Improving the Predictive Validity of NPS in Customer Satisfaction Surveys

Gang Huang and Huai-lin Wang

User Experience Researcher, Alibaba Group No. 969, West Wenyi Road, 311121 Hangzhou, China kaer,hg@tmall.com, kongtuo.whl@tmall.com

Abstract. Though widely used, NPS has been challenged for its doubtful validity in predicting loyalty behaviors and business growth. Its arbitrary assignment of promoters/passives/detractors is one of the factors impacting its predictability. This paper addressed the validity issue by taking into account the scenarios NPS is used: (1) NPS for an organization in developing stage with more of new customers vs. a well established business with more of regular customers; (2) NPS in online survey vs. offline survey. By using the data of an online NPS survey for an online supermarket and tracking the purchase behaviors of all respondents before and after the survey, the authors found that NPS works better for new customers than regular ones in predicting repeat purchase behaviors and changes of purchase volume and value. In addition, in online surveys, a polarized segmentation of promoters/passives/detractors could effectively improve the predictive validity of NPS for new customers.

Keywords: NPS, Customer Satisfaction, Predictive Validity.

1 Background and Research Objectives

1.1 Merits and Criticisms on NPS

The Net Promoter Score (NPS) is one of the simplest customer satisfaction and loyal-ty measures, which only asks customers on a 0 to 10 rating scale: "How likely is it that you would recommend our company to a friend or colleague?" Based on their responses, customers can be categorized into one of three groups: Promoters (9-10 rating), Passives (7-8 rating), and Detractors (0-6 rating). Subtracting detractors from promoters then draws out the "net promoter" score as an estimate customer loyalty behaviors and business growth.

Thanks for the popularization of Fred Reichheld (2006), NPS has been widely adopted in different industries. It is easy to understand and makes intuitive sense. Compared to traditional customer-satisfaction measures such as American Customer Satisfaction Index (ACSI) model, NPS typically focuses on the measurement of WOM that has gained more and more attention with the popularization of Internet usage and the powerful influence of online WOM for organizations.

Despite the merits, NPS has received lots of criticisms that center on its failure to predict loyalty behaviors and business growth and its arbitrary segmentation of promoters/passives/detractors.

Its predictive validity is challenged mainly because this measure is attitudinal rather than behavioral and thus can only reflect present intention than future behavior. Keiningham et al. (2007) found that NPS does not perform better than the ACSI in predicting growth. By examining the robustness of different customer satisfaction and loyalty metrics, Keiningham et al. (2007) also suggested that the NPS alone would not serve as a single predictor of customers' future loyalty behaviors. Instead, multiple indicators performs significantly better.

Regarding its segmentation of three groups, Ken Roberts, CEO of Forethought Research Australia, said this rule-of-thumb score classes is not statically supported and may mask the important changes and potentially mislead management when the organization got a negative NPS whilst this may not be the case. On the other hand, the standard NPS question itself is unipolar (willingness to recommend) but its analysis treats it as bipolar (willingness to detract vs. promote) (blog.verint.com).

Given these criticisms, some researchers have explored different ways to enhance this measure. One of these endeavors is to merge NPS with other loyalty metric and generate a more synthesized measure. For instance, Owyang (2010) has combined NPS with other metrics to create the index of Total Social Customer Value (TSCV). The other part of these efforts focuses on optimizing the 11-point scale given that Reichheld shows flexibility on it. His original work showcases Enterprise Rent-A-Car by using a 5-point scale with 5 as promoters. Schneider, et al. (2007) has also recommended a 7-point bipolar scale for recommendation measuring.

Nevertheless, until now, there is no solid conclusion on either the best scale for NPS or the most effective assignment of promoters/passives/detractors. In addition, the golden rule, if there was, may vary in different survey scenarios which are closely related to business types surveyed, data collection methods used, etc.

1.2 NPS Practice in Online E-Commerce Customer Satisfaction Survey

Facing more and more fierce competition, monitoring customer satisfaction and loyalty is becoming more important for Chinese e-commerce giants. However, when advocating NPS to e-commerce practitioners, two challenges come up:

The first is, the operators in business with well-established e-commerce model and steady sales growth are less likely to accept NPS. The main concern is that their net promoter score may not have major changes over time given that their business is in relatively stable stage. Only those in newly developed e-commerce business pay much attention to NPS and even treat it as a KPI.

The second is that, in the 11-point scale, the point "10" gets much higher ratings than expected. In all the e-commerce customer satisfaction surveys we have done, it counts for around 50% of all responses whilst other ratings are very scattered (see Table 1). Meanwhile, ratings on 0 point are also slightly higher than that of 1, 2, or even 3, though not as obvious as 10 rating. This makes the business operators doubt about the validity of this measure and challenge the design of 11-point scale.

| E-commerce business | % of 10 rating in NPS question |
|----------------------------------|--------------------------------|
| A comprehensive B2C site | 52% |
| Large-scale household appliances | 61% |
| Furniture | 49% |
| Milk powder | 62% |
| Online supermarket | 47% |

Table 1. 10 ratings in e-commerce NPS surveys

Though all these surveys are about different e-commerce businesses, ranging from durables to FMCGs, the common thing is that they were all conducted online via e-mail recruitment and self-administrated web questionnaire. This may be caused the social desirability bias: when respondents fill in the questionnaire online without the presence of interviewers, they may be more likely to give extreme answers if they really think so, without any concern of possible reactions from others.

Bearing these challenges in mind, this paper comes up with the objectives to optimize the design of NPS in online e-commerce customer satisfaction surveys and improve its predictive validity.

2 Key Questions and Hypotheses

With the objectives mentioned above, two questions will be addressed in this paper with respective hypothesis listed below:

Q1. In what kind of e-commerce business scenarios NPS works better in terms of predictive validity?

H1: Our hypothesis is that NPS works better for newly developed business. The first reason is that the likelihood of recommendation fits well with the business objective of growing reputation and enlarging market penetration for this kind of organizations. More importantly, this kind of organizations consists more of new users than regular users. The new users may turn into regular users or lapsed users. Therefore, their responses on NPS may better reflect their future loyalty behaviors. However, for the regular users, their purchase behaviors may not be changed even if they give a very low score in NPS question.

This question is hence turned into a more specific sub-question: does NPS have better predictive validity for new users or regular users? The sub-hypothesis is NPS is more valid in predicting new users retention behaviors than regular users' loyalty.

Q2. What is the better assignment of promoters and detractors when NPS surveys is conducted online?

H2: Our previous experience has led to a hypothesis that the segmentation of promoters/passives/detractors should be more polarized in online surveys, i.e. promoters should be defined with only 10 rating instead of 9-10 rating, while detractors with 0-5 rating or lower instead of 0-6 rating.

3 Research Methodology

To verify our hypothesis, a research-on-research study was conducted by taking an online NPS survey for online supermarket as the showcase.

3.1 Rationale of the Overall Research Design

The core of this research is to test the predictive validity of NPS. Therefore, the purchase behaviors of each respondent in this survey were monitored over 6 months, with 3 months before the survey and 3 moths after. Linking the NPS rating of each respondent with their behavioral changes before the after the survey, we were able to examine the correlations between recommendation intention and purchase behaviors.

To verify the first hypothesis, we tested the predictive validity of NPS for new users and regular users respectively. Each of the testing was not only to compare the behavioral changes between promoters and detractors, but also to do such comparisons between promoters vs. passives, and passives vs. detractors. This is to verify the existence of linear correlation between NPS rating and loyalty behavior. Ideally, promoters should perform better in loyalty behaviors than passives and passives better than detractors.

To verify the second hypothesis, we compared the predictive validity of two segmentations. The first is the standard segmentation advocated by Fred Reichheld as a benchmark. The second is a set of polarized segmentations for comparison. The specific design of the alternative segmentations allows some flexibility for new users and regular users, depending on the actual data got from the survey. This would help us to find out the optimal segmentation in different scenarios.

In addition, to strengthen the second hypothesis, we also looked at the different reasons for recommendation between promoters with 9 and 10 rating. Accordingly, the detraction points were also compared across the detractors with different ratings.

3.2 Choice of the Right Case

To make sure that we can observe the behavioral changes in an acceptable period, the case adopted in this study should be about the organizations with relatively higher purchase frequency. Therefore, we choose one of the leading online supermarkets in China as an example. It sells daily consumption products online, including foods & beverages, personal case and house cleaning products, etc.

3.3 Data Collection

The customer satisfaction survey was conducted online in the last week of August 2013. An email invitation was sent to the current users of this online supermarket with a hyperlink of a web questionnaire. In total, 3,106 valid responses were collected.

The classic NPS question with 11-point scale is adopted in the questionnaire, followed by a set of questions about the reasons to recommend for promoters and the reasons to detract for detractors respectively. Close-end multiple-choice questions were used in this probing session, with a list of attributes generated from qualitative studies as the choices. 30 promotion points and 31 detraction points were covered.

3.4 Sample Coverage

All the respondents surveyed must be customers who had finished at least one purchase order in the past 3 months before the survey, i.e. May – Aug. 2013.

Both the new users and regular users were covered. New users were defined as those who had not bought anything from this site in the 3 months before May, i.e. Feb. - May 2013, whilst the regular users were purchase contributor along the past 6 months including May - Aug. and Feb. - May 2013 (See Table 2 for illustration)¹.

| | Feb May. 2013 | May. – Aug. 2013 | 22 nd -27 th , Aug. 2013 |
|---------------|----------------|------------------|--|
| New users | Non-purchasers | Purchasers | Survey period |
| Regular users | Purchasers | Purchasers | Survey period |

Table 2. Definition of new users and regular users

3.5 Indicators for Predictive Validity Test

The purchase behaviors we monitored covered repeat purchase, volume changes and value changes, which are closely related to business growth.

- Repeat purchase rate (Repeat rate): the percentage of respondents who have purchase in the subsequent 3 months after the survey (Sep. Nov. 2013).
- % of respondents with volume increase (Vol. increase)²: the percentage of respondents who finished more purchase orders in the subsequent 3 months after the survey (Sep. Nov. 2013) than that in the 3 months before the survey (May Aug. 2013).
- % of respondents with value increase (Val. increase): the percentage of respondents who have spend more in this online supermarket in the subsequent 3 months after the survey than that in the 3 months before the survey.

¹ This definition is aligned with business operators. However, we also checked the purchase behaviors 3 months earlier and found that 86% of new users did not purchase anything during Nov. 2012 – Feb. 2013 but only 46% of regular users had no purchase in this period.

² The purchase volume may be increased, decreased, or kept same. In this survey, only very few respondents have exactly the same purchase volume between the first and second 3 months, therefore, unless pointed out, usually the test on volume increase is almost equal to that of volume decrease. It is the same case for the indicator of value increase.

4 Research Result

In this survey, respondents with higher ratings in NPS question are more likely to repeat purchase, increase purchase volume and value in the subsequent 3 months. Looking at the 11 points of the scale, 10 and 5, instead of 9 and 6, appear to be the turning points in this curve (see Table 3). This applies to total respondents and new users. However, for the regular users, rating 9 is still an important turning point.

With this respective, the polarized segmentation of promoters/passives/detractors tested in the following sessions were designed as below:

- Total users: Promoters (10 rating), Passives (6-9 rating), Detractors (0-5 rating)
- New users: Promoters (10 rating), Passives (6-9 rating), Detractors (0-5 rating)
- Regular users: Promoters (9-10 rating), Passives (6-8 rating), Detractors (0-5 rating)

| User | T 1' 4 | NPS rating | | | | | | T-4-1 | |
|------------------|---------------|------------|-----|-----|-----|-----|-----|-------|-------|
| group | Indicators | 0-43 | 5 | 6 | 7 | 8 | 9 | 10 | Total |
| | Repeat rate | 55% | 58% | 64% | 67% | 67% | 71% | 76% | 70% |
| Total users | Vol. increase | 38% | 40% | 47% | 49% | 50% | 53% | 56% | 52% |
| | Val. increase | 37% | 43% | 50% | 53% | 53% | 55% | 59% | 54% |
| | Repeat rate | 49% | 52% | 56% | 60% | 60% | 60% | 69% | 62% |
| New | Vol. increase | 32% | 39% | 44% | 45% | 46% | 45% | 51% | 46% |
| users | Val. increase | 32% | 40% | 48% | 47% | 48% | 46% | 52% | 48% |
| Regular users | Repeat rate | 71% | 75% | 81% | 82% | 79% | 87% | 86% | 83% |
| | Vol. increase | 52% | 43% | 53% | 58% | 58% | 65% | 64% | 61% |
| | Val. increase | 50% | 48% | 57% | 68% | 62% | 69% | 68% | 64% |

Table 3. Behavioral changes across the respondents with different NPS ratings

4.1 The Predictive Validity at Total Level

By using the standard segmentation, NPS shows very good predictive validity in general. In the subsequent 3 months after the survey, the tendency of promoters to repeat purchase, increase purchase volume and value are significantly higher than passives, and passives significantly higher than detractors (see Table 4).

Nevertheless, the polarized segmentation was still examined and it turned out to yield perfect prediction for purchase behaviors as well, with its t-test result is slightly more significant than that of standard segmentation (see Table 5).

³ The 0-4 rating was analyzed as whole due to the small sample size of individual rating.

| Tests | Indicators | Test groups | s (standard) | T-test | result (2-t | ailed) |
|--------|---------------|-----------------|-----------------|--------|-------------|--------|
| | | Detractors(0-6) | Passives(7-8) | t | df | Sig. |
| Test 1 | Repeat rate | 59% | 67% | -2.992 | 1160.0 | 0.003 |
| | Vol. increase | 42% | 50% | -3.002 | 1210.5 | 0.003 |
| | Val. increase | 43% | 53% | -3.581 | 1204.3 | 0.000 |
| | | Passives(7-8) | Promoters(9-10) | t | df | Sig. |
| m . 2 | Repeat rate | 67% | 75% | -4.164 | 1617.8 | 0.000 |
| Test 2 | Vol. increase | 50% | 56% | -2.885 | 1728.6 | 0.004 |
| | Val. increase | 53% | 58% | -2.453 | 1720.9 | 0.014 |
| | | Detractors(0-6) | Promoters(9-10) | t | df | Sig. |
| | Repeat rate | 59% | 75% | -6.798 | 869.5 | 0.000 |
| Test 3 | Vol. increase | 42% | 56% | -5.856 | 968.7 | 0.000 |
| | Val. increase | 43% | 58% | -6.116 | 958.4 | 0.000 |

Table 4. The predictive validity test for total users with standard segmentation

Table 5. The predictive validity test for total users with polarized segmentation

| Tests | Indicators | Test groups | (polarized) | T-test result (2-tailed) | | |
|--------|---------------|-----------------|---------------|--------------------------|--------|-------|
| | | Detractors(0-5) | Passives(6-9) | t | df | Sig. |
| Т4 1 | Repeat rate | 57% | 67% | -3.526 | 604.1 | 0.000 |
| Test 1 | Vol. increase | 39% | 50% | -3.74 | 642.7 | 0.000 |
| | Val. increase | 40% | 53% | -4.464 | 639.3 | 0.000 |
| | | Passives(6-9) | Promoters(10) | t | df | Sig. |
| Т4-2 | Repeat rate | 67% | 76% | -4.912 | 2585.3 | 0.000 |
| Test 2 | Vol. increase | 50% | 56% | -3.29 | 2663.5 | 0.001 |
| | Val. increase | 53% | 59% | -2.946 | 2659.1 | 0.003 |
| | | Detractors(0-5) | Promoters(10) | t | df | Sig. |
| Test 2 | Repeat rate | 57% | 76% | -6.706 | 541.8 | 0.000 |
| Test 3 | Vol. increase | 39% | 56% | -6.043 | 605.7 | 0.000 |
| | Val. increase | 40% | 59% | -6.525 | 1837 | 0.000 |

4.2 The Predictive Validity for New Users

Looking at the new users, overall the standard segmentation has yield good predictive validity, but it is better for detractors than promoters. The repeat purchase rate, volume and value increase of detractors are significantly lower than any of the other two groups. However, the volume and value changes between promoters and passives have no major difference though they differ in repeat purchase rate (see Table 6).

It should be noted that the tested volume increase between detractors and passives is not that obvious (sig.=0.046). Meanwhile, the test on volume decrease shows different results: no significant difference between them was found (detractors: 60%, passives: 53%, t=1.94, df=867.7, sig.=0.053).

| Tests | Indicators | Test groups | s (standard) | T-test | result (2-t | tailed) |
|--------|---------------|-----------------|-----------------|--------|-------------|---------|
| Test 1 | | Detractors(0-6) | Passives(7-8) | t | df | Sig. |
| | Repeat rate | 53% | 60% | -2.31 | 849.4 | 0.021 |
| | Vol. increase | 39% | 45% | -1.995 | 870.2 | 0.046 |
| | Val. increase | 40% | 47% | -2.184 | 868.5 | 0.029 |
| | | Passives(7-8) | Promoters(9-10) | t | df | Sig. |
| T4 2 | Repeat rate | 60% | 68% | -2.97 | 1114.8 | 0.003 |
| Test 2 | Vol. increase | 45% | 50% | -1.754 | 1162.7 | 0.080 |
| | Val. increase | 47% | 52% | -1.585 | 1159.1 | 0.113 |
| Test 3 | | Detractors(0-6) | Promoters(9-10) | t | df | Sig. |
| | Repeat rate | 53% | 68% | -5.184 | 695.1 | 0.000 |
| | Vol. increase | 39% | 50% | -3.799 | 753.3 | 0.000 |
| | Val increase | 40% | 52% | -3.852 | 748 5 | 0.000 |

Table 6. The predictive validity test for new users with standard segmentation

Fortunately, the polarized segmentation shows perfect predictive validity. In the subsequent 3 months, promoters (10 rating) shows significantly higher repeat purchase rate, purchase volume and value increase than passives (6-9 rating), and passives (6-9 rating) higher than detractors (0-5 rating) in all these indicators (see Table 7). This means, the polarized segmentation does improve the validity.

| Test | Indicators | Test groups | (polarized) | T-test | result (2-t | tailed) |
|--------|---------------|-----------------|---------------|--------|-------------|---------|
| Test 1 | | Detractors(0-5) | Passives(6-9) | t | df | Sig. |
| | Repeat rate | 51% | 59% | -2.49 | 461.0 | 0.013 |
| | Vol. increase | 36% | 45% | -2.618 | 482.5 | 0.009 |
| | Val. increase | 37% | 47% | -3.003 | 482.1 | 0.003 |
| T | | Passives(6-9) | Promoters(10) | t | df | Sig. |
| | Repeat rate | 59% | 69% | -4.032 | 1660.6 | 0.000 |
| Test 2 | Vol. increase | 45% | 51% | -2.292 | 1676.6 | 0.022 |
| | Val. increase | 47% | 52% | -2.201 | 1675.9 | 0.028 |
| | | Detractors(0-5) | Promoters(10) | t | df | Sig. |
| Test 3 | Repeat rate | 51% | 69% | -5.292 | 431.8 | 0.000 |
| | Vol. increase | 36% | 51% | -4.288 | 475.0 | 0.000 |
| | Val. increase | 37% | 52% | -4.608 | 472.7 | 0.000 |

Table 7. The predictive validity test for new users with polarized segmentation

4.3 The Predictive Validity for Regular Users

Regarding the regular users, however, the predictability of the NPS with standard segmentation seems worse than that of new users. Although the promoters have significant behaviors changes compared with detractors in terms all the three indicators, the difference of promoters vs. passives, and passives vs. detractors are not obvious (see Table 8). For example, even among the detractors, 76% of them repeated purchase in the subsequent 3 months and nearly half of them increased their purchase

volume, which has no major difference from the passives. The only difference is that less of detractors increased purchase value than passives, but the absolute rate is still very high (51%). Meanwhile, promoters did not show higher volume and value increase than passives, though they enjoyed higher repeat purchase rate.

| Tests | Indicators | Test groups | Test groups (standard) | | | tailed) |
|--------|---------------|-----------------|------------------------|--------|-------|---------|
| Test 1 | | Detractors(0-6) | Passives(7-8) | t | df | Sig. |
| | Repeat rate | 76% | 80% | -1.048 | 309.3 | 0.295 |
| | Vol. increase | 49% | 58% | -1.919 | 325.0 | 0.056 |
| | Val. increase | 51% | 64% | -2.589 | 317.8 | 0.010 |
| | | Passives(7-8) | Promoters(9-10) | t | df | Sig. |
| T | Repeat rate | 80% | 86% | -2.184 | 512.8 | 0.029 |
| Test 2 | Vol. increase | 58% | 64% | -1.863 | 562.5 | 0.063 |
| | Val. increase | 64% | 68% | -1.271 | 562.7 | 0.204 |
| | | Detractors(0-6) | Promoters(9-10) | t | df | Sig. |
| Test 3 | Repeat rate | 76% | 86% | -2.778 | 213.6 | 0.006 |
| | Vol. increase | 49% | 64% | -3.594 | 235.5 | 0.000 |
| | Val. increase | 51% | 68% | -3.849 | 231.6 | 0.000 |

 Table 8. The predictive validity test for regular users with standard segmentation

By using the refined segmentation with detractors polarized to the lower ratings of 0-5, the predictability does not change much except the improvement in predicting volume increase for promoters vs. passives (see Table 9).

The implication is, for regular users, their recommendation has little relation with their future purchase behaviors. Detractors, passives, and promoters all enjoy very high repeat purchase rate, volume increase and value increase. It is probably because buying groceries in this online supermarket has become a routine for them and they would not change their purchase channel even if they have some complaints.

| Tests | Indicators | Test groups | (polarized) | T-test result (2-taile | | |
|---------|---------------|-----------------|-----------------|------------------------|-------|-------|
| | | Detractors(0-5) | Passives(6-8) | t | df | Sig. |
| | Repeat rate | 73% | 80% | -1.463 | 164.7 | 0.145 |
| Test 1 | Vol. increase | 47% | 57% | -1.934 | 463.0 | 0.054 |
| | Val. increase | 49% | 63% | -2.571 | 174.0 | 0.011 |
| | | Passives(6-8) | Promoters(9-10) | t | df | Sig. |
| Т4-2 | Repeat rate | 80% | 86% | -2.266 | 639.7 | 0.024 |
| Test 2 | Vol. increase | 57% | 64% | -2.211 | 698.6 | 0.027 |
| | Val. increase | 63% | 68% | -1.676 | 696.7 | 0.094 |
| Test 3 | | Detractors(0-5) | Promoters(9-10) | t | df | Sig. |
| | Repeat rate | 73% | 86% | -2.827 | 129.9 | 0.005 |
| | Vol. increase | 47% | 64% | -3.423 | 141.2 | 0.001 |
| | Val. increase | 49% | 68% | -3.754 | 139.4 | 0.000 |

Table 9. The predictive validity test for regular users with polarized segmentation

4.4 An Optimized Comparison between New Users and Regular Users

The above comparison between new users and regular users may be challenged if these two groups have different profiles. It is true that the new users and regular users are different in some aspects: there are more of females (56%), married people (59%) and elder people (21% aged above 35 y.o.) among regular users, whilst more of young (26% aged 25 y.o. and below) and single (54%) people within the new users.

To avoid the impacts of demographic difference on the test result, weighting on age and gender distributions within each of new users and regular has been done, which yielded similar sample structure in calculation.

Looking at the purchase behavior changes across promoters, passives, and detractors again and compare these changes between new users and regular users, we got the same result as before:

For the new users, the polarized segmentation again improved the predictive validity in all the three behavioral indicators; whilst for the regular users, both the standard and polarized segmentation fails to bring a perfect prediction on future purchase behaviors.

Therefore, even when new users and regular users have similar profiles, NPS still works better for new users than regular users⁴.

4.5 The Difference in Promotion and Detraction Points

Given that the polarized segmentation may yield a better predictive validity, especially for new users, the promotion points between respondents with 9 and 10 rating may be different, and the detraction points between respondents with 0-5 and 6 rating may be different. Part of the hypothesis is proved in this research.

For the total respondents, respondents with 9 and 10 ratings have different agreements on 6 out of all the 30 promotion points, which means the strength of this online supermarket are different for promoters rating 9 and 10. However, in the 31 detraction points, only 2 of them have received different agreements between detractors rating 0-5 and 6.

For the new users, there come the similar findings: 5 out of 25 promotion points are different between promoters rating 9 and 10, whilst only 1 out of 20 detraction points differs between detractors rating 0-5 and 6.

For the regular users, however, the case is quite different. Only 1 out of 25 promotion points have different agreements between promotes of 9 and 10 rating. For the detraction points, such comparison is not feasible due to small sample size.

5 Conclusions and Discussions

5.1 Conclusions and Recommendations

In general, the hypothesis addressed at the beginning of this paper are mostly supported:

⁴ Given the limited length of paper, test result with weighted data is not shown here. Please contact the authors if necessary.

NPS has relatively better predictive validity on subsequent purchase behaviors for new users than regular users. Considering that 63% of the respondents are new users in this study, it is not surprising to see the good predictability of NPS at total level. Therefore, it is recommended that NPS should be more promoted in business at developing stage. If used in the mature business, it is suggested to separate the analysis of new users and regular users whenever sample size allows, which will lead to different marketing implications.

To some extend, it is also proved that, in online surveys, a polarized segmentation of promoters/passives/detractors could make the NPS working better in predicting future loyalty behaviors, especially for new users. However, considering that this result is coming from only one case, we would rather suggest using standard segmentation at the very beginning and being open to alternative segments when necessary. Nevertheless, we do suggest separate the analysis of promotion points between respondents with 9 and 10 rating, and the analysis of detraction points across different the detractors as well. This may help to better understand the strength and weakness of our organization for different customers and thus yield more specific and practical marketing measures.

5.2 Discussions and Suggestions for Future Researches

Although most of the hypothesis has been proved in this case study, there remains some room for further exploration given the limitations of this research.

Firstly, the result of this research is only drawn from a customer satisfaction survey on FMCG e-commerce site. Whether this result is applicable for other categories or other business types is doubtful.

Secondly, this study only tested the predictive validity of purchase behaviors in 3 months right after the survey. However, whether NPS can predict behaviors in a longer period? In predicting behavioral changes of a longer period, will NPS works differently for new users and regular users? Can we suppose that NPS can better predict long-time changes of regular users and short-time changes of new users? All these questions are yet to be explored but can be done by continue tracking the purchase behaviors for a longer time.

Finally, the showcase in this study only contains one wave of survey. Therefore, it is not feasible to monitor the changes of "Net Promoter" score between new users and regular users, as well as its relation to their purchase behaviors. If possible, that will further strengthen the conclusion of this paper.

References

- Blog of Verint, http://blog.verint.com/net-promoter-score-npscriticisms-and-best-practices
- Keiningham, T.L., Cooil, B., Andreassen, T.W., Aksoy, L.: A Longitudinal Examination of Net Promoter and Firm Revenue Growth. Journal of Marketing 71(3), 39–51 (2007)

- Keiningham, T.L., Cooil, B., Aksoy, L., Andreassen, T.W., Weiner, J.: The Value of Different Customer Satisfaction and Loyalty Metrics in Predicting Customer Retention, Recommendation, and Share-of-wallet. Managing Service Quality 17(4), 361–384 (2007)
- 4. Owyang, J.: Enhancing Net Promoter Score (NPS) with Total Social Customer Value (TSCV). Web Strategy (June 20, 2010), http://www.web-strategist.com/blog/2010/06/20/enhancing-net-promoter-score-nps-with-total-social-customer-value-tscv/
- 5. Reichheld, F.F.: The Ultimate Question: Driving Good Profits and True Growth. Harvard Business School Press, Boston (2006)
- Schneider, D., Berent, M., Thomas, R., Krosnick, J.: Measuring Customer Satisfaction and Loyalty: Improving the "Net-Promoter" Score. Paper Presented at the Annual Conference of the World Association for Public Opinion Research (WAPOR), Berlin, Germany (2007)