



Factors Influencing Customer Addictive Purchase Behaviours of Toy Blind Boxes

Wenluo Duan^(✉)

Department of Digital Humanities, Dital Economy, London WC2R 2LS, UK
duan_wenluo@163.com

Abstract. This article will examine which factors influence consumer behaviour when they buy toy blind boxes in the digital age. Data collection of both the observation group and control group are extracted through the questionnaire, and qualitative analysis such as Cluster analysis and the Chi-square test is used to analyse the differences between the two groups to ascertain which factors affect consumers' addiction status. The analysis of these 169 valid questionnaires yielded the finding that average costs, average purchase frequency, average purchase quantity, understanding of toy blind boxes, and ways of retrofitting unused toy blind boxes have a substantial impact on customer addiction and purchasing behaviours.

Keywords: toy blind boxes · customer behaviour · digital age · addiction

1 Introduction

Blind boxes is a fancy product that widespread in China, its main feature is the product sold in the form that placing unknown items of various values in the box with the same price, and the blind box whose contents or items are random within a certain range [1]. Consumers don't know what product they will get until they buy it, so this kind of box is called Blind Box [2]. Blind boxes can be categorized as loot boxes and toy blind boxes. In recent years, whether it is the loot box in the video game field or the toy blind box with IP toys as its main selling point, there are many loyal buyers, and the factors behind it indirectly affect the behaviour of many consumers.

There have been many articles analysing whether the blind box is gambling and how it affects consumer behaviour, especially the relationship between the gambling nature of loot boxes and game player addiction, including articles that specifically study adolescents and adult gamers. For example, Gonz'alez-Cabrera et al. [3] and Brooks and Clark [4] have studied loot box buying addiction among Spanish adolescents and loot box addiction among adults in general, respectively. There are also articles devoted to how loot boxes affect consumer psychology to stimulate persistent irrational consumption [5].

However, there are few studies on the blind box economy that has been successful in China recently, especially for Chinese toy blind box companies. Compared to loot boxes, these toy blind boxes have entities, and these toys themselves have a certain value

in real life, they are different from the characteristics of loot boxes that are limited to a certain online game. In addition, the articles on Chinese toy blind boxes are mainly from the perspective of corporate marketing, such as how POP MART promotes the purchase of toy blind boxes in China's Z era through marketing methods [6]. Therefore, there is a lack of research on toy blind boxes, especially from the perspective of consumer behaviour to explain the great success of the economy of blind boxes in China. Hence, this article focuses on what kinds of factors influencing consumer purchase behaviour of toy blind boxes in the digital era, causing people become addicted to toy blind boxes.

2 Research Methods

This study uses qualitative analysis. It is based on the data collected by the questionnaire, using both Cluster analysis and Chi-square analysis to research the purchase problem of toy blind-box consumers full of subjective perspectives. The questionnaire is a very direct and intuitive way of research. If the questionnaire is designed reasonably, meaningful answers can be obtained because it is necessary to think as carefully as possible about each question and the connection between the questions and the questions in order to obtain valid data [7, 8].

In addition, since the raw data collected in real life will always have many defects, and the quality of the data directly affects the analysis results, the quality of data pre-processing is crucial [9]. Even if the amount of data is large, the randomness is strong, the data is normally distributed due to the law of large numbers, and the influence of the disturbance term is small, it is a better choice to perform data preprocessing first and then analyse the questionnaire [10]. Therefore, the data from the questionnaire must be processed first. In order to better carry out the subsequent Cluster analysis, it is necessary to convert the qualitative questions in the questionnaire into quantitative questions, and the quantified data can be obtained through Cluster analysis [11].

Furthermore, the Chi-square test is a very effective way to analyse the possible differences between different groups in turn, so it is often used in articles on blind boxes. The Chi-square test is a test based on the Chi-square distribution, and it mainly tests if there is no difference between the observed value and the predicted value. It is assumed that there is no difference from the null hypothesis and if it is significant, the null hypothesis is overturned to prove that there is a difference between the two [12]. Moreover, the Chi-square test also can be used to check if the two factors have relations. The null hypothesis is that there is no relationship between these two factors. If the result is significant, the null hypothesis is overturned, and it is proved that there is a certain relationship between the two factors [13]. In Kristiansen and Severin's article [14], on loot boxes and adolescent problem gambling, Chi-squared test was used to analyse the relationship between gender and age and addiction to loot box buying.

3 Research Design

This article mainly uses a questionnaire to obtain the possible attitudes of consumers towards the consumption of toy blind boxes and the possible addiction mechanism behind them. A total of 169 questionnaires were collected, and the target group was

mainly Chinese. Most of the questions in this questionnaire are subjective, especially with multiple-choice questions. Therefore, it is necessary to reasonably transform the questions before extracting valid data. The main purpose of data processing is to delete the sample data that cannot be converted, and at the same time divide the existing questionnaire questions again, and reassign them to be sorted, or become 0 and 1 questions [11, 15]. The allocation of options can be transformed according to the standard of the Likert scale, which converts categorical options into quantitative options, and sorts them according to their actual meaning [16]. For example, the average cost is ranked from 0 to 4 by size, and a four-point or seven-point system will also be adopted according to the situation [17].

The first is to remove the people in the first question who chose “prefer not to say” or “other”, then there were 160 questionnaires left. In the second question “To what extent do you know about toy blind boxes”, the options below are graded Master is 4, very Much is 3, and so on, Null corresponds to 0. In this way, consumers with different degrees of understanding of the blind box can be expressed numerically. Next is the question of purchase frequency, which is also sorted according to different frequencies. The options such as few purchases and other (the people answer other is mainly not to buy or unclear, so it is properly counted as 0) correspond to 0, once a quarter is 1, and so on, many times a week corresponds to 4. For the question How many toy blind boxes do you buy per time? It is the same to use the method of rank order to assign values to different quantities, 0 corresponds to 0, and 1 corresponds to 1. The next question is to average the cost each time, and the same way to sort, from the least costs to the most expensive number from 1 to 4 one-to-one.

Moreover, the question of why buying blind boxes, where the first two options are “Just like it, no reason” and “The items in the boxes are cute or very special”, while the last two options are “Following the people around you” and “other”. The first two options are relatively positive purchases, and the latter two are relatively negative, so the first two options are combined into the positive purchase, and the latter two options are combined into the negative purchase, and each option is an assignment of 1 point. Since it is a multiple-choice question, the points can be stacked, for example, selecting the first two options at the same time will get two points. The following question “If failed to get the items you like when you open toy blind boxes, what would you like to do?” is the same as the one above, with options “Selling to others”, “Giving it to others”, “Retrofitting the items” is merged into the active treatment of this option, while the options “Staying it at home” and “Other” are merged into passive treatment. Again, each option is a point, and points can be accumulated. In addition, if both passive and positive options are selected, which is considered as 0 points.

4 Analysis

4.1 Qualitative Analysis

For whether people will be addicted to the toy blind box, it is first necessary to further classify the consumer groups, such as gender and extent of understanding about toy blind boxes in the questionnaire. Different gender and level of understanding may produce different results. The result shows the ratio of men and women in this questionnaire. It

can be seen that 62.13% of women who are the main consumer group and only 32.54% of men in contrast.

In addition, 46.75% of people know about toy blind boxes in general, and 29.59% and 11.24% excess normal level. The remaining small number of people do not know toy blind boxes, but the number is very small, which proves that toy blind boxes are indeed popular among Chinese.

Next is the purchase frequency. The sum of the frequency of purchases once a month or more is 46.74%, accounting for about half of the sample.

Corresponding to the purchase frequency is the number of a single purchase, the probability of purchasing two or less toy blind boxes in a single purchase is 69.57%, the rest is consumers who purchase a large quantity blind boxes in a single purchase, accounting for 30.43%.

In addition, concerning the question about average consumption, consumers who buy 12 lb or less each time account for 52.66%, while the rest of consumers with higher consumption account for 46.33%.

Next is the reason for buying the toy blind box. This questionnaire has an internal logic for the design of this question, which is active purchase and passive purchase. People buy the blind boxes because boxes are very cute (66.27%) and they just like them (36.09%), such a behavior is active purchase, people buy the toy blind boxes just because they want to follow other people around them are passive buying purchase.

The next question is also a test of consumers' purchasing attitudes. It is about how to deal with the toy blind box after buying it. Since there are multiple choices, the pie chart can only reflect the general choice of consumers, and it needs the data processed for further analysis. However, it also can be clearly seen from the result that compared to actively processing toy blind boxes, choosing passive processing and placing them at home is the main choice for consumers.

The next question is the possible emotions of consumers who do not obtain their favourite toy blind box. Except for the B option, nothing special, which is a relatively positive attitude, the rest are relatively strong emotional fluctuations. It can also be seen from the result that the proportion of the number of people with a calm state of mind and the proportion of strong psychological fluctuations are similar, and even many consumers will have several emotions at the same time.

The next question is to test whether consumers would make irrational retaliatory consumption if failed to obtain the favourite toy blind boxes. It can be seen from the questionnaire result that most people (67.46%) do not choose retaliatory consumption. Only 32.54% of people have a strong irrational consumption impulse.

The last question is to ask the consumers level of addiction on toy blind boxes. There are three levels of problems, it is nothing addiction that accounts for 34.32% of the population. A little is that the second level accounts for 50.3%, which is the status of the main consumers, 14.2% which is a small fraction of the population, and the left part is "other" which accounts for 1.18%. This question set is mainly to obtain a judgment of consumers on whether they are addicted to the toy blind box. The main analysis is the previous question, and the conclusion of this question is for assistance.

4.2 Further Analysis

It is essential part to research the connections among the different questions in the survey. Among them, consumers' different understanding of the toy blind box, the number of each purchase, and the average amount of each purchase may have an impact on whether the toy blind box will eventually produce addictive behaviour. Moreover, these questions are also related to how to dispose of the toy blind box after purchase, and the negative emotions caused by the condition of the buyers not buying the desired toy blind box. For further research, this research intends to apply Cluster analysis to divide consumers into two groups and see which factors can make consumers divided into two groups.

Furthermore, the result of the Cluster analysis made by SPSS is in Table 1 below.

According to above Table 1, there are some outcomes that can be seen. For the variable Gender, the significant P value is 0.662, which is not significant at the level, and the null hypothesis cannot be rejected, indicating that the variable Gender has no significant difference between the categories divided by Cluster analysis.

For the variable "To what extent you know about toy blind box", the significant P value is 0.011**, and the level is significant, rejecting the null hypothesis, indicating that there are significant differences between the categories classified by the class analysis.

For the variable "How often do you purchase toy blind boxes?", the significant P value is 0.000***, showing significance at the level, rejecting the null hypothesis, indicating that there is a significant difference.

For the variable "How many toy blind boxes do you buy per time?", the significant P value is 0.000***, and the level is significant, rejecting the null hypothesis, indicating that there is a significant difference between the categories.

For the variable "What are the average costs of your consumption on toy blind boxes every time?", the significant P value is 0.029**, and the level is significant, rejecting the null hypothesis, indicating that there is a significant difference between the categories divided by Cluster analysis.

For the variable "Negative purchase", the significant P value is 0.251, which is not significant at the level, and the null hypothesis cannot be rejected, indicating that there is no significant difference between the categories divided by Cluster analysis.

For the variable "Positive dealing", the significant P value is 0.000***, which is significant at the level, rejecting the null hypothesis, indicating that there are significant differences among the categories divided by Cluster analysis.

For the variable "Negative dealing", the significant P value is 0.701, which is not significant at the level, and the null hypothesis cannot be rejected, indicating that negative dealing has no significant difference between the categories divided by Cluster analysis.

For the variable "Positive purchase", the significant P value is 0.126, which is not significant at the level, and the null hypothesis cannot be rejected, indicating that the variable positive purchase has no significant difference between the categories divided by Cluster analysis.

For the variable "Emotion when failing has the item you want", the significant P value is 0.220, which is not significant at the level, and the null hypothesis cannot be rejected, indicating that the variable is not a significant difference.

For the variable "Would you choose retaliatory consumption for not getting the toy blind boxes you want?", the significant P value is 0.022**, which is significant at

Table 1. Cluster analysis result

	Cluster categories (mean \pm standard deviation)		F	P
	Category 1 (n = 140)	Category 2 (n = 20)		
Gender	0.65 \pm 0.479	0.7 \pm 0.47	0.192	0.662
To what extent you know about toy blind boxes ?	2.293 \pm 0.941	2.85 \pm 0.587	6.625	0.011**
How often do you purchase toy blind boxes ?	1.129 \pm 1.274	2.55 \pm 0.999	22.836	0.000***
How many toy blind boxes do you buy per time ?	1.575 \pm 0.887	6.625 \pm 2.422	319.214	0.000***
What are the average costs of your consumption on toy blind boxes every time ?	1.529 \pm 0.694	1.9 \pm 0.788	4.848	0.029**
Negative purchase	0.336 \pm 0.595	0.5 \pm 0.607	1.327	0.251
Positive purchase	0.75 \pm 0.779	1.5 \pm 0.827	15.993	0.000***
Negative dealing	0.693 \pm 0.463	0.65 \pm 0.489	0.148	0.701
Positive dealing	0.993 \pm 0.556	1.2 \pm 0.616	2.364	0.126
Emotion when failing has the item you want	0.607 \pm 0.49	0.75 \pm 0.444	1.519	0.220
Would you choose revenge spending because you didn't get the toy blind boxes you want ?	0.293 \pm 0.457	0.55 \pm 0.51	5.386	0.022**

Note: ***, **, * represent the significance level of 1%, 5%, and 10%, respectively

the level, rejecting the null hypothesis, indicating that there are significant differences between the categories divided by Cluster analysis.

For the variable "Are you addicted to buying toy blind boxes?", the significant P value is 0.041**, the level is significant, rejecting the null hypothesis, indicating there are significant differences between the categories divided by Cluster analysis.

Therefore, it can preliminarily be concluded that consumers' different understanding of toy blind boxes, purchase frequency, purchase quantity, average purchase cost, active transactions, revenge consumption psychology and the degree of addiction to toy blind boxes (self-evaluation) will affect the outcome of Clustering. It is an important result, and

Table 2. Cluster category

Cluster category	Frequency	Percentage (%)
Cluster Category 1	140	87.5%
Cluster Category 2	20	12.5%
Total	160	100%

according to the above factors, consumers can be mainly divided into two categories, and it is shown in Table 2 below. Consumers with a high degree of understanding, a large number of purchases, a high average purchase cost, active transactions, retaliatory consumption psychology and a high degree of self-addiction are Category 2, which is the minor consumer group of toy blind boxes. Category 1 is ordinary people who are the contrary group comparing with Category 2.

Moreover, toy blind boxes do have the potential to affect consumer behaviour, which first depends on different types of consumers. There are two main categories divided by consumers' understanding of toy blind boxes, purchase frequency, average costs, average purchase quantity, and positive dealing. However, if the factors really influence consumer behaviour needs the following Chi-square test to find out in the discussion part.

5 Discussions

Toy blind boxes have achieved huge commercial success as a new business model. This is not only due to the exceptional quality of the items but also some hidden mechanisms of gambling addiction reinforced by the Internet, which has impacts on customer behaviour, causing people addicted to purchase toy blind boxes. This essay primarily examines how the mechanism of toy blind boxes influences customer behaviour and what aspects of consumers are most readily impacted in the digital age.

Therefore, it is assumed that the two selected factors are independent as the null hypothesis, and the Chi-square test is performed on the factors that may affect consumption behaviour and whether consumers are addicted in turn. If the result significantly overturns the null hypothesis, it means the different levels of the factor may cause the corresponding extent of addiction, if it is not significant, the factor does not have effects on consumer addiction to affect behaviour.

The first is the Chi-square test of whether knowledge of the toy blind box is associated with addiction. As it can be seen from Table 3 below, if the confidence interval is 95%, the result is that the degree of understanding of the toy blind box has a significant effect on whether the toy blind box will be addicted, because the significant P value is 0.011**, and the level is significant, rejecting the null hypothesis. This means that the degree of knowledge about toy blind boxes is an important factor affecting consumer addiction. It is probably the more consumers know about toy blind boxes, the higher possibility they addicted when buying toy blind boxes, and the less they know, the lower possibility they will become addicted.

Table 3. Relations between knowledge about toy blind boxes and addiction

Topic	code	To what extent you know about toy blind boxes ?					Total	X ²	Correction X ²	P
		0	1	2	3	4				
Are you addicted to buying toy blind boxes ?	0	3	7	22	5	6	54	19.825	19.825	0.011**
	1	1	6	31	35	9	82			
	2	1	3	9	8	3	24			

Note: ***, **, * represent the significance level of 1%, 5%, and 10%, respectively

Next is whether purchase frequency is significantly associated with toy blind box addiction. It can be concluded from Table 4 below that there is a significant relationship between the two, because the X² value is 50.298, and the p-value is infinitely small, so against the null hypothesis, there is a strong relationship between the frequency of purchases and addiction. This shows that purchase frequency also significantly affects consumers' addiction status. People with higher purchase frequency are probably addicted to toy blind boxes, and possibly buy a large number of toy blind boxes because of irrational behaviours.

In the same way, it can be concluded that there is also a strong correlation between the average purchase cost and addiction. The more the average cost of the toy blind box, the more addicted to the toy blind box. Since the p-value of the Chi-square test from Table 5 below is extremely small, the null hypothesis is rejected, so there is a link between average spending and addiction, and the more spending, the greater the likelihood of addiction. This shows that the average consumption and addiction may be positively correlated. The higher the average consumption each time, the more possibly it will affect the behaviour of normal consumers addicted to the toy blind box.

Next is whether the average number of toy blind boxes purchased affects the behaviour of toy blind box addiction, which is also a significant effect. As shown in Table 6 below, the X² value is 53.972, which is significantly higher than the normal value, and the p-value is 0.001 to reject the null hypothesis, so the average number

Table 4. Relations between purchasing frequency and addiction

Topic	code	How often do you purchase toy blind boxes ?					Total	X ²	Correction X ²	P
		0	1	2	3	4				
Are you addicted to buying toy blind boxes ?	0	42	6	2	3	6	54	50.298	19.825	0.000***
	1	23	6	26	16	9	82			
	2	1	3	9	3	3	24			

Note: ***, **, * represent the significance level of 1%, 5%, and 10%, respectively

Table 5. Relations between average costs per time and addiction

Topic	code	What are the average costs of your consumption ob toy blind boxes every time ?					Total	X ²	Correction X ²	P
		0	1	2	3	4				
Are you addicted to buying toy blind boxes ?	0	43	8	2	1	54	54	25.565	19.825	0.000***
	1	35	6	6	2	82	82			
	2	1	3	9	0	24	24			

Note: ***, **, * represent the significance level of 1%, 5%, and 10%, respectively

of purchases each time will also affect addictive behaviour, and the more the average quantity of purchases, the easier it will appear to be addicted.

The next is whether active treatment affects addiction. The active treatment itself also reflects the attitude towards the toy blind box. It first assumed that there was no relationship between active processing and the hypothesis, and then performed the Chi-square test. It is found from Table 7 below that the X2 value is 17.864, and the p-value is 0.007, which overturned the null hypothesis. The outcome is significant, and the attitude of active processing would affect the degree of addiction. This means that consumers who are willing to actively deal with the toy blind box after buying it are probably more addicted to the toy blind box than ordinary people.

Finally, whether there is a relationship between negative emotions and the toy blind box addiction, it can also be seen from Table 8 below that the X2 value of 5.932 is not significant, and the p-value is 0.052. If 5% is used as the significance standard, it cannot be rejected. The null hypothesis is that there is no significant relationship between the negative emotions caused by buying the toy blind box and the addition of the toy blind box, so even if there are negative emotions, it is not necessarily addictive. Even if some people are feasible to have negative emotions and make irrational consumption, one or two short-term impulsive behaviours will not make them possibly addicted to the toy blind box.

Hence, consumer addiction is influenced by aspects such as knowledge of toy blind boxes, average expenditure, frequency of toy blind box purchases, average quantity per purchase, and active handling.

Table 6. Relations between purchasing quantity of toy blind boxes and addiction

Topic	Are you addicted to buying toy blind boxes?			
	Code	0	1	2
How many toy blind boxes do you buy per time?	0	8	3	2
	1	28	25	2
	2.5	3	5	0
	2	6	23	7
	2.5	1	3	3
	3	5	11	4
	10	1	0	0
	12	2	0	0
	4	0	1	3
	5	0	4	2
	6	0	2	0
	7	0	2	1
	7.5	0	1	0
	8	0	2	0
Total		54	82	24
X2		53.972		
Correction X2		53.972		
P		0.001***		

Note: ***, **, * represent the significance level of 1%, 5%, and 10%, respectively

Table 7. Relations between positive dealing of toy blind boxes and addiction

Topic	code	Positive dealing				Total	X ²	Correction X ²	P
		0	1	2	3				
Are you addicted to buying toy blind boxes ?	0	31	17	3	3	54	17.864	53.972	0.007***
	1	3	38	5	4	82			
	2	2	2	0	0	24			

Note: ***, **, * represent the significance level of 1%, 5%, and 10%, respectively

Table 8. Relations between purchase emotions and addiction

Topic	code	Emotions when failing to acquire the items consumers want		Total	X ²	Correction X ²	P
		0	1				
Are you addicted to buying toy blind boxes ?	0	27	27	54	5.932	53.972	0.053*
	1	3	55	82			
	2	2	2	24			

Note: ***, **, * represent the significance level of 1%, 5%, and 10%, respectively

6 Conclusion

In this article, qualitative research methods, such as Cluster analysis and Chi-square test, are used as tools to explore the main influencing factors that affect consumers' behaviours when buying toy blind boxes. The following five factors were found to significantly affect whether consumers are addicted to toy blind boxes, including the consumer's knowledge of toy blind boxes, the average cost per time, the average number of purchases per time, the frequency of purchases, and positive retrofitting. Therefore, the more consumers know about the toy blind box, or the higher the average cost of the toy blind box, the more frequent the purchase, the greater the number of purchases, and the more positive retrofitting will cause consumers easier addicted to the toy blind box.

Moreover, in the Internet age, the rise of e-commerce platforms and the precise advertising brought about by big data analysis make it accessible for every potential consumer to further comprehend toy blind boxes, and it is liable for consumers to impulse purchases or frequent purchases. Therefore, in the internet era, according to the previous analysis, consumers are possibly addicted to blind boxes, which makes toy blind boxes headed by POP MART, not only popular offline but also online. To some extent, the gambling nature of the toy blind box itself and consumer behaviour have a two-way impact. The toy blind box has certain risks and predictable benefits, which increases the probability of many consumers indulging in it. In the meantime, consumers are gradually indulging in the consumption of toy blind boxes, as their understanding of toy blind boxes deepens or their average consumption increases, they may eventually contribute to more consumption amount.

However, there are some limitations of the article. The scope of the research object is relatively small because the research question of the thesis is based on the toy blind box in China, so the research survey object is mainly on Chinese, and there is a lack of research on how the toy blind box affects consumer behaviour in other regions under the digital age. Additionally, although this article concentrates on toy blind boxes, anything can be a blind box, not only toys or loot boxes in video games, there are other kinds of blind box that deserve to research.

In a word, the blind box is a kind of thinking, and the toy blind box has achieved unimaginable success in the Internet era. No matter the small group of consumers who invest much time and money in the toy blind box and get addicted to it or the ordinary consumers, toy blind boxes, as a new business model with a gambling nature, deeply affect the behaviour of consumers.

References

1. Xiao, L. Y., 2022. Blind Boxes: Opening Our Eyes to the Challenging Regulation of Gambling-Like Products and Gambification and Unexplained Regulatory Inaction.
2. Wang, D.J and Zhou. C.C., 2021. Manghexiaofei: dangdai qingnian xiaofei shenghuofangshi de xinxiangxiang [Blind Box Consumption: A New Phenomenon of Contemporary Youth Consumption Lifestyle]. *Gansushehuixue* (02), pp.120-126. doi:<https://doi.org/10.15891/j.cnki.cn62-1093/c.2021.02.017>.
3. González-Cabrera, J., Basterra-González, A., Montiel, I., Calvete, E., Pontes, H.M. and Machimbarrena, J.M., 2022. Loot boxes in Spanish adolescents and young adults: Relationship with internet gaming disorder and online gambling disorder. *Computers in Human Behavior*, 126, p.107012.
4. Brooks, G.A. and Clark, L., 2019. Associations between loot box use, problematic gaming and gambling, and gambling-related cognitions. *Addictive behaviors*, 96, pp.26-34.
5. Larche, C.J., Chini, K., Lee, C., Dixon, M.J. and Fernandes, M., 2021. Rare loot box rewards trigger larger arousal and reward responses, and greater urge to open more loot boxes. *Journal of gambling studies*, 37(1), pp.141-163.
6. Gao, J. and Chen, R., 2022. Understanding consumer behaviors of Generation Z under China's blind box economy: case company: POP MART.
7. Groves, R.M., Fowler Jr, F.J., Couper, M.P., Lepkowski, J.M., Singer, E. and Tourangeau, R., 2011. *Survey methodology*. John Wiley & Sons.
8. Boynton, P.M. and Greenhalgh, T., 2004. Selecting, designing, and developing your questionnaire. *Bmj*, 328(7451), pp.1312-1315.
9. Ramírez-Gallego, S., Krawczyk, B., García, S., Woźniak, M. and Herrera, F., 2017. A survey on data preprocessing for data stream mining: Current status and future directions. *Neurocomputing*, 239, pp.39-57.
10. García, S., Luengo, J. and Herrera, F., 2015. *Data preprocessing in data mining* (Vol. 72, pp. 59–139). Cham, Switzerland: Springer International Publishing.
11. Duran, B.S. and Odell, P.L., 2013. *Cluster analysis: a survey* (Vol. 100). Springer Science & Business Media.
12. Sharpe, D., 2015. Chi-square test is statistically significant: Now what?. *Practical Assessment, Research, and Evaluation*, 20(1), p.8.
13. Ugoni, A. and Walker, B.F., 1995. The Chi square test: an introduction. *COMSIG review*, 4(3), p.61.
14. Kristiansen, S. and Severin, M.C., 2020. Loot box engagement and problem gambling among adolescent gamers: Findings from a national survey. *Addictive Behaviors*, 103, p.106254.
15. Mishra, P., Biancolillo, A., Roger, J.M., Marini, F. and Rutledge, D.N., 2020. New data preprocessing trends based on ensemble of multiple preprocessing techniques. *TrAC Trends in Analytical Chemistry*, 132, p.116045.
16. Joshi, A., Kale, S., Chandel, S. and Pal, D.K., 2015. Likert scale: Explored and explained. *British journal of applied science & technology*, 7(4), p.396.
17. Chomeya, R., 2010. Quality of psychology test between Likert scale 5 and 6 points. *Journal of Social Sciences*, 6(3), pp.399-403.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

