Analysis Report

This report is structured as follows.

Contents

Data Screening

Before moving to the execution of the models, data was screened for multivariate and univariate outliers. A pragmatic approach to identify multivariate outliers is suggested by Hair et al. (2014): Mahalanobis distances. These are calculated for the variables to be entered on the multiple regression analysis and their results are divided by the number of variables. When sample sizes are large (100+), coefficients above 3.5 or 4.0 can be considered outliers (Hair et al., 2014). In this study, Mahalanobis distances were calculated for all the survey items. Results were divided by 21 and the maximum distance was 3.000, suggesting no multivariate outliers were present in the data.

Missing cases were analyzed and 12 data cells were empty on 11 different variable. The maximum number of missing cases on a single variable was 2. This amount of missing data can be negligible. The empty cells were replaced by the variable's means.

Sample Characterisation

The table below shows the frequency of responses for each categorical variable being studied.

		Count	Column N %
Gender	Female	168	72.1%
	Male	61	26.2%
	Non-binary/third gender	1	0.4%
	Prefer not to say	3	1.3%
Age	18-25 year	36	15.5%
	26-33 year	44	18.9%
	34-41 year	28	12.0%
	42-49 year	34	14.6%
	50-57 year	41	17.6%
	58-65 year	38	16.3%
	66 year or older	12	5.2%
Education	Higher professional education (HBO)	117	50.4%
	No education completed	1	0.4%
	Pre-university education (VWO)	1	0.4%
	Pre-vocational secondary education (VMBO)	5	2.2%
	Secondary vocational education (MBO)	64	27.6%
	Senior general secondary education (HAVO)	14	6.0%
	University education (WO)	30	12.9%

Descriptive Statistics

The table below shows descriptive statistics for all variables under study. Skewness was within ±2 range in most of the cases and so was kurtosis, suggesting normality. A minor depart from normality was observed for AT2 and ESB3. The total sample size was 233 (N = 233). Items coded with the suffix '_INV' indicate items that were reverse-scored.

Descriptive Statistics

	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis
ESB1	1	4	3.343	.703	-1.045	1.395
ESB2	1	4	3.500	.630	-1.201	1.785
ESB3	1	4	3.463	.792	-1.612	2.253
ESI1	1	4	3.352	.758	-1.167	1.256
ESI2	1	4	1.664	.957	1.256	.380
ESI2_INV	1	4	3.336	.955	-1.258	.395
ESI3	1	4	3.004	.763	-.419	-.169
AT1	1	4	3.392	.599	-.546	.083
AT2	1	4	3.524	.695	-1.677	3.200
AT3	1	4	1.584	.745	1.037	.245
AT3_INV	1	4	3.416	.745	-1.037	.245
SN1	1	4	3.077	.659	-.631	1.269
SN2	1	4	1.961	.767	.354	-.471
SN2_INV	1	4	3.039	.767	-.354	-.471
SN3	1	4	3.176	.629	-.467	.819
PBC1	1	4	2.952	.744	-.302	-.244
PBC2	1	4	2.216	.924	.089	-1.024
PBC2_INV	1	4	2.784	.922	-.090	-1.016
PBC3	1	4	1.721	.806	1.047	.707
PBC3_INV	1	4	3.279	.806	-1.047	.707
PC1	1	4	2.189	.918	.020	-1.155
PC2	1	4	3.180	.789	-.969	.916
PC3	1	4	2.267	.936	.108	-.958
NC1	1	4	1.781	.820	.804	-.032
NC2	1	4	1.707	.825	.872	-.189
NC3	1	4	1.573	.773	1.191	.645

The test of the Measurement Model (Confirmatory Factor Analysis)

The main objective of testing a measurement model is to test construct validity. Construct Validity is defined as "the extent to which a set of measured variables actually represent the theoretical latent construct they are designed to measure" (Hair et al., 2014).

Confirmatory Factor Analysis (CFA) was used in the analysis, as it is an adequate method to be used as evidence of construct validity of theory-based instruments (Li, 2016). This type of analysis is used when a researcher wishes to confirm a specific pattern of variables that are predicted based on theory or previous analytical studies (DeVellis, 2012). That is, based on knowledge of the theory, he or she assumes the a priori factorial structure and then tests this hypothetical arrangement statistically (Byrne, 2016).

CFA was conducted using SPSS AMOS software v22, which uses Maximum Likelihood (ML) algorithm to estimate the results. ML is the most common method used to estimate parameters in CFA, because of its attractive statistical properties (i.e., asymptotic unbiasedness, normality, consistency, and maximal efficiency) (Li, 2016). After defining the model in the software and executing the analysis, four main phases were conducted to examine construct validity (1) assessment of model fit; (2) assessment of convergent validity; and (3) respecification of the model (if necessary). The statistics that were used to assess model fit and their rules of thumb are presented in the table below.

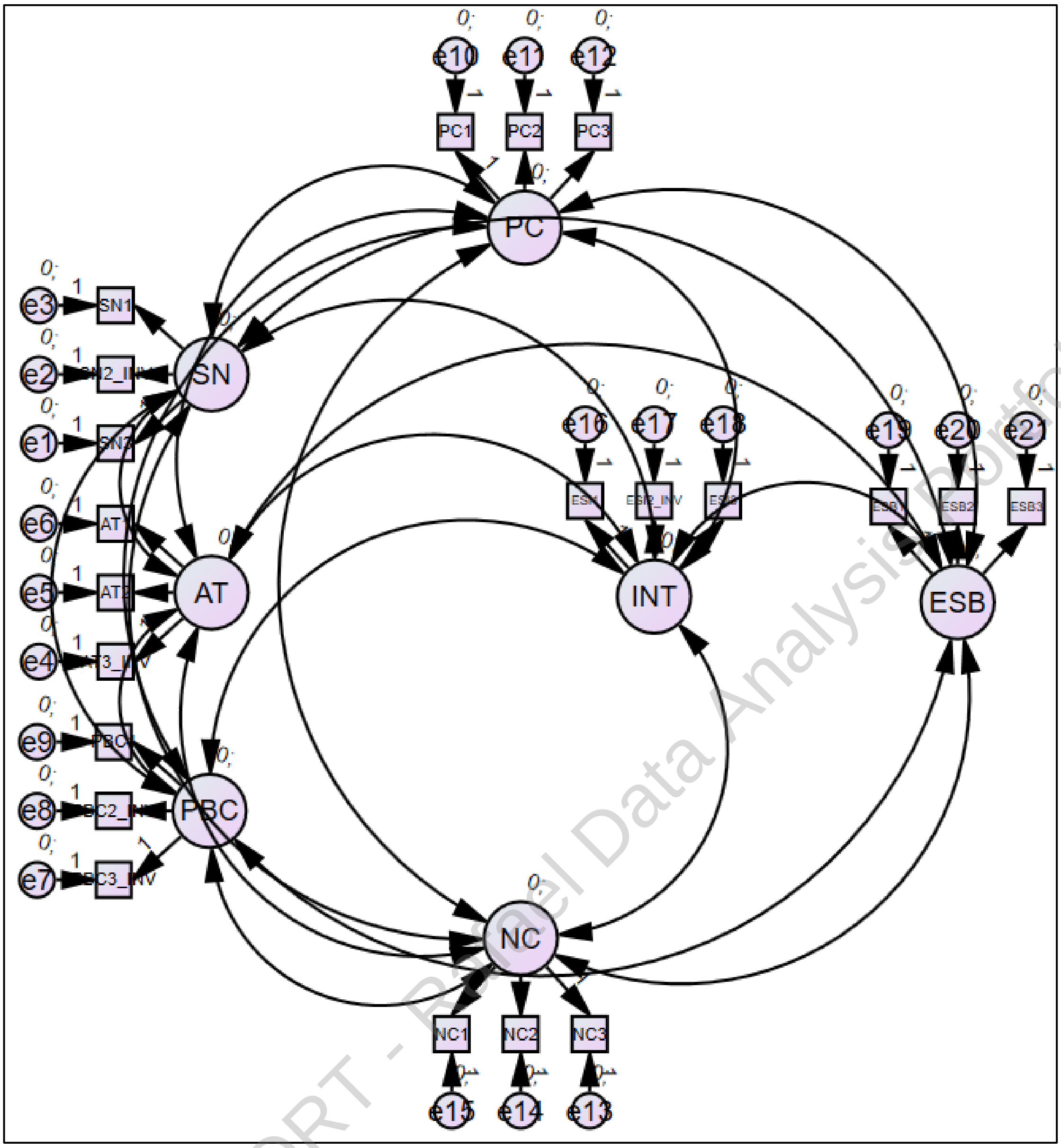
Fit index	**Rules of thumb**
Normed chi-square (χ^2/df)	The division between the chi-square value and the model's degrees of freedom should be less than 4.
Root mean square error of approximation (RMSEA)	RMSEA < 0.08
Comparative fit index (CFI)	CFI > 0.90
Normed fit index (NFI)	NFI > 0.90

Convergent validity refers to the "extent to which indicators of a specific construct converge or share a high proportion of variance in common" (Hair et al. 2014, p. 601). A second important concept to test is construct's reliability. This was done using the Composite Reliability index (which is based on factor loadings) and Cronbach's Alpha (which is based on correlations).

A summary of the indicators used to measure constructs' validity and reliability are detailed below (table below).

Indicator of convergent validity	Definition	Rules of thumb
Factor loadings (λ)	Correlation between the original variables and the factors, and the key to understanding the nature of a particular factor. Squared factor loadings indicate what percentage of the variance in an original variable is explained by a factor.	In the case of high convergent validity, high one-factor loadings would indicate that they converge on a common point, the latent construct. At a minimum, all factor loadings must be statistically significant. Because a significant load can still have quite weak strength, a good rule of thumb is that standardized loading estimates should be 0.5 or higher and ideally 0.7 or higher.
AVE	A summary measure of convergence among a set of items representing a latent construct. It is the average percentage of variation explained (variance extracted) among the items of a construct.	An AVE of 0.5 or higher is a good rule of thumb suggesting adequate convergence. An AVE of less than 0.5 indicates that, on average, more error remains in the items than variance explained by the latent factor structure imposed on the measure.
Indicator of internal consistency	**Definition**	**Rules of thumb**
Construct Reliability (CR)	Measure of reliability and internal consistency of the measured variables representing a latent construct. Must be established before construct validity can be assessed. It is computed from the squared sum of factor loadings for each construct and the sum of the error variance terms for a construct.	0.7 or higher suggests good reliability. Reliability between 0.6 and 0.7 may be acceptable, provided that other indicators of a model's construct validity are good.
Cronbach's Alpha	Cronbach's Alpha is a coefficient that represents the proportion of total variance among items that are due to the construct that they intend to measure	0.7 is the minimum acceptable level (Pallant, 2010).

The constructs present in the model below were tested for reliability and validity.



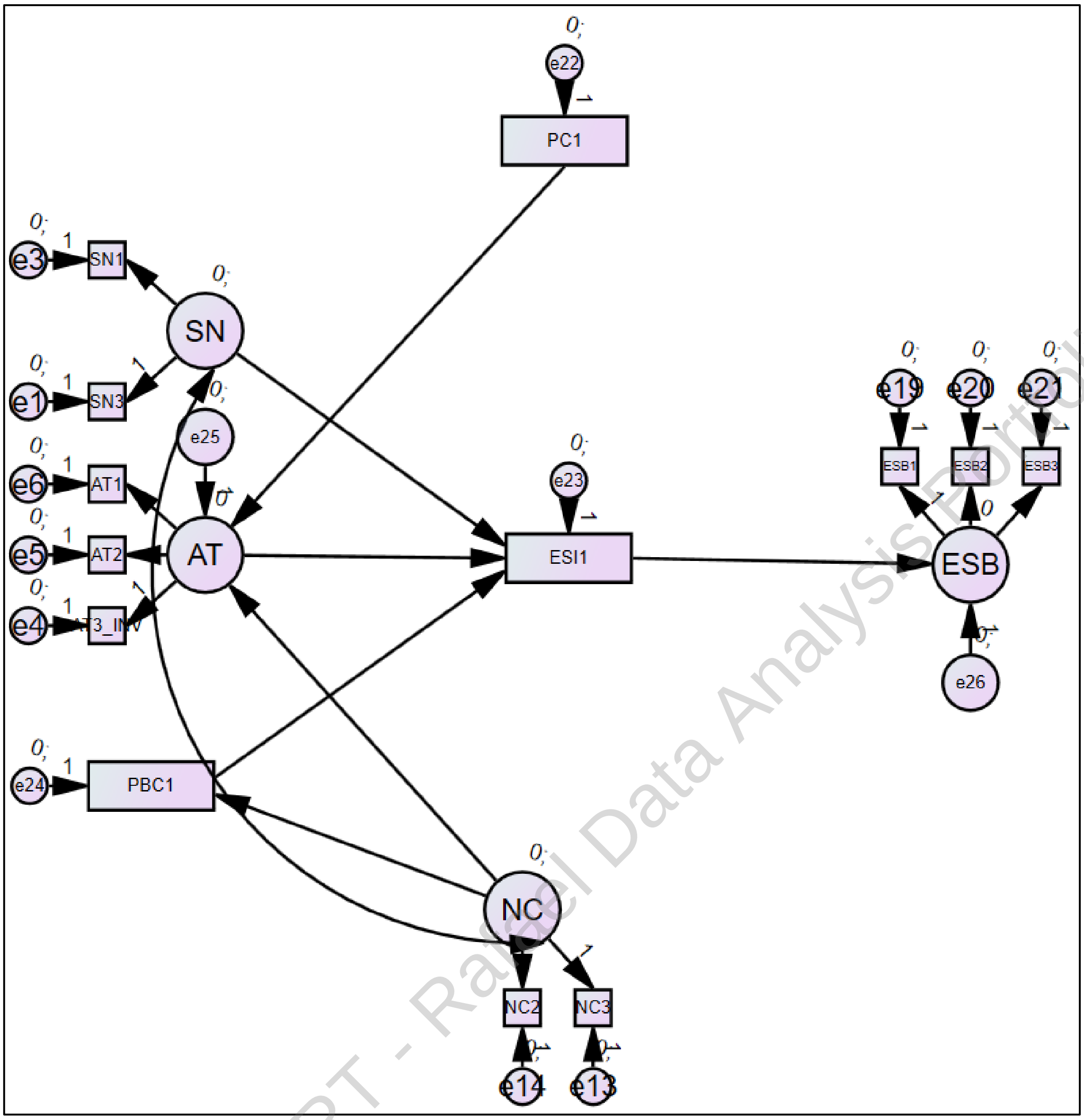
The measurement model did not fit due to a negative definite covariance matrix. The factor loadings were estimated and several factor loadings were below 0.400, indicating lack of convergent validity on several constructs. For instance,

NC1 (λ = 0.339), PC1 (λ = 0.303), PC3 (λ = 0.280), ESI2_INV (λ = 0.325), SN2_INV (λ = 0.299) and PBC2_INV (λ = 0.399) showed low factor loadings. These items were dropped and the model executed again, which showed acceptable fit. The PC construct was removed from the measurement model since two of its three items showed poor loadings, resulting on a single-item construct. On the second model, two constructs showed unacceptable indices of reliability: PBC (CR = 0.354) and ESI (CR = 0.470). They also did not demonstrate to have validity (AVE < 0.500). Since they were already composed by only two items, there was no option but to use single-item constructs to represent PBC and Intention. The final measurement model (without PC, PBC and Intention) showed good-fit ($\chi^2/df = 1.731$, RMSEA = 0.056, CFI = 0.957, NFI = 0.906, IFI = 0.958). Reliability was considered acceptable (CR > 0.600). All Cronbach's Alpha (α) were above 0.700, average variance extracted (AVE) were above 0.500 and composite reliabilities (CR) were above 0.700 (table below). All factor loadings (λ) were above 0.500, also suggesting appropriate convergent validity.

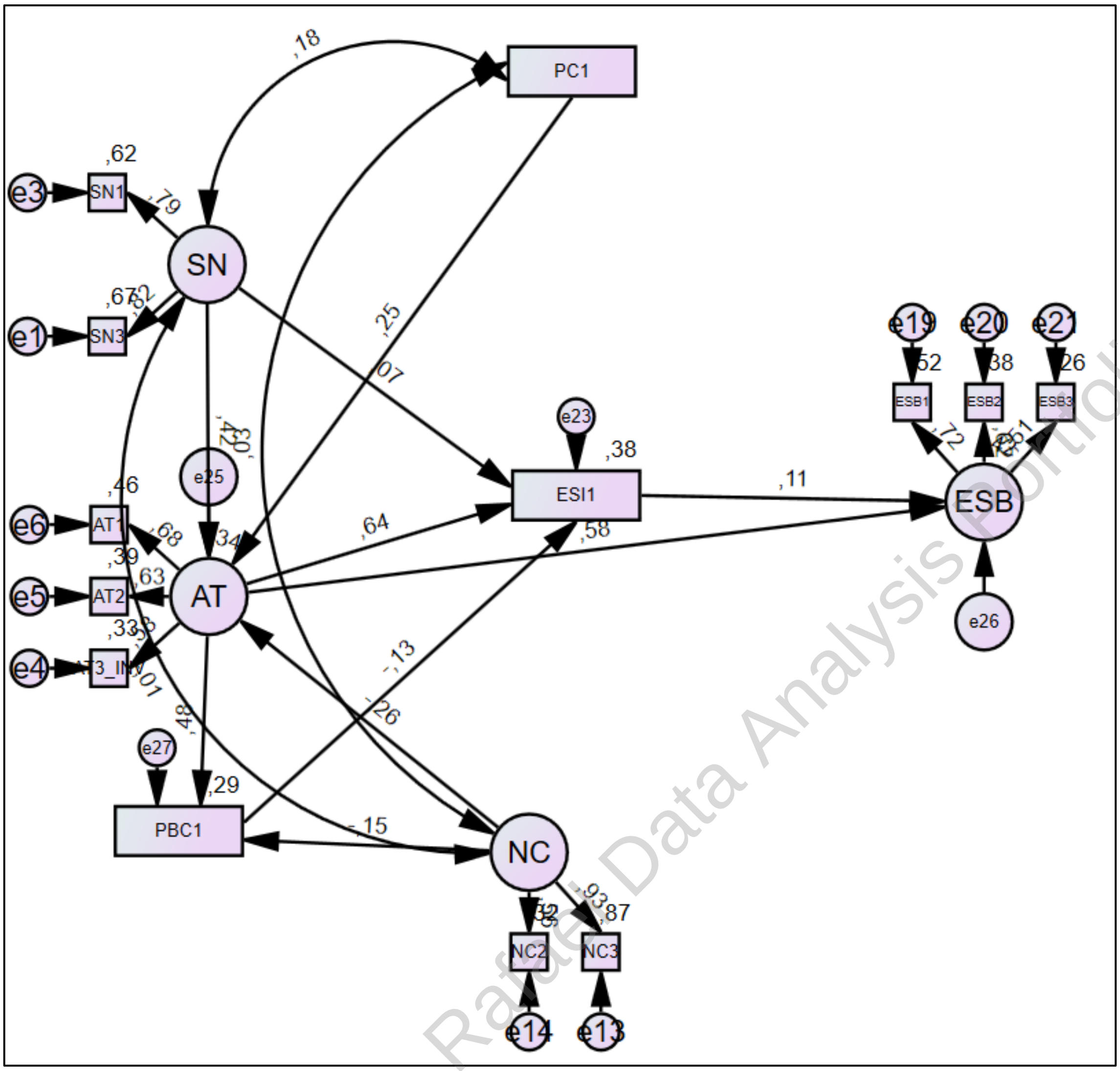
Item		Construct	λ	AVE	CR	α
SN1	<---	Subjective Norm	0.771	0.645	0.784	0.783
SN3	<---		0.834			
AT1	<---	Attitude	0.585	0.411	0.676	0.667
AT2	<---		0.702			
AT3_INV	<---		0.631			
NC2	<---	NC	0.718	0.526	0.690	0.690
NC3	<---		0.733			
ESB1	<---	Behavior	0.706	0.392	0.653	0.639
ESB2	<---		0.665			
ESB3	<---		0.484			

The test of the Structural Model (Confirmatory Factor Analysis)

The first tested model is represented in the figure below.



The model did not reach appropriate fit (χ^2/df = 3.187, RMSEA = 0.097, CFI = 0.802, NFI = 0.742, IFI = 0.807). The examination of Modification Indices suggested two paths that could modify the Chi-Square by more than 10.000. Subjective Norm predicting Attitudes (MI = 17.618) and Attitudes predicting PBC (MI = 13.441). These two paths were added to the model but CFI was still below 0.900, suggesting poor fit (χ^2/df = 2.283, RMSEA = 0.074, CFI = 0.892, NFI = 0.827, IFI = 0.895). The next largest modification index was inserting a path from Attitudes directly to Behavior (MI = 9.510). The resulting model finally showed good fit (χ^2/df = 1.861, RMSEA = 0.061, CFI = 0.929, NFI = 0.862, IFI = 0.931). The figure below represents the model with standardized beta coefficients.



Model coefficients are given in the table below. 42.4% of the variance of Behavior was explained by the model ($R^2 = 0.424$). NC had no effect on PBC ($p > 0.05$), but had a negative effect on attitudes ($\beta = -0.256$, $p < 0.05$). SN had no effect on Intention ($p > 0.05$). PBC did not affect Intention ($p > 0.05$) and, surprisingly, Intention has no effect on Behaviour ($p > 0.05$). Attitude has a direct effect on Behavior ($\beta = 0.578$, $p < 0.001$). PBC and SN both have positive effects on Attitude ($p < 0.05$).

			B	β	S.E.	C.R.	R^2	p
AT	<---	NC	-.152	-.256	.062	-2.450		.014
AT	<---	PC1	.115	.246	.034	3.378	.341	***
AT	<---	SN	.351	.422	.081	4.331		***
PBC1	<---	NC	-.155	-.151	.081	-1.910	.292	.056
PBC1	<---	AT	.837	.482	.153	5.478		***
ESI1	<---	SN	.101	.069	.119	.847		.397
ESI1	<---	AT	1.125	.637	.219	5.139	.379	***
ESI1	<---	PBC1	-.129	-.127	.076	-1.697		.090
ESB	<---	ESI1	.075	.112	.068	1.109	.424	.267
ESB	<---	AT	.685	.578	.158	4.326		***

***: $p < 0.001$

An examination of indirect effects showed that SN has an indirect effect on Behavior (z = 0.279, p = 0.001), as well as NC (z = -0.162, p = 0.004) and PC (z = 0.158, p = 0.002). PBC had no indirect effect on ESB whatsoever (z = -0.014, p = 0.161).

The model suggests that Behavior is highly explained by attitudes, directly. It also suggests that Norm, PC and NC are direct predictors of attitudes and indirect predictors of behavior.

References

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* Routledge.

DeVellis, R. F. (2012). *Scale development: Theory and Applications* (Third). SAGE Publications.

Hair, J. F., Black, W., Babin, B., & Anderson, R. (2014). *Multivariate data analysis* (Seventh). Pearson Education, Inc.

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Pallant, J. (2010). *SPSS Survival Manual* (4th ed.). McGraw-Hill.