Influencer: | NPS-F.2 | 23% |
| Factory – User: | NPS-F.3 | 31% |

The total NPS taking all customer segments into account is 30%. Responses received from the survey yield a profile when NPS is calculated for each customer segment, see **Table 1**. In this sample case, the NPS profile has six entries, one for any combination of {Decider, Influencer, and User} with {Enterprise, Factory}.

NPS scores are not continuous satisfaction ratio scales. The eleven selectable scores between 0 and 10 collapse into three group percentages promoters, passives,

and detractors. Only promoters and detractors are explicitly included in the calculation. However, when passives move into promoter or detractor status, this also affects the overall score, accounting for shifts in the distribution of responses. Stable scores require sufficiently large samples.

2.3 Verbatim Analysis

The second question in the NPS survey asks for an explanation for the given score. It is free text that needs individual reading and interpretation. *Verbatim Analysis* means that comments and score justifications are scanned for affinity groups found in the collection of free text. Based on psychology experience [11], recommended practice is to use not more than 7 ± 2 affinity groups; otherwise humans are not able to consistently recognize affinity group patterns. The process is hard to automate. The affinity groups are the suspected business drivers; thus the selection of business drivers strongly impact analysis. Source for business driver suspects is the customer; they usually aren't known before the first NPS survey.

Answers can be positive or negative. This must be separated from pattern recognition; whether for instance 'fitness for purpose' is affirmed or challenged, in both cases it is deemed important enough to justify the given score. Thus verbatim analysis firstly returns results regarding low, medium, or high importance regardless whether the respondent made a positive or negative comment.

When performing verbatim analysis with 7 ± 2 business drivers, three points are assigned to each respondent, which can be distributed among the business drivers. This rule forces the analyst to investigate into the response focus in the free text supplied by customers. Points can be distributed 1-1-1 among three business drivers, or 2-1 if one driver is strongly accentuated; even 3-0 is possible if the answer is single focused on one aspect only. Weak affirmation or answers too general can reduce the total number of points given to two or one. The analysis returns a frequency metric that measures the importance given to each of the business drivers. More details and references about verbatim analysis can be found in [10].

When importance points are combined with affirmation, the verbatim analysis also evaluates customer's perception of performance. Negative affirmation yields negative points, and sometimes answers can contain both positive and negative affirmation, indicating customer's satisfaction with the performance of the supplier company meeting the respective business drivers. The limit of points given remains three. The satisfaction profile can become quite different from importance; the transfer function from the customer satisfaction profile to the NPS profile is a different one, based on a different frequency metric.

2.4 Analyzing Correlation with Referrals or Using a Constructive Approach?

Standard regression analysis compares the correlation index between referrals and referenced business drivers with the importance of the respective business driver. However, regression analysis can only prove certain events eventually happen

together. It does not establish a causal relationship between two events, and it does not help organizations uncover business drivers that were unknown before.

The constructive approach of detecting and selecting business drivers that explain the observed NPS profile allows managing and influencing customer response and loyalty towards a software product when developing or evolving it with new releases. Product marketing, product management, and software development benefit from this approach alike.

3 Transfer Functions

3.1 The Concept of Transfer Functions

Transfer functions origin from analysis of signals and systems such as Fourier transforms [8]. In DfSS they are instrumental when designing processes that deliver responses at defined variability levels [2]. They describe how a profile vector \underline{x} of controls impacts the response profile vector \underline{y} of the process: $\underline{y} = \mathcal{T}(\underline{x})$.

The process under question is the survey process \mathcal{T} . The result of the process is the survey NPS profile \underline{y} ; the controls are the business drivers, expressed as profile \underline{x} , that had been implemented in the software². The survey process \mathcal{T} is the transfer function that maps business drivers into NPS scores. It models the mind of respondents.

Given the survey \mathcal{T} , it is straightforward to compute the profile \underline{x} . These are the weights given by the respondents to the suspected business drivers as identified in verbatim analysis. Thus $\mathcal{T}(\underline{x})$ describes the NPS profile that the respondents would have given reflecting the suspected business drivers.

Thus, the difference between the vector profile $\mathcal{T}(\underline{x})$ and the observed profile \underline{y} is an indication whether the business drivers explain the observed response. It can be calculated as a vector difference $\|\mathcal{T}(\underline{x}) - \underline{y}\|$ and is called the *Convergence Gap*.

3.2 Using Eigenvectors

The method is taken from QFD. In QFD, expert teams predict the controls \underline{x} of a process by calculating the transpose $\underline{x} = \mathcal{T}^T(\underline{y})$ in a QFD workshop, and calculate the convergence gap to see whether the transfer function was assessed right.

Thus, if $\mathcal{T}(\underline{x}) = \mathcal{T}(\mathcal{T}^T(\underline{y})) = \underline{y}$, \underline{y} is an *Eigenvector* with *Eigenvalue* 1 of the composed symmetric and rectangular matrix $\mathcal{T} \cdot \mathcal{T}^T$, the result of matrix multiplication between \mathcal{T} and \mathcal{T}^T . Similar to Saaty's *Analytic Hierarchy Process* (AHP) [14], the eigenvector theory eliminates the measurement errors made in the survey and in the consequential analysis. For a theoretical background regarding QFD matrices, see [4].

² The profile \underline{x} of business drivers detected in the survey become in turn the profile of the response requested from the software product development process. This is the essential idea behind Six Sigma for Software and the Deming Chain; see [5], [6].

When deriving business drivers from an NPS survey result, a small convergence gap proves that the verbatim analysis of the responses is correct in the sense that it matches respondents' intent.

3.3 Attenuation of the Signal

In NPS surveys, passives also add a verbatim to their score, but the metric collapses the referral scores onto two percentages. High ratio of passives dampens the NPS while their verbatim signal is the same as from low ratio. Thus it is necessary to absorb the excess verbatim signal that originates from the total number of points given during verbatim analysis. **Table 2** shows attenuation for four verbatim counts, with 200 respondents and a ratio of two to one of promoters versus detractors.

Table 2. Sample Attenuation of Verbatim Signal

| Promoters | Passives | Detractors | Total | NPS | Signal Strength | Attenuation Factor | Adjusted Strength |
|-----------|----------|------------|-------|-----|-----------------|--------------------|-------------------|
| 133 | 0 | 67 | 200 | 33% | 600 | 1.00 | 600 |
| 90 | 77 | 33 | 200 | 29% | 600 | 0.62 | 369 |
| 67 | 100 | 33 | 200 | 17% | 600 | 0.50 | 300 |
| 44 | 133 | 22 | 200 | 11% | 600 | 0.33 | 200 |

The attenuation of the signal reflects the behavior of NPS for passives by weakening the total signal received. Thus the points given by passives contribute less to both NPS and the analysis points of the verbatim; the maximum strength being three points.

3.4 Finding Measurement Errors in the Verbatim Analysis

The convergence gap identifies measurement errors. When reading and analyzing the verbatim of customers, there is room for interpretation and this is often not done correctly. 'Correct' means that the points given to the verbatim explain the NPS as given by the respondent, closing the convergence gap.

If the convergence gap is not close enough to zero, either the business drivers assumed for the analysis are not able to explain the observed score – this is the case if gaps are important and not easily to close – or some verbatim have been assigned to the wrong business drivers. Revising the analysis knowing where the gap comes from usually uncovers such glitches in the analysis easily.

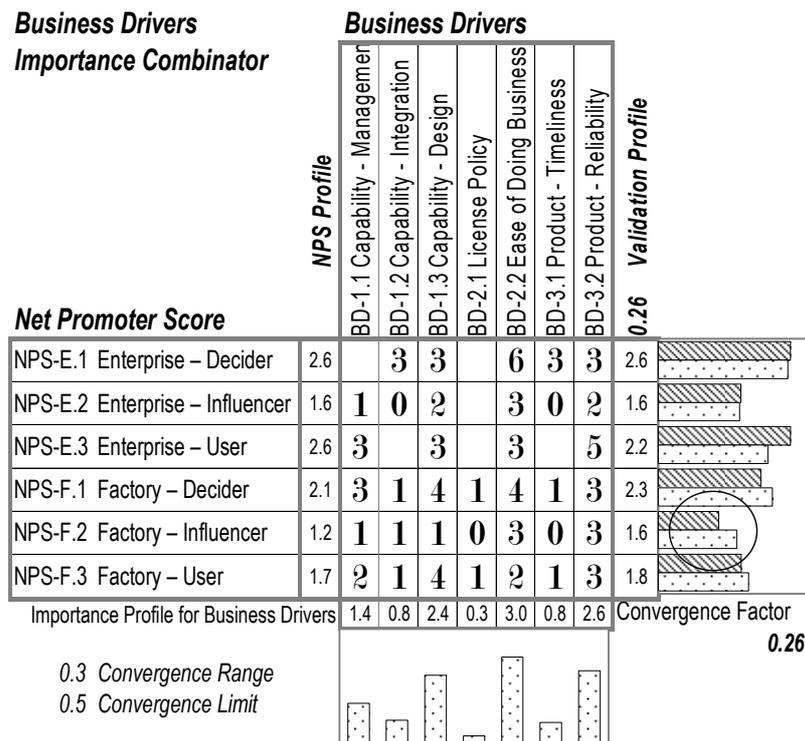
3.5 A Sample Analysis – Validating Importance

Table 3. Suspected Business Drivers for Verbatim Analysis in the Sample Case

| | <i>Topics</i> | <i>Attributes</i> |
|----------------|--------------------------|-------------------------------------|
| – Capabilities | Capability – Management | Ability to track and manage actions |
| | Capability – Integration | Integrate with other applications |
| | Capability – Design | Appearance and relevance of output |
| – Processes | License Policy | Licenses, contracts, T's & C's |
| | Ease of Doing Business | Support, responsiveness, competence |
| – Qualities | Product – Timeliness | Deliver new features in time |
| | Product – Reliability | High product quality, no bugs |

In this sample case, we use seven business drivers for verbatim analysis, see **Table 3**. The data are from an actual survey of a software company; the analysis is contrived. The business drivers reflect the customer's viewpoint; interestingly, it is not needed to use nationality or cultural influences as driving the response to the survey.

Fig. 2. QFD Transfer Function for Analyzing the Business Drivers' Importance for NPS Survey



For each of the NPS second question answers, the business driver points' frequency counts have been normalized in **Fig. 2** to the scale 0 – 9 known from QFD. Each of

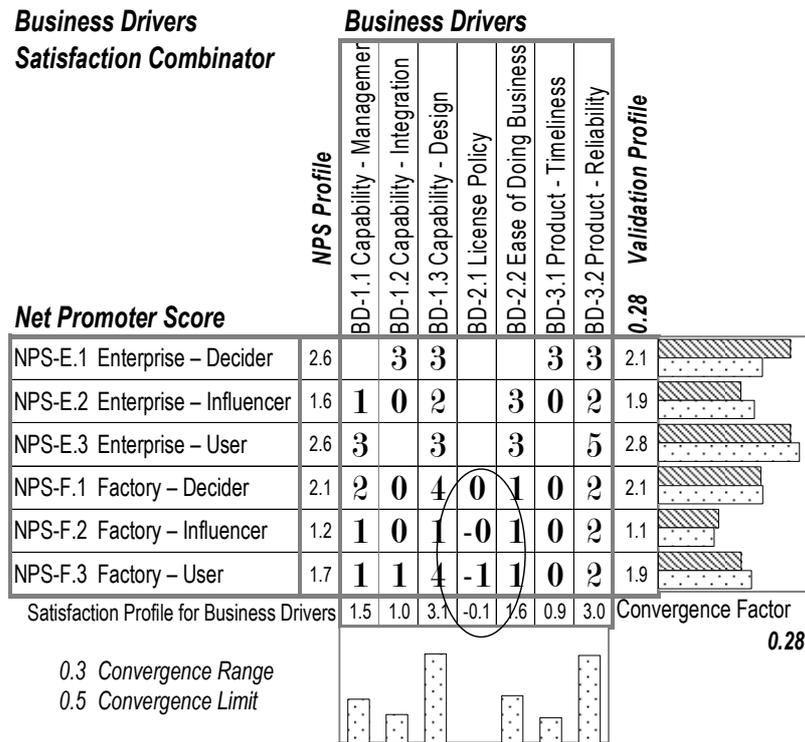
the respondent categories according **Table 1** is weighted equally, even if the number of respondents differs. These normalized weights for each pair of {NPS, Business Driver} fill a QFD matrix that defines the transfer function as a linear mapping. For the details of profile normalization, see [4].

The profile vector length is normalized to five; the convergence gap of 0.26 thus indicates that the profiles \underline{y} and $\mathcal{T}(\mathcal{T}^T(\underline{y}))$ differ by some 5% only, and thus the gap is within tolerance range, set by the convergence limit of 0.5. For a survey, this is within acceptable limits, proving that the seven business drivers are valid assumptions for the decision criteria that respondents used for choosing their score.

3.6 Validating Customer Satisfaction

There is a noteworthy deviation for NPS-F.2: Factory – Influencers in **Fig. 2** who’s actual NPS is below what had to be expected from business drivers’ importance. This hints at some specific dissatisfaction, as their NPS (see **Table 1**) does not match importance given to the business drivers. Despite its minor importance, BD-2.1: License Policy is a candidate for weak spots in software product management.

Fig. 3. QFD Transfer Function for Analyzing Satisfaction with Business Drivers



Thus, when re-doing the analysis with the performance evaluations in verbatim analysis (taking negative numbers for negative comments), the dissatisfaction conjecture for Factory – Influencers is confirmed. Note that matrix cells now can become negative, since performance can be evaluated negatively by respondents. Satisfaction with business drivers is particularly affected for BD-2.1: License Policy. Consequentially, business management of our software company may want to review license policy in order to improve loyalty among Factory – Influencers; however, taking into account that overall importance of this business driver is limited.

3.7 Innovation

Sometimes, initial selections for business drivers won't work and convergence gaps will not close. Thus, creativity is needed for guessing what other business drivers are needed to understand the customer's NPS voice.

This creativity is guided by the need to find suitable drivers that allow closing the gap, i.e., that explain importance and satisfaction profile measured in the free survey comments. Base for finding other business drivers is again the verbatim: very often, some notions pop up in the comments that don't fit well into initial drivers chosen for the analysis. That might cause the need to redo the analysis with different drivers. The effort to analyze hundred of voices is significant; however, with a little experience, it is doable since comments start repeating after a while.

Once the business drivers that fit the answers from customers are found, innovation in the software product should be guided and controlled by the new set of business drivers, as known for instance for Lean Six Sigma software development [6].

3.8 Findings from This Analysis

The data and the eigenvector analysis presented here support some of Satmetrix' claims. First, the verbatim analysis provides the transfer function that transfers the importance profile given by the survey respondents into the NPS. Second, another transfer functions also maps the satisfaction profile onto the NPS score. Satisfaction and referral probability are not the same indeed. The customer satisfaction is completely different from the NPS profile. This observation nicely explains Fred Reichheld's original finding. Both transfer functions return the NPS signal as a response, and thus allow for two different explanations both in terms of importance to the customer of, and satisfaction with, business drivers.

Using median of the survey scores by segments, or averages, does not yield any comparable clear signal for the respondents. Detecting a transfer function for the sample data presented here would have been impossible, even from satisfaction profile into NPS profile, because the median NPS in all segments was eight, and neither average nor geometric means did yield any other result. It looks like NPS models customers' gut feelings [7] against the product or service. Customers won't give scores according ratio scale criteria. Customers seem to flip-flop at certain scale positions from enthusiastic to neutral or reserved. This matches the experience made by business practitioners and seems to be modeled in the definition of the NPS [12].

Another question is whether both, importance and satisfaction, must yield valid explanations in order to validate business drivers; i.e., their respective transfer function found in verbatim analysis yield a small convergence gap each. While a transfer function must exist between satisfaction profile and NPS profile, and the mention of reasons for giving the score is most probably uncovering that transfer function, this is less obvious for importance. If the transfer function that maps importance onto NPS cannot be validated by its convergence gap, it means customers do not refer based on what they deem important. Such a gap indicates a major strategic failure in either the positioning of the software product, or in the target market approached with the product. A close investigation of both would be advisable in such case. Sometimes customers use a product for totally different purposes than what it was originally designed for.

3.9 Limitations

However, there is one inherent limitation when asking existing customers. They seldom give hints what new capabilities are needed for the software to meet future demands, if not asked directly with suggestions. This analysis of a relational-type survey alone cannot decide which capabilities to develop in future releases of the software product. Much better would be to ask customers after transactions such as solving a support case, providing training or installing new releases, or by explicitly asking customers and prospect for their preferences regarding new capabilities. Such transactional surveys can follow the same NPS patterns and use similar analysis methods, and in fact, they should complement them.

As a general caveat, and overall limitation of all such prediction attempts, future buying behavior might still depend from other factors such as economic outlook, or sudden change in paradigms, which are not detectable with verbatim analysis, and sometimes inherently unforeseeable.

4 Outlook and Conclusion

Business Drivers make software valuable for the customer. NPS measures the probability of referrals, resulting from satisfaction. The transfer function from satisfaction to NPS is found by verbatim analysis and validated by its eigenvector.

Eigenvector theory is a constructive, not empirical, method for analyzing cause-effect relationships. It is useful in many business areas, as presented here when analyzing business drivers that cause high NPS in surveys for a software vendor. Eigenvectors also enable software project managers to establish self-control for reaching business goals, such as to deliver the business driver's importance profile found in the survey, see [6]. The method is easy to use and requires no more investments than doing an NPS survey right, knowing your customer base, and requires some rather simple vector calculations involving linear algebra.

The analysis method is not specific to NPS surveys but particularly well suited. It can be easily adapted to any free-text analysis, provided the voice verbatim is linked to customer's business perceptions. Analyzing support cases is yet another very

promising application of the method. Since customer feedback is freely available through social media, and chat forums are a valuable source for business innovation, analyzing such input could become an important new application area for the eigenvector analysis method.

References

1. Akao, Y.: Quality Function Deployment: Integrating Customer Requirements into Product Design. Productivity Press, University Park, IL (1995)
2. Creveling, C.M., Slutsky, J.L., Antis, D.: Design for Six Sigma, Prentice Hall, NJ (2003)
3. Denney, R.: Working Smart to Deliver Quality. Booch–Jacobson–Rumbaugh – Series, Addison-Wesley, New York, NY (2005)
4. Fehlmann, Th.: The Impact of Linear Algebra on QFD. In: International Journal of Quality & Reliability Management, Vol. 21 No. 9, pp. 83--96. Emerald, Bradford, UK (2005)
5. Fehlmann, Th.: Six Sigma in der Software-Entwicklung, Vieweg-Verlag Wiesbaden, Germany (2005)
6. Fehlmann, Th.: Agile Software Projects with Six Sigma. In: 3rd European Research Conference on Continuous Improvement and Lean Six Sigma, Glasgow (2011)
7. Gigerenzer, G.: Gut Feelings. The Intelligence of the Unconscious. Viking, New York (2007)
8. Girod, B., Rabenstein, R., Stenger, A.: Signals and Systems, 2nd ed., Wiley (2001)
9. Likert, R.: A Technique for the Measurement of Attitudes. In: Archives of Psychology 140: 1--55 (1932).
10. Mazur, J., Bylund, N.: Globalizing Gemba Visits for Multinationals. In: Transactions from the 21st Symposium on Quality Function Deployment, Savannah, GA (2009)
11. Miller G., The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information, The Psychological Review, Vol. 63(2), pp. 81--97 (1956)
12. Reichheld, F.: The Ultimate Question: Driving Good Profits and True Growth. Harvard Business School Press, Boston, MA (2007)
13. Reichheld, F.: The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value. Paperback Ed. Harvard Business School Press, Boston, MA (2001)
14. Saaty, Th.L.: Decision-making with the AHP: Why is the principal eigenvector necessary. In: European Journal of Operational Research, Vol. 145, pp. 85--91, Elsevier Science B.V. (2003)