

"Can You Give Me A Five Star Please?" - How Human Interactions Influence Online Service Ratings

Jiayi Wu¹

¹University of Toronto, Toronto, Canada, M1C 1A4

Abstract. When talking about giving “five stars” in the quality of human interaction between service provider and receiver should not be omit. In the field of marketing studies, this is relatively less talked about. This study examines the relationship between the existence of human interaction in the service experience and the consumer ratings in online services, how the quality of the interaction affect the ratings and to what extend. Three studies were conducted to test out if consumers are more likely to rate when the human interactions take place, and they are more likely to rate when the interaction is not favourable. Furthermore, they are more likely to rate to the extreme. The results are predicted to agree with the hypotheses. The study results are valuable for brands to interpret the received ratings and to work to provide a better user experience.

1. Introduction

Online services are becoming more commonly used in our daily life. Home delivery service platforms like UberEats and Instacart allow us to use mobile phone applications or websites to help us get our food or groceries without us ever getting out of our homes; platforms usually hire independent contractors to get your food or buy your groceries and deliver them to your doorstep. Other kinds of service providers like Uber and Lyft enable us to find a ride nearby on our phones, and the independent contractors would give us the ride to wherever we want. These service platforms would always provide a rating system to collect consumer satisfaction rates after the service is rendered. The ratings are crucial to both the media and the independent contractors; platforms rely on the rating to gather consumer experience data; contractors depend on good ratings to get distributed with better service deals. Studying consumers' rating behaviour is essential for brands to build relations with their customers and help them interpret the ratings. Especially now the pandemic hits, more and more rely on this kind of non-contact service, and more service providers like this would appear and thrive. Existing platforms should differentiate themselves in order to stay comparatively competitive and gain more market shares. However, gathering quality consumer reviews would not be easy sometimes.

Typically, the consumers would not even give a rating and just continue placing their next order. Consumers will willingly leave their ratings if the experience is impressive, whether negative or positive. We might think the impressive services with high ratings are those with fast delivery speed, and accurate items; this is only partially true. The writer observed in being both the consumer and the independent contractors of these service platforms that consumers are more likely to give a better

rate when they feel there are human interactions in the service experience. Since home delivery services rarely require face-to-face interactions with the consumers and delivery staff, the staff would drop off the delivery, and the service is completed; most of the time, there is only online chatting between the two. The lack of interaction might result in consumers forgetting that an actual human is fulfilling their orders. Being contactless is convenient to the consumers; however, making consumers think there is a human working for their orders and working hard to cater to their needs makes the delivery staff and the platform stand out.

Based on this circumstance, the writer proposes that consumers willingly leave a higher rate if human connections are involved in the service experience. To study this matter, this paper mainly discusses a number of previous studies relating to consumer satisfaction first; and introduce the conceptual model followed by three proposed studies, and discuss the result of the research and its implications.

2. Related Literature Review

2.1. Consumers' Ratings

Consumers' ratings reflect consumers' satisfaction and how the rating system is designed; the comprehensiveness of the rating system also influences the ratings [1]. Sometimes, consumer ratings do not reflect all consumers' evaluations of the service or product, and the ratings are sometimes biased. The biases are in at least two ways [1]: Firstly, on certain platforms, consumers tend to rate when they were satisfied rather than dissatisfied [2]. Secondly, ratings are mostly extreme. Meaning that the star ratings are often extremely positive or negative; this suggests that

Kaiyee.wu@mail.utoronto.ca

consumers who are highly favourable or unfavourable towards the product or service tend to make the effort of rating, consumers with moderate satisfaction are less likely to devote the effort [3].

2.2. Expectation-disconfirmation Theory

Expectation-disconfirmation theory (EDT) [4] suggests that consumers sometimes compare the product or service performance against their expectation of the product or service they received. The performance should meet or go beyond the original expectations to obtain a satisfactory rating [5]. EDT was widely used in studying consumer behaviors, as Ho, Wu, and Tan mentioned in their research [6]. In studying post-purchase evaluations, scholars use the expectation-disconfirmation theory to reason consumer reactions like repurchase decisions [7] and post-purchase complaints [8]. However, factors that influence consumer ratings are more than listed criteria in the service-providing platforms.

2.3. Consumer Satisfaction

Studying consumer satisfaction is also crucial to this paper; consumer satisfaction is a certain consumer response that carries a specific focus determined at the particular time [9]. Giese and Cote's research [6] specified the focus in the response being expectations, products, consumption experience, and so on. Consumers' satisfaction is often derived from constantly comparing to some standard, whether the standard lies in the rating systems or consumers' past experience of similar purchases.

2.4. Human Interaction and Happiness

Social interactions contribute to the happiness of individuals, regardless of individuals' construal of self [10]. In this way, interactions between service providers and consumers might result in higher consumer ratings.

From these previous researches and works, we can infer that consumer ratings reflect consumers' evaluation of the product or service; however, it can be influenced by other factors such as the rating system design, and the ratings are sometimes biased. Based on the expectation-disconfirmation theory, we learnt consumer evaluation behaviors, that is, consumers would often evaluate based on their expectation of the product or service. And this evaluation behaviour helps brands to justify consumers' post-purchase behaviors. As for consumer satisfaction, we acknowledged that it is a response that consumers come up with at a certain point of time during the service. It results from consumers' comparison of a certain standard. When studying the influences of human interactions in service evaluations, no significant empirical studies were found. However, researches imply that positive human interactions contribute to ones' overall happiness.

3. Conceptual Model

This paper argues that, especially in online home delivery service, when contact with consumers and service providers are not required, more human interactions would result in more ratings received by the service provider, and positive human interactions might result in higher ratings.

Our model would conduct three studies to study the relationship between human interaction between service-providing individuals and the consumers.

In Study 1, we try to test out if the existence of human interactions in the service experience would influence the willingness to rate. This is to say if there is a human reaction that exists during the service, chance to say the consumer would be more or less willing to leave a rate for the service provider. In the study we use the existence of human interactions as the independent variable and whether consumers rate or not being the dependent variable. However, in this study, we need to eliminate the influence of consumers' personal experience on rating behaviour. Elaborating on personal experience, it would be optimal to make sure the consumers do not have the experience of being an independent contractor of these platforms nor do they understand how important ratings are to the independent service providers. Since the previous experiences of being the contractors for the platforms would incent consumers to willingly rate for the service providers, doing this we could minimize other factors that could potentially affect study results.

For Study 2 and Study 3, we study if the positivity in the interaction would affect the rating regarding quantity and quality. In learning the quantity of the rating, we would base on the previous study of Dellarocas and Wood [2], to see in what case consumers are more likely to devote the effort to rate when positive interactions occur or negative ones. Based on Ho et al.'s study [3], these consumers might rate more extremely; under this circumstance, consumers might give more 5-star and 1-star ratings rather than moderate ratings like 3 to 4 stars. For studying rating quality, we look into if the interaction is positive would the consumers' rate the service higher or lower than those with negative human interactions? In this study, the key concern is to define what positive means in the human reaction. We need to make sure that human interaction is what we are studying. If an outside issue results in a negative experience for the consumers, this would still lead to a positive human interaction regardless. A negative service experience does not necessarily mean a negative human interaction took place. For instance, a consumer called an Uber to get home during rush hours. Inevitably, the consumer was stuck in a traffic jam while in the Uber ride. In this case, this ride might be an awful experience to the customer; however, if the Uber driver communicates with the rider, comforting the consumer and showing that the driver cares about the value of consumers' precious time or even trying to take short-cuts to avoid traffic. Consumers might consider human's interactions as positive. Take the Instacart shopping experience as another example. When the shopper is shopping for the customer, the shopper discovers that numerous products that the consumer asked for are out of

stock in the designated stores. Now the consumer might consider this unfavourable, but if now the shopper attentively chatting with the consumer assuring her/his/their need would be taken care of and provide the consumer with the suitable alternatives. After this consumer would think she/he/they had a positive interaction with the shopper although the store is unfavourable. Making sure the concepts of human interactions are clear. The evaluation of positivity should only be limited to human interaction for independent variables helps us study the notion of positive interaction better in regard to learning consumer rating behaviors.

3.1. Hypothesis

In Study 1 where we test out the consumers willingness in rating when interactions take place, we propose that:

H1: For the service experiences in which human interactions take place, the consumers are more likely to devote the effort to rate the experiences after the service is complete.

As for the second and third studies, we study consumer rating quantity and quality-wise, with respect to the positivity of the interactions between the service personnel and the consumers when the human interactions are already established.

H2: When human interactions occur during the service, consumers would be more likely to willingly rate when the interactions are negative rather than positive.

H3: When human interactions occur during the service, consumers would more likely to give higher extreme ratings when positive human interactions occur, lower extreme ratings when negative human interactions occur.

3.2. Study 1: Willingness in consumer rating and human interactions

This study would distribute questionnaires about individuals' experience of online service platforms such as Uber and UberEats. After receiving the questionnaire results, we would select 50 participants with zero experience being an independent contractor of those online service platforms and unaware of how important receiving a rating is to the contractors to avoid bias in the research results. Other than the subjects, we would also invite 25 Instacart service providers as experimenters to fulfill the services. To conduct the study, we would divide the participants into two groups. The first group is the experiment group of 25 randomly chosen participants from the 50 participants, whom we ask the service providers to interact with. The second group of participants would act as the controlled group, with every condition being the same as the first group, just without interactions with the service providers.

The experimenters would ask participants to choose five items at a designated convenience store through the Instacart app. The store would guarantee full stock of all the commodities available in the Instacart app. Then the Instacart services providers would take their orders to shop for them. In controlling the human interaction

variable, for the experiment group, shoppers would first greet consumers warmly by sending them a chat through the app telling them they are happy to shop for them and ready to begin shopping. Secondly, shoppers would be asked to tell the participants that two of the items they chose are out of stock, and they would be working to provide them with the best alternative. For the controlled group, shoppers would not greet the participant, and they would get all the items that the participant ordered then complete the whole service without chatting with the participant. We would not ask the participants to rate in the end.

Lastly, we would gather participants' ratings to study, thank the participants for participating and compensate the shoppers.

3.3. Study 2 & 3: Willingness and quality of consumer ratings and quality of human interactions

The second and third studies are about testing out whether the quality of the human interaction would affect consumers' willingness to rate and the quality of the ratings. Similar to the first study, we would ask 50 participants and 25 independent contractors to complete these studies. This study split the 50 participants into two groups. One group is the positive interaction group, and the second is the negative interaction group. However, this time we use the Uber app as our primary platform. This time, the participants would be asked to take a 10-minute Uber ride to the designated destination; we make sure there are no traffic jams along the ride. For both groups, we would ask the contractors to greet the participants and have a chat with them. For the group of positive interaction, we would ask the drivers to welcome the participants energetically, chat with them about light-hearted events like Christmas or other things; when the ride completes, the driver should say goodbye and wish the participants a good one. For the negative interaction group, we also ask the drivers to greet the participants but in a plain tone, chat with the participants but complain about their work or traffic or other bothersome events; furthermore, the drivers would also act tired and not satisfied during the ride. In the end, the "negative" drivers would not say goodbye to the riders and would drive off as soon as the participants got off from the vehicle. Lastly, we would gather participants' ratings to study, thank the participants for participating and compensate the shoppers.

3.4. Results

As for the first study, the results should be consistent with the hypothesis. We would gather more consumer ratings from the experiment groups and less from the controlled group with no human interaction involved in the experience.

For Study 2, more ratings would come from the participants from the negative group than the participants of the positive group. Based on Study 2's result, we evaluate the quality of the ratings to get the results for

Study 3. Similar to what we proposed, we think that we would receive more extreme ratings from both groups. For example, the positive group ratings should consist of a majority of 5-star ratings, and the majority of the negative group ratings would be 1-star or 2-star ratings.

4. Conclusion

The results of the studies should be consistent with the previous researches and tell a lot about consumer rating behaviors. Due to the fast internet development and the pandemic's hit, more people are less willing to go out of their homes; the need for online services is increasing, and relative platforms are thriving. This study would help the service platforms consider whether or how to train their independent contractors to make the brand stand out. Service providers can also consider how they should behave when providing services in order to get more and higher ratings.

The studies might be hard to conduct since there are a lot of outside factors that can affect consumers' satisfaction. Consumers' personal preferences could affect the accuracy of the results. As far as the writer observed in real life, some consumers do not like interacting with service providers. Some Uber users do not appreciate chatting with the drivers because that makes them nervous. Some Instacart users also just want their groceries delivered to their doorstep hassle-free; they do not care to interact with the shoppers. In this case, we might also look into why consumers are not in favour of human interactions; and how to avoid unsatisfied service experiences caused by unwanted interactions.

Limitations for brands might be that they do not know whether the consumer favors human interactions in the service and how to measure human interactions in service providers' performance.

However, the limitations might be avoidable. For uncertainty in consumer service preference, brands can develop different types of services for consumers to choose from. For example, for consumers with social awkwardness, platforms can ask whether they need contactless service or quiet service, and then alert service providers after the consumer has revealed their preference. Secondly, Uber and UberEats do a good job in service measurements. Instead of just star ratings, they would provide icons that indicate "Great Communication" and "Delivered with Care" to measure the quality of human interaction in the service.

Other applications could be brands that can use the study findings to offer training sessions to individual service providers, to help them achieve better ratings through performing positive interactions with customers.

References

1. Schneider, Christoph, Markus Weinmann, Peter N.C. Mohr, and Jan vom Brocke.: "When the Stars Shine Too Bright: The Influence of Multidimensional Ratings on Online Consumer Ratings." *Management Science*, October (2020). <https://doi.org/10.1287/mnsc.2020.3654>.
2. Dellarocas, C., & Wood, C. A.: The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Science*, 54(3), 460–476 (2008). <https://doi.org/10.1287/mnsc.1070.0747>
3. Ho, Yi-Chun (Chad), Junjie Wu, and Yong Tan.: "Disconfirmation Effect on Online Rating Behavior: A Structural Model." *Information Systems Research* 28 (3): 626–42 (2017). <https://doi.org/10.1287/isre.2017.0694>.
4. Oliver, Richard L.: "Effect of Expectation and Disconfirmation on Postexposure Product Evaluations: An Alternative Interpretation." *Journal of Applied Psychology* 62 (4): 480–86 (1977). <https://doi.org/10.1037/0021-9010.62.4.480>.
5. Zhang, Jiasheng, Wenna Chen, Nicolai Petrovsky, and Richard M. Walker.: "The Expectancy-Disconfirmation Model and Citizen Satisfaction with Public Services: A Meta-Analysis and an Agenda for Best Practice." *Public Administration Review*, May (2021). <https://doi.org/10.1111/puar.13368>.
6. Giese, Joan, and Joseph Cote.: "Defining Consumer Satisfaction." *Academy of Marketing Science Review* Volume 2000 (No. 1) (2002). <http://www.amsreview.org/articles/giese01-2000.pdf>.
7. Anderson, Eugene W., and Mary W. Sullivan.: "The Antecedents and Consequences of Customer Satisfaction for Firms." *Marketing Science* 12 (2): 125–43 (1993). <https://doi.org/10.1287/mksc.12.2.125>.
8. Bearden, William O., and Jesse E. Teel. "Selected Determinants of Consumer Satisfaction and Complaint Reports." *Journal of Marketing Research* 20 (1): 21(1983). <https://doi.org/10.2307/3151408>.
9. Fisher, Matthew, George E Newman, and Ravi Dhar.: "Seeing Stars: How the Binary Bias Distorts the Interpretation of Customer Ratings." Edited by Gita Johar and Stijn van Osselaer. *Journal of Consumer Research*, March (2018). <https://doi.org/10.1093/jcr/ucy017>.
10. Oishi, Shigehiro, Minkyung Koo, and Sharon Akimoto.: "Culture, Interpersonal Perceptions, and Happiness in Social Interactions." *Personality and Social Psychology Bulletin* 34 (3): 307–20 (2008). <https://doi.org/10.1177/0146167207311198>.