

Samsung: Do AI generated replies on positive reviews work?

Context

A lot of reviews

Samsung BeNeLux receives a lot of reviews, up to 50.000 a year. This is because of a proactive approach that's been developed for a number of years, resulting in best in class numbers worldwide. From research we know that 98% of consumers read reviews before making a purchase¹. So focussing on reviews is a good idea.

But since the number of reviews is so huge, it's impossible for the local team to reply to every review, so naturally only the negative ones are replied. Leaving the vast majority of replies unanswered, which didn't feel right. It just wasn't feasible.

AI to the rescue

To solve the issue with unreplied reviews we created an AI bot trained with the Samsung product data and some more data to be able to respond to reviews with a thank you, an answer to a question or even upsell opportunities.

At the beginning of 2025 we started replying using the AI bot to reviews within the 3-5 star range, those that are not handled manually. This is in 2 languages, Dutch and French. Initially we started with the mobile product group to later extend to the AV category as well. This bot is running smoothly, leaving no reply unanswered.

Integration with Bazaarvoice

To ensure authenticity and credibility, we integrated everything within Bazaarvoice, a trusted third-party review platform. The custom-trained AI bot was connected to a middleware that integrates with the Bazaarvoice engine via APIs. Whenever a 3-, 4-, or 5-star review is submitted on selected products, the review is downloaded into our system, the AI generates a fitting reply, and it is pushed back after a credible time. This integration allows Samsung to respond at scale while maintaining the trustworthiness and visibility of reviews.

¹<https://www.forbes.com/councils/forbesbusinessdevelopmentcouncil/2024/07/11/how-reviewsand-ratings-affect-clients-buying-decisions/>

Does it work?

Of course we have a natural feeling that responding to a good review is something positive. This is basic human psychology. The question is if it also works in terms of business metrics.

These are the results that were quickly measured:

- The average time on site doubled from 31 seconds to over a minute
- Organic views to the review pages multiplied by 260%

It seems having more relevant content on the website is something that keeps prospects longer hooked. Similarly, since we drop in relevant keywords in the replies, the findability improves. This seems logical and the numbers prove the assumptions are proven to be true. But does it also have an effect on the people leaving reviews. This we had to investigate.

A survey

To find out the impact on the reviewers themselves we conducted a survey. Bear in mind that we only surveyed the positive reviewers, 4 and 5 stars. This is a disadvantage in terms of survey results because we're already in the high regions, so differences are typically smaller there. Our initial idea to do a NPS - Net Promoter Score, where the percentage of detractors (scoring 6 or below) is subtracted from the promoters (scoring 9 or 10). This turned out to be unusable because in this selection of people (4 and 5 star reviews) there were hardly any detractors.

But still we structured the questionnaire the same way where we asked respondents to answer questions on a 0-10 scale. In the NPS fashion we asked the first question as: how likely are you to recommend (..) to friends and family, whereas a 0 is "very unlikely" and a 10 is "very likely".

We selected 1000 people that recently left a reply. 500 of which did receive a reply and 500 who didn't. Also 500 who gave a 4-star rating and 500 with a 5-star rating. Equally spread between Belgium and the Netherlands. A very balanced and big enough base to start from.

To incentivize the participation we raffled 2 €500 gift cards to the Samsung webshop, but made clear that (positive) answers to the questionnaire are not related to the chances of winning such a gift card.

These are the questions that we asked, all on a 0-10 scale:

1. How likely is it that you would recommend the product you wrote a review for to friends and family?
2. How likely is it that you would recommend Samsung as a brand to friends or family?
3. How likely is it that you would write another review on Samsung.com in the future?
4. How satisfied are you with how Samsung handled your review?
5. How well does Samsung's follow-up align with what you expect from a brand?
6. How well do you think Samsung listens to its customers?
7. To what extent do you feel that Samsung values your opinion when you write a review?

The survey ran from September 9th 4PM local time until September 12th eod.

After sending we got these email results:

- 99,6% delivery rate (very accurate list)
- 453 (45,48%) open rate
- 277 (27,81%) clicked

Results in the survey

Out of the people selected for the survey 280² (28%) respondents answered. Of these 130 respondents received a reply on their review, and 150 didn't. Across the board we see significantly better scores when answered by respondents who got an AI reply. Here are the overall results.

General	q1	q2	q3	q4	q5	q6	q7
No reply - avg scores on 10	8.09	8.08	7.98	7.68	7.79	7.71	7.57
AI reply - avg scores on 10	9.32	9.50	8.81	9.03	9.12	8.63	9.01
Increase	13%	15%	9%	15%	15%	11%	16%

Example to read this, on question 1 (q1) "How likely is it that you would recommend the product you wrote a review for to friends and family?" respondents who didn't get an AI reply averaged 8,09 out of 10 on this question, while those who answered gave an average response of 9,32 out of 10. A 13% increase on already high numbers.

² This is a higher number than the unique clicks and is due to privacy settings on the user side that prevent accurate click measuring. Each link was unique so no errors are possible.

Here's the full breakdown with results compared to the control group that didn't get a reply:

1. How likely is it that you would recommend the **product** you wrote a review for to friends and family? - 13% increase
2. How likely is it that you would recommend Samsung as a **brand** to friends or family? - 15% increase
3. How likely is it that you would write **another review** on Samsung.com in the future? - 9% increase
4. How satisfied are you with how Samsung **handled your review**? - 15% increase
5. How well does Samsung's **follow-up align with what you expect** from a brand? - 15% increase
6. How well do you think Samsung **listens** to its customers? - 11% increase
7. To what extent do you feel that Samsung **values your opinion** when you write a review? - 16% increase

When we drill down deeper and separate the respondents in groups based on the number of stars their review had (4 or 5) the results are even more profound, with an even stronger impact on the 4-star reviewers.

4-star reviews	q1	q2	q3	q4	q5	q6	q7
No reply - avg scores on 10	7.16	6.90	7.63	7.05	7.32	6.68	6.21
AI reply - avg scores on 10	9.08	9.67	8.08	8.25	8.17	7.83	8.42
Increase	21%	29%	6%	15%	10%	15%	26%

5-star reviews	q1	q2	q3	q4	q5	q6	q7
No reply - avg scores on 10	8.23	8.25	8.03	7.77	7.86	7.85	7.75
AI reply - avg scores on 10	9.34	9.48	8.88	9.11	9.22	8.71	9.07
Increase	12%	13%	10%	15%	15%	10%	15%

³ 7 and 8 in NPS are considered neutrals and don't affect the score

Promoters

Net Promoter Score (NPS) matters because it is a key predictor of business growth and a simple, effective measure of customer loyalty. A high NPS score indicates strong customer satisfaction, leading to positive word-of-mouth, increased customer retention, and higher revenue.

Since the NPS is the percentage of promoters (score 9 or 10) minus the percentage of detractors (scoring 6 or less) there is no use in calculating the NPS since we hardly have any detractors as said. This is logic since we only questioned high score reviewers.

Still the principle of NPS is making sure to have enough promoters, and this we can distill from the dataset and make a comparison between the two groups. We can also see if we turned people into promoters, because the group that scored 4 stars out of 5 are technically not promoters as that translates to a 8 out of 10, whereas a promoter is a 9 or a 10³.

PROMOTORS	#	Product promoters (q1)	%	Brand promoters (q2)	%
No reply	150	45	30%	50	33%
Score 4	19	0	0%	0	0%
Score 5	131	45	34%	50	38%
AI reply	130	106	82%	113	87%
Score 4	12	11	92%	12	100%
Score 5	118	95	81%	101	86%

Overall we see that the control group that didn't get an AI reply has 30% promoters on the product question ("Would you recommend the product ..") and 33% promoters on the brand question ("Would you recommend Samsung as a brand ..").

If we dig deeper we see that all these promoters are in the group that gave a 5-star rating.

If we compare these results with the results of the group that got an AI reply the story is quite different. Overall we see that 82% of respondents are confirmed as promoters on the product question, and even 87% on the brand question. What's more, we see that we converted almost all respondents that initially gave a 4-star review (so, no promoters) have been converted to promoters.

Conclusion

Replying to positive reviews clearly works, and AI is a great technology to handle the high number of reviews that Samsung BeNeLux generates. Overall this is a great project that seems to generate brand advocates, and is highly scalable over product categories, languages and countries.

Report made and numbers validated by:

A handwritten signature in blue ink, appearing to read 'M. Wantens', with a long horizontal stroke extending to the right.

Michael Wantens
Digital ATL Manager Samsung Mobile

A handwritten signature in blue ink, appearing to read 'S. Verbruggen', with a long horizontal stroke extending to the right.

Steven Verbruggen
Managing Director AdSomeNoise
steven@adsomenoise.com