Andrew J. Roback
Survey Results Write-up
20 October 2010

Introduction

In order to address a subset of our research questions, we fielded a survey targeted at users of the Dominick's self-service checkout kiosk which was designed to gauge user perception of the system and collect demographic information from our target audience. Despite some shortcomings in the design of the survey, we were able to extract useful information that will impact our redesign process. This report will:

1. Present our survey results,
2. Discuss how those results are applicable to our research questions,
3. Discuss the shortcomings of our survey design and how we might redesign this survey in the future, and,
4. Discuss statistical methods that might be useful in obtaining further data from our results.

Survey Results

The primary indicator of overall satisfaction was a five point Likert scale (5=very satisfied), which returned a mean satisfaction rating of 3.68 (standard deviation=0.96, mode=4). The distribution is represented graphically in Figure 1.

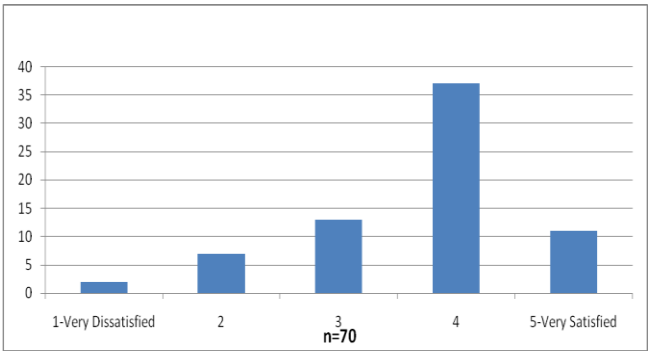

Fig. 1: Distribution of satisfaction ratings

Ease of use was rated on a four point Likert scale for each feature that we determined users would most frequently encounter when using a kiosk (in our own estimation, and based on results of a pilot study). Comparing each individual response to other results would require more sophisticated statistical methods, an issue that will be further discussed later in this report. In order to have numbers with which to make comparisons, we found the mean ease of use value for each category, and the mean ease of use value for all users across all categories. The mean ease of use value for each category is displayed below in table 1 and graphically in figure 2. The mean ease of use value for each user across categories was used for questions that compare ease of use to other user dimensions.

Ease of Use (responses on a 4 point Likert scale)	n	mean	standard deviation	median	mode
Getting started	70	3.83	0.42	4	4
Scanning items with UPC codes	68	3.75	0.50	4	4
Searching product codes for produce and other items without barcodes	62	2.90	0.82	3	3
Weighing produce or other by-the-pound products	62	3.21	0.85	3	4
Paying with credit or debit card	68	3.84	0.41	4	4
Paying with cash	43	3.70	0.51	4	4
Paying with check	4	2.50	0.58	2.5	3
Paying with another payment method	6	3.67	0.52	4	4
Bagging your groceries	68	3.34	0.80	4	4
Receiving change or cash back	45	3.84	0.37	4	4
Entering coupons	24	3.29	0.91	4	4
Purchasing liquor, tobacco, or other age-restricted sale items	25	2.84	0.85	3	3

Table 1: Ease of use by category; shading is for visibility purposes only.

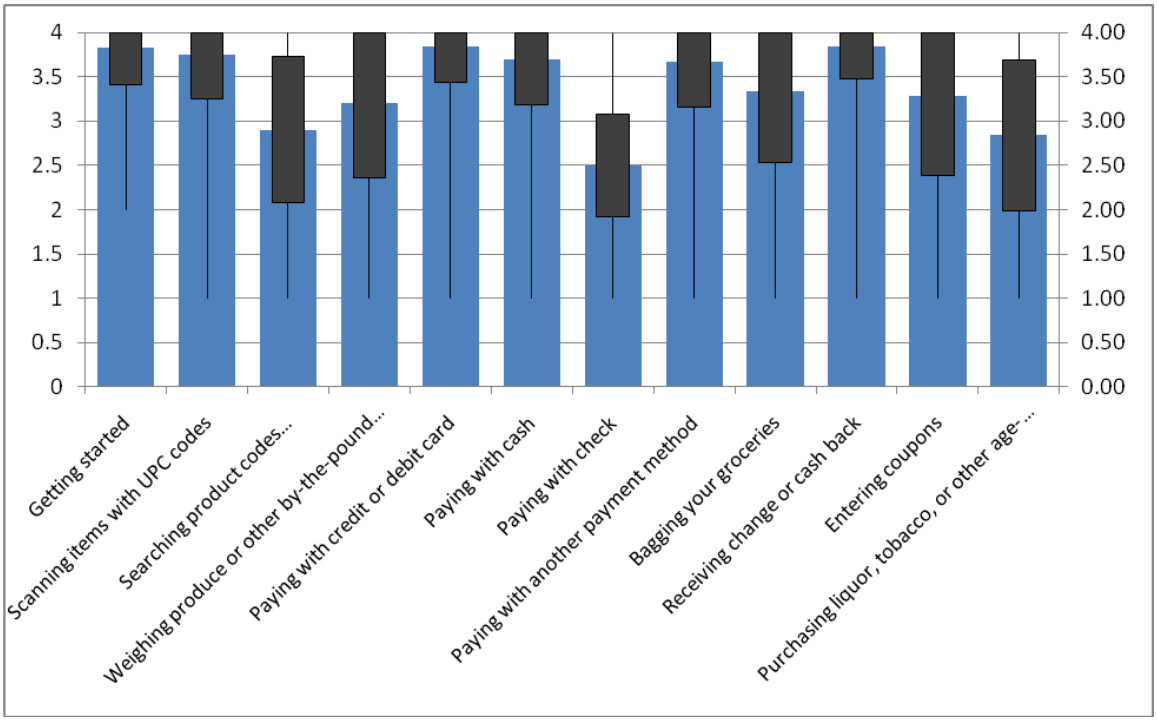

Fig. 2: A histogram of mean ease of use values with range (thin lines) and standard deviation (solid boxes). Possible responses for users were whole numbers 1-4 (there was no response analogous to zero; users could opt out of rating a feature they had never used).

Research Questions

This section will present statistics that attempt to answer the questions raised during the design of the survey. In the Results section, a summary description on the left of the page describes the significance of the statistical results on the right. A detailed discussion of these statistics follows in the Discussion section.

Results

Does frequency of use correlate with satisfaction?

A correlation of the frequency with which a shopper uses the kiosk (every time, sometimes, rarely) and the customer's associated satisfaction rating shows a positive correlation.

correlation coefficient	0.390
degrees of freedom	68
critical value	0.232
significant?	Yes

Does frequency of use correlate with mean perceived ease of use (for each user across all features)?

A correlation of the frequency with which a shopper uses the kiosk (every time, sometimes, rarely) and the shopper's mean perceived ease of use across all categories of use shows a positive correlation.

correlation coefficient	0.530
degrees of freedom	68
critical value	0.232
significant?	Yes

Does perceived ease of use impact how frequently customers use the self-service kiosk?

Frequency of use	n	Mean Ease of Use Rating (across all features)	p values: Rarely	p values: Sometimes	p values: Every time	Significant?
Every time	17	3.50	0.165	0.647		No
Sometimes	41	3.56	0.0592		0.647	No
Rarely	12	3.23		0.0592	0.165	No

Table 2: Results of a t-test to determine whether users grouped by frequency of use have statistically significant differences in mean perceived ease of use (across all categories)

As represented in the above chart, the t-test revealed that differences in mean perceived ease of use across shoppers grouped by frequency of use were not statistically significant.

Does the amount of times per week that a person shops for groceries correlate with satisfaction?

A positive correlation was established between the amount of times per week that a person shops for groceries and satisfaction, but the results are only barely significant. The correlation coefficient value is very close to the minimum critical value (the value at which the correlation coefficient is significant with p=0.05).

correlation coefficient	0.251
degrees of freedom	68
critical value	0.232
significant?	Yes

Does the amount of times per week that a person shops for groceries correlate with the mean perceived ease of use?

There is no evidence of a correlation between these two values.

correlation coefficient	0.020
degrees of freedom	68
critical value	0.232
significant?	No

Does the number of people the user shops for correlate to the frequency of use of self-service kiosks?

There is no evidence of a correlation between these two values.

correlation coefficient	-0.071
degrees of freedom	68
critical value	0.232
significant?	No

Is there a difference between primary shoppers and non-primary shoppers in terms of frequency of use?

A t-test of frequency of use results across the primary and non-primary shopper groups revealed that there is no significant difference in how frequently primary and non-primary shoppers use the kiosk.

p	0.251
significant	No

Is there a difference between primary shoppers and non-primary shoppers in terms of satisfaction?

A t-test of overall satisfaction results across the primary and non-primary shopper groups revealed that there is no significant difference in how satisfied primary and non-primary shoppers are with the kiosk.

p	0.251
significant	No

Does the amount spent by a shopper correlate with frequency of use of the self-service kiosk?

An inverse correlation exists between the amount a shopper spends and the frequency with which a shopper uses the kiosk.

correlation coefficient	-0.358
degrees of freedom	68
critical value	0.232
significant?	Yes

Does the amount spent by a shopper correlate with satisfaction with the self-service kiosk?

There is no evidence of a correlation between these two values.

correlation coefficient	0.133
degrees of freedom	68
critical value	0.232
significant?	No

Does the number of features a shopper reports using correlate with frequency of use of the self-service kiosk?

A correlation exists between number of features used and frequency of use.

correlation coefficient	0.437
degrees of freedom	68
critical value	0.232
significant?	Yes

Does the number of features a shopper reports using correlate with satisfaction with the self-service kiosk?

There is no evidence of a correlation between these two values.

correlation coefficient	0.165
degrees of freedom	68
critical value	0.232
significant?	No

Discussion

Our descriptive statistics indicate a need to address two types of interaction with the kiosk: searching for UPC codes for non-barcode items and purchasing age-restricted items. Customers reported a lower perceived ease of use for these tasks as can be discerned from figure 2 above. Although customers reported difficulty paying by check, a very small subset of users (n=4, or roughly 5% of respondents) reported using this method of payment making it a low priority item. Our prioritization of redesign efforts will be based, in part, on these findings.

Customers who use the self-service checkout kiosks frequently are on average more satisfied with the kiosk and also have a higher average perceived ease of use ($p \leq 0.05$, respectively). In terms of how often customers shop for groceries, we see that customers have a higher degree of satisfaction the more often they shop for groceries; however, these results are tenuous since the values were very close to the minimum level of statistical significance. There is no data that suggests shoppers who shop for groceries more often also have a higher perceived ease of use rating.

In order to determine if there were differences in the mean overall satisfaction and mean frequency of use between primary shoppers and non-primary shoppers, three t-tests were conducted. The results show that their mean responses are not significantly different.

There was no evidence to suggest a correlation between number of persons shopped for and frequency of use; however, an inverse correlation is evident between the amount a person spends on groceries and the frequency of use. Hence, we can say with confidence that the more a person spends on groceries, the less likely they are to use a self-service kiosk and vice versa ($p \leq 0.05$). No connection was found between the amount spent on groceries and overall user satisfaction.

Finally, we defined "usage patterns" as the number of features used by each shopper. The number of features used did not correlate with satisfaction but did show a strong correlation with frequency of use. We can say with confidence that there is a connection between the number of features used and the frequency with which shoppers use the self-service kiosk ($p \leq 0.05$).

It is important to remember when interpreting these results that correlations only show connections between the results and do not demonstrate that one factor influences the other in any way. Further usability research methods must be employed prior to speculating on causation.

Survey Shortcomings

Our survey had two design problems which limited our ability to answer questions proposed during the survey design:

1. We anticipated that how users began using the kiosk (select language, scan item, etc.) would relate to how satisfied they were with the system and what their "error frequency" was. We allowed users to select multiple starting techniques, which would make several different combinations of results that would have to be compared to one satisfaction value, a problem that needs a more complex statistical analysis than can be provided here. We also did not define "error frequency" or ask a question that provided a metric (e.g. Out of ten times using the self-service kiosk, how often do you experience an error). Determining information about user errors is better left to observation and interview, where the observer/interviewer can be sure that the user understands questions and can code data according to error type (as opposed to relying on the user to translate his or her perceptions into our coding schema).
2. We expected to correlate satisfaction with other kiosks to satisfaction with the Dominick's kiosk, but we allowed users to input their own perceptual responses in the text box and did not limit their choices. Data would have to be coded and then compared in order to answer this question.

Collecting future data and statistical methods

In the future, questions dealing with satisfaction on a survey should be calibrated to the same scale (using defined and validated responses) if comparisons are to be made.

Even though satisfaction is not analogous to usability, an analysis of perceived ease of use ratings with user satisfaction and frequency of use might demonstrate a relationship between one or more features and overall satisfaction or frequency of use (or an inverse relationship, which might help us to prioritize redesign efforts). A regression analysis might be appropriate for finding relationships between these variables.

Since fielding a revised survey is not possible, analysis of the relationships between multiple groups using other statistical methods might help us to make more significant connections using the data that we have (e.g. satisfaction of persons who experience three or more types of errors versus persons who experience less than three). Further time to code our data and better knowledge of statistical methods and interpretation of results will most likely lead to further connections between results, which will in turn help us to move forward in the redesign process and help us to better support our answers to our research questions.

Findings in this report should be seen as a foundation for future statistical work on the survey results. Although the survey contained some design flaws that limited our ability to answer specific research questions, the results we were able to obtain answered several questions posed in the design phase and will help us in designing future usability tests in this study.