

Analysis Report

This report presents the results of a multivariate within-subjects ANOVA, which was conducted to examine the differences in two measures of brand trust before and after an intervention had taken place. The report is structured as follows.

Contents

Descriptive statistics	2
Reliability Analysis.....	2
Tests of Assumptions for ANOVA	3
Normality	4
Outliers.....	4
Repeated-measures ANOVA	5
Multivariate Effects.....	6
Univariate Within-Subjects Effects	6
References.....	8

Descriptive statistics

This section presents descriptive statistics and graphs of the variables of interest. The section starts by presenting a frequency table, with the proportion of the sample by sociodemographic category (table below).

<i>Descriptive Statistics</i>			
Variable		Count	Column N %
And what's your gender?	Female	76	42.70%
	Male	96	53.93%
	Non-binary	6	3.37%
	Total	178	100.00%
What about education?	Higher education (PhD, WO, HBO)	147	82.58%
	MBO	16	8.99%
	Secondary school	15	8.43%
	Total	178	100.00%
Have you ever flown with KLM before?	No	25	14.04%
	Yes	153	85.96%
	Total	178	100.00%

The majority of the sample is male (53.93%), with higher education (82.58%) and have flown with KLM before (85.96%).

The sample size was initially composed by 188 participants, but 10 of these participants were outside the age range of 21 to 42 and were therefore excluded from the study. Final sample size was 178 (N = 178).

Reliability Analysis

Before any hypotheses were tested, the scales needed to be tested for reliability. Reliability is an assessment of the degree of consistency between multiple measurements of a variable. One form of reliability is test–retest, by which consistency is measured between the responses for an individual at two points in time. The objective is to ensure that responses are not too varied across time periods so that a measurement taken at any point in time is reliable. A second and more commonly used measure of reliability is internal consistency, which applies to the consistency among the variables in a summated scale. The rationale for internal consistency is that the individual items or indicators of the scale should all be measuring the same construct and thus be highly intercorrelated (Hair et al., 2014). The analysis in this study was done using Cronbach's Alpha. The table below shows the mean and

standard deviation of each survey item, along with the calculated alpha for each scale, which should be higher than 0.700 to indicate good reliability (Hair et al., 2014).

Construct	Item	Mean	Std. Deviation	N	Cronbach's Alpha
Fiability - Pre	With KLM, I can obtain what I look for in a flight	3.65	1.053	178	0.818
	KLM is an airline that meets my expectations	3.57	1.030	178	
	I feel confident with KLM	3.61	1.116	178	
	KLM is a brand that never disappoints me	2.57	1.252	178	
	KLM is not constant in satisfying my needs	3.11	0.942	178	
Fiability - Post	With KLM, I can obtain what I look for in a flight	3.47	1.053	178	0.829
	KLM is an airline that meets my expectations	3.41	1.128	178	
	I feel confident with KLM	3.43	1.088	178	
	KLM is a brand that never disappoints me	2.68	1.204	178	
	KLM is not constant in satisfying my needs	3.20	0.917	178	
Intentionality - Pre	KLM would be honest and sincere in addressing my concerns	2.60	1.101	178	0.845
	KLM would make any effort to satisfy me	2.69	1.074	178	
	I could rely on KLM to solve problems	2.69	1.116	178	
	KLM would be interested in my satisfaction	2.97	1.139	178	
	KLM would compensate me in some way for a problem with a flight	3.08	1.008	178	
	KLM would not be willing in solving a problem I could have with a flight	3.29	0.922	178	
Intentionality - Post	KLM would be honest and sincere in addressing my concerns	2.98	1.068	178	0.853
	KLM would make any effort to satisfy me	2.94	1.085	178	
	I could rely on KLM to solve problems	2.92	1.134	178	
	KLM would be interested in my satisfaction	3.15	1.055	178	
	KLM would compensate me in some way for a problem with a flight	3.10	1.018	178	
	KLM would not be willing in solving a problem I could have with a flight	3.28	0.900	178	

All the four scales were successfully tested for reliability, since the lowest Alpha was 0.818, above the minimum threshold.

Tests of Assumptions for ANOVA

The ANOVA method (to be applied in this study) needs the data distribution of the dependent variables to be normal. An additional assumption is that there are no outliers in the data (extreme values). The next sections assess these conditions.

Normality

One of the methods to examine normality is to look at values of skewness and kurtosis (Hair et al., 2014). Both values should remain between -1 and 1 to indicate normality. As can be seen in the table below, none of the values surpass these thresholds, which indicates no depart from normality.

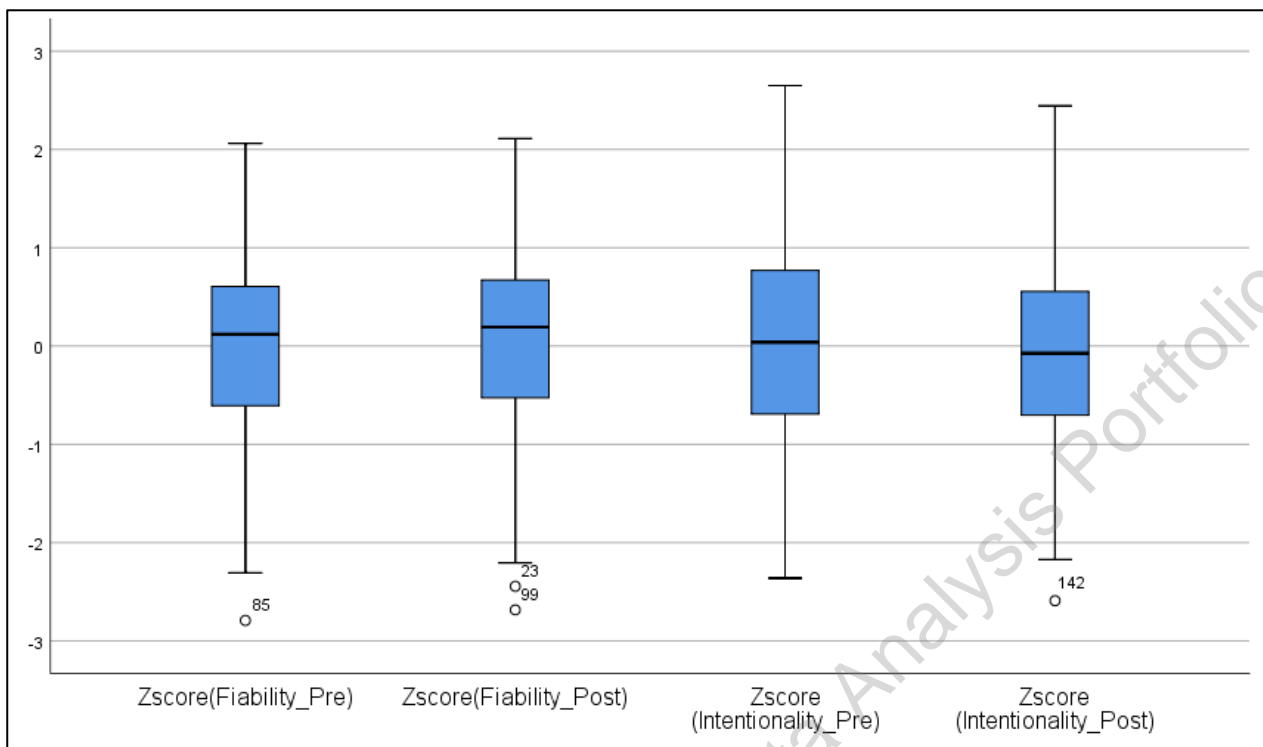
<i>Descriptive Statistics</i>					
	N	Skewness		Kurtosis	
	Statistic	Statistic	Std. Error	Statistic	Std. Error
Fiability_Pre	178	-0.327	0.182	0.072	0.362
Fiability_Post	178	-0.317	0.182	-0.021	0.362
Intentionality_Pre	178	0.029	0.182	-0.079	0.362
Intentionality_Post	178	-0.011	0.182	-0.091	0.362
Valid N (listwise)	178				

In order to provide more evidence of normality, a Shapiro Wilk's test was conducted, which is a test of hypothesis to assess normality. The table below shows that the results are non-significant considering the 99% confidence level ($p > .001$) for the four scales, which indicate no substantial depart from normality.

<i>Tests of Normality</i>			
	Shapiro-Wilk		
	Statistic	df	Sig.
Fiability_Pre	0.976	178	0.004
Fiability_Post	0.977	178	0.005
Intentionality_Pre	0.990	178	0.237
Intentionality_Post	0.987	178	0.110

Outliers

The existence of outliers can distort the results of ANOVA, so these need to be examined carefully. A first assessment was done using Box Plots. Box plots are graphical representations of the distribution of values in a particular variable. The graph literally box in observations that are around the median (horizontal line in the middle of the box). The box edges represent the interquartile range of values. That is, the 25th percentile (lowest edge) and the 75th percentile (highest edge). 50% of values lie inside the box. The whiskers (lines protruding from the box), represent the minimum and maximum values observed among the cases. The scores of the variables were standardized and the box plots are shown in the figure below.



Values above 4 or below -4 would represent outliers and should be deleted according to the guidelines of Hair et al. (2014). The figure above shows that the values are within the range of -3 and 3 which indicates the absence of extreme values in the data. Therefore, there were no cases excluded.

Repeated-measures ANOVA

The repeated-measures ANOVA method is useful when the same participants are being measured under different conditions (or measured at different points in time) (this is also referred to as a within-subjects design) (Pallant, 2010). When more than one dependent variable is under study and the researcher is interested on examining multivariate effects, then the test is called a repeated-measures multivariate ANOVA (MANOVA). When the dependent variable was measured only twice (which is the case here), the test may also take the form of a Paired-Samples T-test.

The table below shows the mean scores of the two scales for the pre and post intervention assessments. These differences will be tested for their statistical significance later in the report.

<i>Descriptive Statistics</i>					
	N	Minimum	Maximum	Mean	Std. Deviation
Fiability_Pre	178	1.00	5.00	3.301	0.824
Fiability_Post	178	1.00	5.00	3.239	0.834
Intentionality_Pre	178	1.00	5.00	2.886	0.798
Intentionality_Post	178	1.00	5.00	3.059	0.794
Valid N (listwise)	178				

The table above shows that scores of Fiability decreased with the intervention, while scores of Intentionality increased.

Multivariate Effects

The first test was to measure if there are significant effects of the intervention on the two brand trust scales together (multivariate analysis).

The within-subjects line of the following table shows that there is a significant multivariate effect of the Intervention, $F(2) = 18.732$, $p < .001$, which means that the scores of brand trust are significantly different between pre and post intervention.

Multivariate Tests^a

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
	Intercept	0.950	1654.802 ^b	2.000	176.000	0.000	0.950
Within Subjects	Intervention	0.176	18.732 ^b	2.000	176.000	0.000	0.176

a. Design: Intercept

Within Subjects Design: Intervention

b. Exact statistic

The partial eta squared column indicates the effect sizes, which can be interpreted as the following:

- 0.01: small
- 0.06: medium
- 0.138: large (Cohen, 1988)

The intervention's multivariate effect size was large ($\eta^2_p = .176$)

Univariate Within-Subjects Effects

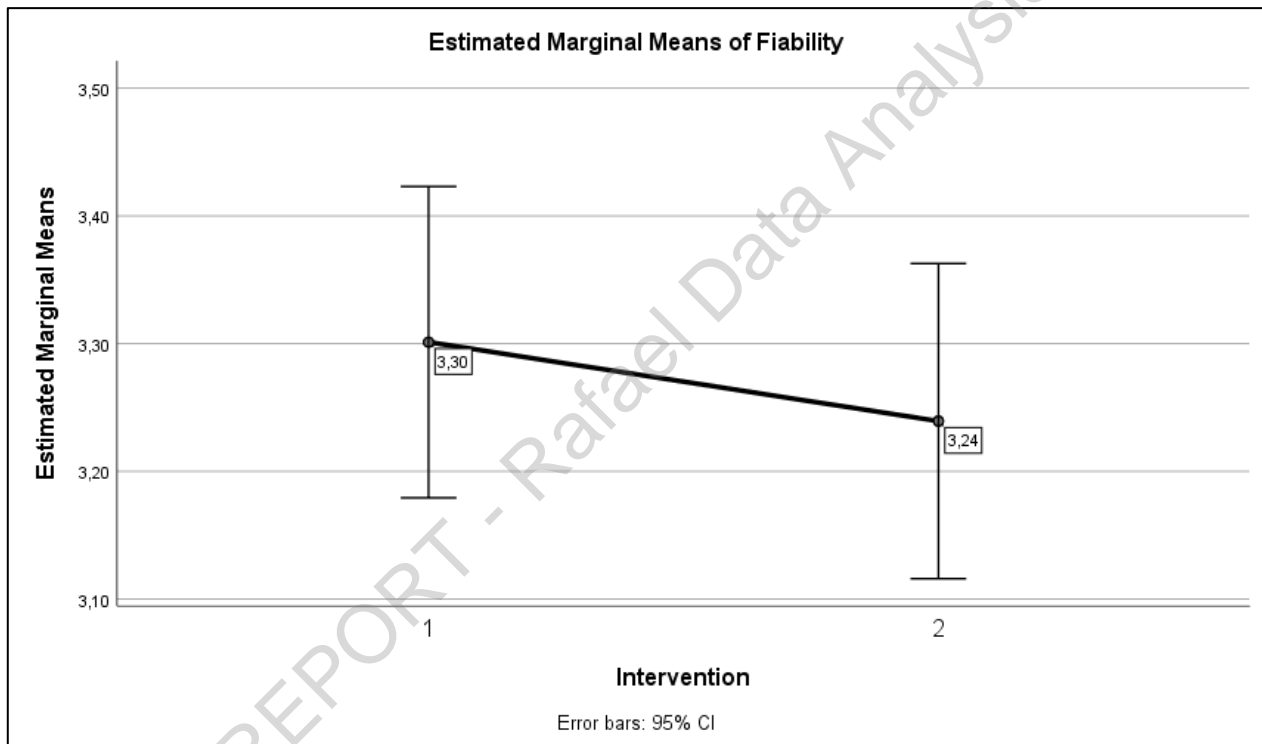
This section presents the univariate effects, which means looking at each brand trust indicator separately. The table below shows the results of both ANOVA models. The effect of the intervention on Intentionality was large and highly significant, $F(1) = 28.726$, $p < .001$, $\eta^2_p = .140$. Since earlier in the report it was shown that the scores of intentionality increase, we can conclude that the intervention has a positive effect on intentionality.

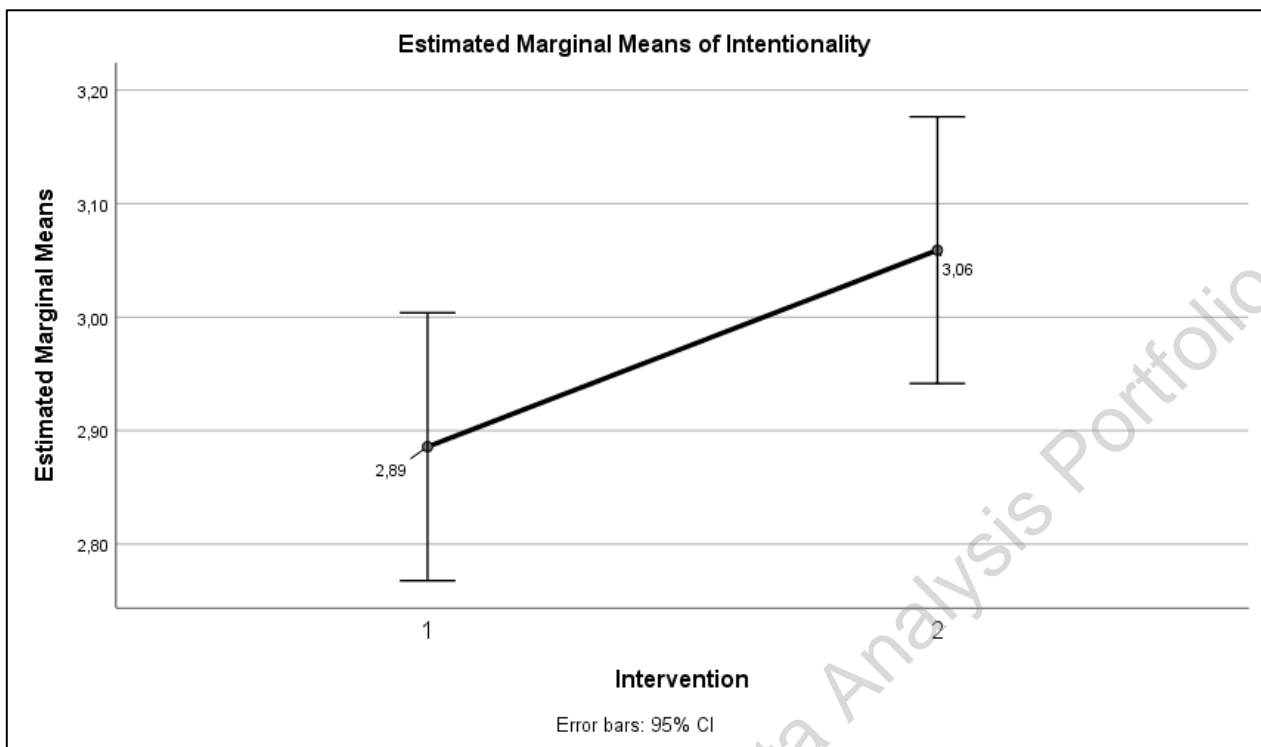
The effect on Fiability was small (weak) and the effect is only significant at the 90% confidence level, $F(1) = 2.755$, $p < .10$, $\eta^2_p = .015$. Since scores decrease, the conclusion is that the effect is negative.

Tests of Within-Subjects Contrasts

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intervention	Fiability	0.340	1	0.340	2.755	0.099	0.015
	Intentionality	2.670	1	2.670	28.726	0.000	0.140
Error(Intervention)	Fiability	21.840	177	0.123			
	Intentionality	16.455	177	0.093			

In order to better visualize the differences from pre to post intervention, the following figures were generated. Number 1 represents the pre-intervention period and 2 the post-intervention. The whiskers represent 95% confidence intervals for the mean scores.





The graph shows the decrease of fiability after intervention and the increase of intentionality, which was more substantial than the decrease of fiability. The main conclusion of the study is that the intervention significantly increases intentionality ($p < .001$) while significantly decreases fiability ($p < 0.10$).

References

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- Hair, J. F., Black, W., Babin, B., & Anderson, R. (2014). *Multivariate data analysis* (Seventh). Pearson Education, Inc.
- Pallant, J. (2010). *SPSS Survival Manual* (4th ed.). McGraw-Hill.