180425 lolaarowolo.com A4

White Paper

THE EFFECTS OF SOCIAL CONSTRUCTS ON THE BEHAVIOR OF FEMALE STUDENTS

A Structural Equation Model Specification Search

Olufunlola Arowolo, Ph.D.

CONTENTS

Recommended citation:

Arowolo, O. A. (2018). *The effects of social sonstructs on the behavior of female students: A structural equation model specification search* (Vol. A4.180425) [Working paper]. https://lolaarowolo.com/white-papers

1. Introduction

This study is a specification search in the context of path analysis. The goal of this study is to examine how social constructs affect the cognitive behavior of female students in the eighth grade. This study is based on the empirical analyses conducted by Felson and Bohrnstedt (1979), which assesses how physical attractiveness influences impression formation in middle-school-aged children – who were asked to rate their peers on physical appearance. Felson and Bohrnstedt aimed at examining the possibility of reciprocal feedback between two sociocognitive correlates: physical attractiveness and academic ability. The significance of this study hinges on its potential to identify social constructs that affect self-esteem, and make adjustment to the school environment difficult for adolescent students. The researchers were able to study the influence that physical attractiveness has on impression formation by using children's ratings to examine the possibility of reciprocal feedback between physical attractiveness and academic ability. The data for this project is based on the correlation matrix obtained from the Felson and Bohrnstedt study.

2. Literature Review

Experimental studies in the field of psychology point to the importance of impression formation as a variable in measuring physical appearance (Felson and Bohrnstedt 1979). Boyatzis, Baloff, and Durieux (1998) explain that adolescents place a premium on attractiveness, perhaps because gender barriers begin to weaken during early adolescence (Schofield 1981), making friendships across genders, romance, and intimacy more common, and more important as children begin to mature. According to Dion, Berscheid, and Walster (1972), physical attractiveness influences how individuals are perceived – the more attractive a person is, the more likely said person will be judged positively on a wide range of personal characteristics (as cited by Felson and Bohrnstedt 1979).

Clifford and Walster (1973) found that teachers were likely to judge fifth-graders, of both sexes, as intelligent when the children were good-looking (as cited by Felson and Bohrnstedt 1979). A more recent study by Boyatzis, Baloff, and Durieux (1998) varied attractiveness and academic performance in four conditions – high attractiveness/high grades; high attractiveness/low grade – and found that students rated as attractive were more popular, irrespective of whether their grades were high or low.

Scrabis-Fletcher and Silverman (2017) studied dynamic processes that account for the interaction and filtration of the stimuli that students receive, and how said stimuli affects their behavior (p. 85). According to Scrabis-Fletcher and Silverman, differences exist between male and female students concerning perceptions of competence. Byrne et al. (1968) found a positive relationship between attractiveness and perceived intelligence for females, but a negative relationship for males (as cited in Felson and Bohrnstedt 1979, p. 386). However, contrary to the popularized hypothetical expectation that attractiveness is more important to girls than boys, Boyatzis, Baloff, and Durieux (1998) find that popularity ratings by girls and boys were not significantly different.

Based on Boyatzis, Baloff, and Durieux's findings, the overarching research question that emerges from the literature is: *Do perceptions of physical attractiveness affect how female students perceive the academic ability of their peers?*

2.1. General Model Configuration

2.1.1. Path Diagram

The direct and indirect effects of four exogenous predictors – *GPA, height, weight*, and *strangers' ratings of physical attractiveness* – were measured against two endogenous variables: *Perceptions of academic ability* and *perceptions of physical attractiveness*. The general path diagram to be evaluated is illustrated in Figure 1, which depicts the overall model representing academic ability, as proposed by Felson and Bohrnstedt (1979).

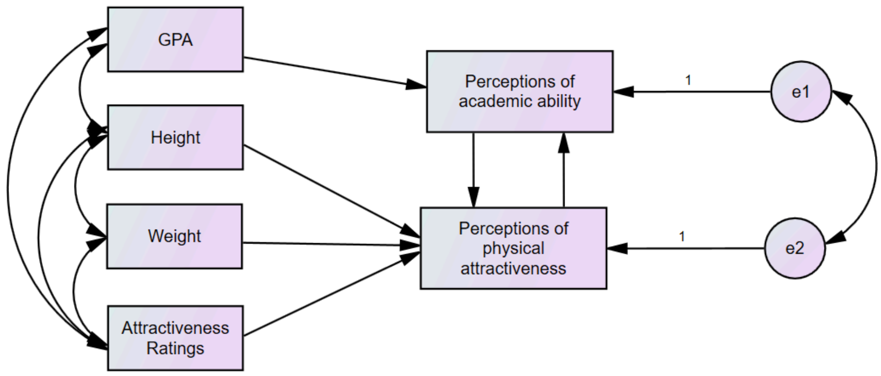


Figure 1. A nonrecursive model specification (Felson and Bohrnstedt 1979).

GPA predicts perceptions of academic ability. Height, weight, and strangers' ratings of physical attractiveness predict perceptions of physical attractiveness. There is a bi-directional relationship between the two endogenous variables, since they are predicting each other. To the right of the model is a pair of correlated disturbances.

2.1.2. Maximum Likelihood Equations

The goal of this study is to replicate the Felson and Bohrnstedt (1979) empirical analysis for females, as studies (Boyatzis, Baloff, and Durieux 1998) have found non-significant differences between genders, concerning attractiveness ratings. The following hypothetical expectations are tested:

H_1: *Female students' perceptions of physical attractiveness strongly affects their perceptions of academic ability.*

H_2: *A female student's perception of physical attractiveness can be predicted by a subject's height, weight, and physical appearance rating.*

The configuration of the models to be tested is based on a general maximum likelihood method for estimating a system of linear structural equations, which make it feasible to estimate a nonrecursive system once the order condition for identifiability[1] is met, as shown in Figure 1 (Joreskog 1973, as cited in Felson and Bohrnstedt 1979, p. 388). The structural equations representing Figure 1, as illustrated by Felson and Bohrnstedt 1979 are as follows:

$$\zeta_1 = \beta_2\zeta_2 + \gamma_1\nu_1 + \epsilon_1$$

$$\zeta_2 = \beta_1\zeta_1 + \gamma_2\nu_2 + \gamma_3\nu_3 + \gamma_4\nu_4 + \epsilon_2$$

where γ is the structural coefficient linking the endogenous variables to the exogenous variables; and, β represents the structural coefficient between the endogenous variables. The disturbance terms, ϵ, are allowed to correlate freely, and are assumed to be uncorrelated with the independent variables, with mean values of zero.

[1] Identifiability implies that there is one best value for each parameter in the model whose value is not known (Davis, 1993)

3. Methods

3.1. Data: Sample and measurement

The skewness values for normally distributed data should fall within the range of -1 and 1, which indicates that there are no issues with skewness in the data. Values for asymmetry and kurtosis between -2 and +2 are considered acceptable in order to prove normal univariate distribution (George & Mallery, 2010). For this study, procedural CFA for normal data was performed.

Correlation analysis was used as the initial test, after which, the relationship model was tested with an structural equation modeling (SEM) approach, using AMOS software. Model testing was performed with a two-step approach, as suggested by Byrne (2001). First, it validates the measurement model in terms of assessing the relationship between hypothesized latent constructs and clusters of observed variables underlying each construct (de Carvalho & Chima 2014). Validation of the measurement model is often conducted by using Confirmatory Factor Analysis (CFA) (de Carvalho & Chima 2014). The second step centers around fitting the structural model by measuring the significance of the relationship between latent variables, which is often accomplished through path analysis (Hoyle, 1995; Kaplan, 2000).

The data for this study comes from the article by Felson and Bohrnstedt (1979), which focuses on the relationship between children's perceptions of both ability and physical attractiveness. The authors sampled middle-schoolers – 209 girls and 207 boys – in a small midwestern city. Of the 416 student participants, 25 (or 6%) were black. Graduate students gave out questionnaires to the students, who then filled them out during a fourth period afternoon class. According to the school authorities, the students hailed from diverse socioeconomic backgrounds – the median incomes, in 1970 USD, ranged from $6,803 to $11, 638. Felson and Bohrnstedt gauged perceptions of physical attractiveness and ability (academic and athletic) using several sociometric items, as shown in Table 1 (1979, p. 387).

Grade-point Average (GPA) was measured by taking the sum of grade points that the students earned in four different academic subjects during the semester in which the questionnaire was distributed. Athletic skill was only assigned to boys and was based on a grade from a physical education class (basketball). Weight and height were self-reported. The average height of children of the same sex and weight were subtracted from each individual reported height, and weight was controlled in relation to the determined height, since weight is a function of height (p. 387). By viewing color photographs, strangers of the same grade, but from another city, rated attractiveness on a scale of 1 (least good-looking) to 5 (most good-looking).

Table 1. Measurements of study variables.

Variables	Description
Academic	Perceived academic ability, a sociometric measure based on the item: ***"Name who you think are your three smartest classmates".***
Athletic	Perceived athletic ability, a sociometric measure based on the item: ***"Name three of your classmates who you think are best at sports".***
Attract	Perceived attractiveness, a sociometric measure based on the item: ***"Name the three girls in the classroom who you think are the most good-looking (excluding yourself)".***
GPA	Grade-point average.
Height	Deviation of height from the mean height for a subject's grade and sex.
Weight	Weight, adjusted for height.
Rating	Ratings of physical attractiveness obtained by having children from another city rate photographs of the subjects.

Since direct access to the raw data used by Felson and Bohrnstedt (1979) was not an option, data for this study is based on the matrix summary data of 209 girls. To re-analyze the original study, a correlation matrix (with the relevant means and standard deviations) was imported directly into AMOS – this allowed for SEM analyses to be performed without having access to the actual raw data. Table 1.1 shows the summary matrix of sample correlations, means, and standard deviations of the six variables of interest, which were entered into an SPSS file, and then loaded into AMOS, for analyses.

Table 1.1. Matrix Summary Data.

	rowtype_	varname_	academic	athletic	attract	gpa	height	weight	rating
1	n		209	209	209	209	209	209	209
2	corr	academic	1	.	.	.	.	.	.
3	corr	athletic	0.43	1	.	.	.	.	.
4	corr	attract	0.5	0.48	1	.	.	.	.
5	corr	GPA	0.49	0.22	0.32	1	.	.	.
6	corr	height	0.1	-0.04	-0.03	0.18	1	.	.
7	corr	weight	0.04	0.02	-0.16	-0.1	0.34	1	.
8	corr	rating	0.09	0.14	0.43	0.15	-0.16	-0.27	1
9	stddev		0.16	0.07	0.49	3.49	2.91	19.32	1.01
10	mean		0.12	0.05	0.42	10.34	0	94.13	2.65

3.2. Model Testing: Initial Run

3.2.1. Fit Statistics

The fit statistics in Tables 2, 3, and 4 show a group of results that compare the study model against a baseline model. Essentially, the baseline model is a null model that assumes no relationship among the variables.

The Chi-square Goodness of Fit (χ^2) test compares the model-implied VCV matrix against the sample VCV matrix. If the chi-square test is statistically significant, then it signifies that we have a poor model fit of the data, which is indicative of model mis-specification. If the p-value is greater than 0.05, this signifies a good fit/specification of the model to the data. In Table 2, the first-line output for the default model shows CMIN = 2.761 is the chi-square value (with DF =2, $p > 0.05$). Generally, good-fitting models have low chi-square values, as shown in Table 2. Since the chi-square value is statistically insignificant, the model is a good fit of the data.

Table 2. Model Fit Summary.

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	19	2.761	2	.251	1.381
Saturated model	21	.000	0		
Independence model	6	228.800	15	.000	15.253

Table 2. Model Fit Summary (ct'd).

Root Mean Square Residual (RMR), GFI.

Model	RMR	GFI	AGFI	PGFI
Default model	.102	.996	.954	.095
Saturated model	.000	1.000		
Independence model	4.582	.731	.624	.522

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.988	.909	.997	.973	.996
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

One disadvantage of the Chi-square goodness of fit test is that it is impacted by sample size (as is the case for all inferential tests), and given that SEM is a large sample procedure, that often translates into statistical significance for this test, even when the discrepancy between the model-implied VCV and the sample VCV is fairly small. As a result, more descriptive indices are assessed, in conjunction with the chi-square value, to evaluate how well the model fits the data. CMIN/DF (χ^2/df) is a ratio of chi-square value and the degrees of freedom (DF) – it takes the chi-square value of the model and adjusts it for the DF.

The DF value reflects the complexity of the model, in terms of the estimated parameters. As more parameters are estimated, the degrees of freedom gets closer to zero. Theoretically, the overall fit of the model (to the data) can be increased by increasing the number of parameters to be estimated in the model. This means less parsimony for the model – in other words, model fit can be increased by decreasing parsimony. Thus the CMIN/DF ratio is a descriptive index that adjusts the chi-square value, based on the complexity of the model. Although not typically used, a CMIN/DF ratio close to zero represents a better model fit than higher values, but thresholds usually fall between 0-2 and 0-5 (Wheaton et al, 1977). In this study, the CMIN/DF ratio is 1.381, which falls well within either of threshold ranges, meaning that the model meets the standard, in terms of representing a good fit for the data.

According to Byrne (2010), the Goodness of Fit Index (GFI), shown at the second level of the Table 2 output, represents the "amount of variance and covariance in the sample covariance matrix" accounted for by the model (p.77). The GFI ranges between 0 and 1. The Adjusted Goodness of Fit Index (AGFI) is the GFI, adjusted for model complexity. So, when increasing/decreasing the number of parameters to fine tune the model, the AGFI adjusts for model complexity, whereas the GFI does not. For both indices, values above 0.90 are generally acceptable, and values of .95 or greater are of optimal fit. Based on these standards, Table 2 shows that the model GFI of 0.996 shows optimal fit of the model to the data, as does the AGFI of 0.954. The comparative fit index (CFI) and the rest of the remaining fit indices are all close to 1, indicative of a good model fit.

The null model tested in this study indicates poor fit with the data. This is indicated by the low goodness-of-fit index (GFI) value (GFI <0.90). In addition, the baseline comparisons in Table 2, show that the difference between the GFI, AGFI, CFI is negligible – this indicates that the model does not need to be modified anymore.

The Normed Fit Index (NFI), is not adjusted for model complexity, whereas the NNFI/TLI is adjusted for complexity. Typically, null model indices range from 0 (no fit) to 1 (optimal fit) – however, the Tucker-Lewis Index (TLI), also known as the Non-Normed Fit Index (NNFI), can surpass 1 (Kenny 2015). The Comparative Fit index is used more often when comparing the fit of two different models.

Barrett (2007) argues that fit indices add nothing to analyses and allow researchers to justify a misspecified model as acceptable. Some scholars (Hayduk, Cummings, Boadu, Pazderka-Robinson, & Boulianne, 2007) contend that fit index cutoffs can be misleading and subject to misuse. The counter argument is that only the chi-square value should be interpreted. However, Kenny and Mccoach (2003) find that fit indices favor models with a small number of variables, as is the case with this study.

Looking at the fit values for the null model – based on the standard that values above 0.90 are a good, and values of .95 or greater are of optimal fit – the NFI of 0.988 and the CFI of 0.996 can both be considered a good fit.

The root mean square error (RMSEA), shown in Table 3, takes into account the sample size and model complexity. Values between 0 and .05 represent a ”close’’ fit, and values around .08 represent and ”acceptable” fit. A poor fit would have a value greater than .08. These standards show that the model for this study, with an RMSEA of .043, shows a good fit of the model to the data.

Table 3 also shows the confidence intervals and a test of close-fit for both models. The LO90 and HI90 values represent the 90% confidence interval of the RMSEA value. The null value for a ’close’ fitting model is – an output value that falls within the range of the confidence interval means that the model exhibits a close fit to the data (Kenny, 2015).

Table 3. Model Fit Summary (Ct’d)

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.043	.000	.151	.417
Independence model	.262	.232	.292	.000

For the RMSEA confidence intervals, the lower 90% value should either include or be very near zero (or no worse than 0.05) and the upper value should be near .08 (Kenny, 2015). The LO90 value of .000 and HI90 value of .151 suggest a good fit.

The PCLOSE value is useful for evaluating the RMSEA sampling error, and represents a significance value (p-value), which tests whether the RMSEA value of the output is significantly different from 0.05, indicating a close fit (Kenny, 2015). A PCLOSE value less than 0.05 shows a lack of fit. The default has an insignificant PCLOSE value of .417, indicative of a close of fit.

Model estimations can then be compared using the AIC (Akaike Information Criterion), Browne-Cudeck criterion (BCC) or the Bayesian Information criterion (BIC). To evaluate the models based on these criteria, the better fitting model will be closer to zero, in relation to the worst fitting model. The AIC for the default model is 40.761 – this appears to be the better fitting model, as shown in Table 4.

Table 4. Model Fit Statistics (Cont'd)

AIC

Model	AIC	BCC	BIC	CAIC
Default model	40.761	42.084	104.265	123.265
Saturated model	42.000	43.463	112.189	133.189
Independence model	240.800	241.218	260.854	266.854

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	.196	.192	.238	.202
Saturated model	.202	.202	.202	.209
Independence model	1.158	.940	1.411	1.160

The Expected Cross-Validated Index (ECVI) is usually reported with the AIC. Bryne (2010) explains that the ECVI is a means of evaluating a single sample to assess the likelihood that a model cross-validates across similar sized samples from the same population. Since ECVI coefficients can take on any value, the model showing the smallest ECVI is most desirable, since there is no pre-determined threshold value. The default model has the lowest ECVI. This is consistent with the AIC fit determination, which also shows the default as the preferable model.

3.2.2. Model Estimates and Path Analyses

To assess the overall fit of the model, various fit statistics are considered to help with evaluating the fit of the model at the global level. In addition to assessing the global fit, as discussed in the preceding, the individual parameters can be assessed to determine how well the model parameters have been specified. Table 5 shows the unstandardized regression coefficients for the paths in the model for this study, as well as the standard errors (S.E.), and the critical ratio (C.R.). Three of the model estimates are insignificant: *height* predicting *attract, weight* predicting *attract, and attract* predicting *academic*.

Table 5. Maximum Likelihood Estimates

Unstandardized Regression Weights: (Group number 1 - Default model)

			Estimate	S.E	C.R.	P
attract	<---	height	.000	.010	.050	.960
attract	<---	weight	-.002	.001	-1.321	.186
attract	<---	rating	.176	.027	6.444	***
academic	<---	gpa	.023	.004	6.241	***
attract	<---	academic	1.607	.349	4.599	***
academic	<---	attract	-.002	.051	-.039	.969

The critical ratio is formed by dividing the regression estimate by the standard error. The C.R. is compared against the unit normal distribution or the z-values within the unit normal distribution (Kenny, 2015). The p-values shown are for a two-tailed significance test. The regression coefficient showing the relationship between *rating* and *attract* is .176 (S.E. = .027 C.R. = 6.444). A C.R. value of 1.96 or higher (or -1.96 and lower) means two-tailed significance at the 0.05 level, for the value in question (Hox and Bechger 1998). The C.R. of 6.444 for the relationship between *rating* and *attract* is statistically significant at the .01 level of significance. The paths from *gpa* to academic, and from from *academic* to *attract,* are also statistically significant at the .01 level of significance with a C.R. values of 6.241 and 4.599, respectively.

Figure 2 shows the initial run estimates from the null model. The p-values for the unstandardized estimates, shown in Table 5, also apply to the standardized regression weights shown in Table 6. Standardized coefficients represent regression coefficients and covariance metrics in terms of R^2. Standardized estimates and standard errors are independent of parameterization choices and on the choice of identifying scale parameters – i.e., whether or not the model is identified by setting ϕ=1, for example (Suhr, 2008).

Figure 2, the unstandardized solution, shows the exogenous predictor variables and their respective covariances. In the standardized model, the arrows between the predictors are the Pearson's Correlations. The correlations among the exogenous variables assume a standardized mean = 0, SD = 1; and, range from -1) perfect negative association) to +1 (perfect positive association). A zero value indicates no relationship. Cohen's (1988) standard is used to interpret the effect size of the correlations. A small effect will have an absolute value of .10 (|r| = .10); for a medium effect, |r| = .30; and, for a large effect |r| = .50 (Cohen 1988). Since both *e1* and *e2* are unobserved, their variances are unexplained by the predictors.

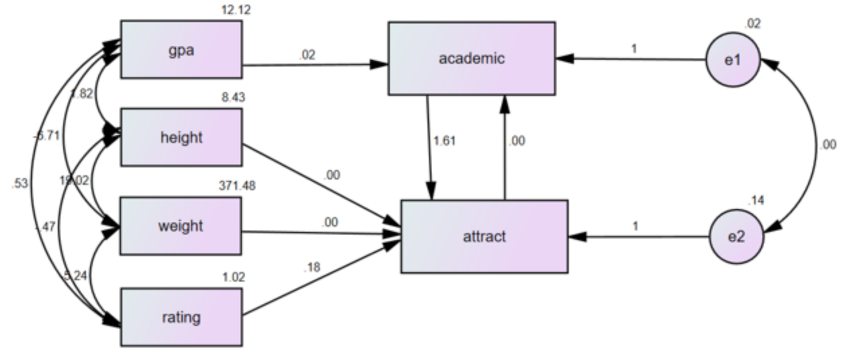


Figure 2. Initial run: Unstandardized estimates.

The Squared Multiple correlations show the *R^2* value associated with each of the endogenous variables in the model. The R^2 values in Table 6 represent the percent of variance explained by the predictor variables. In the initial run (Figure 3), the predictors of *attract* explains 40.2% of the variance in *height, weight,* and *ratings,* with 1 - .402 = .598 or 59.8% of the variance in these exogenous variables left unexplained.

Table 6. Maximum Likelihood Estimates

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
attract	<---	height	.003
attract	<---	weight	-.078
attract	<---	rating	.363
academic	<---	gpa	.492
attract	<---	academic	.525
academic	<---	attract	-.006

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
academic	.236
attract	.402

Figure 3 shows the initial run estimates of the standard model. The top-right values for the endogenous variables, *attract* and *academic*, are the squared multiple correlations (R^2 values).

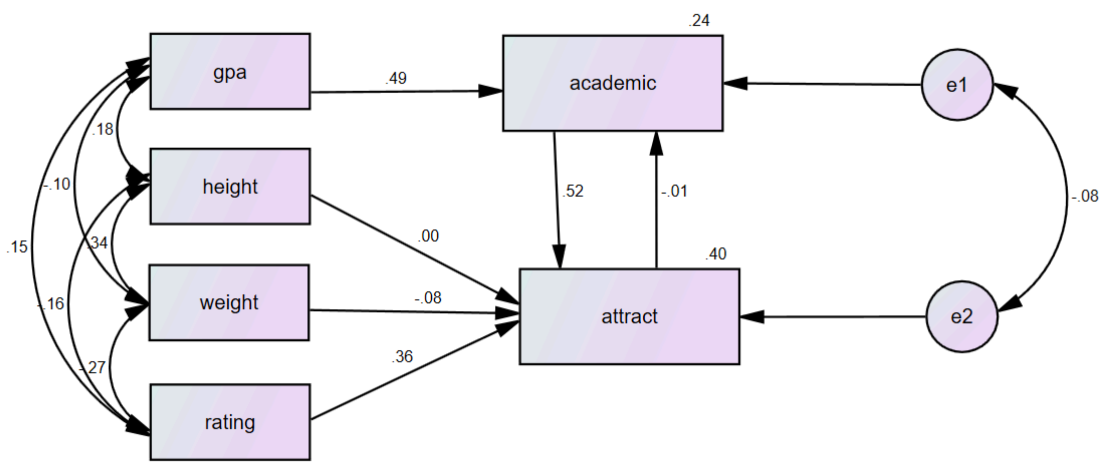


Figure 3. Initial run: Standardized estimates.

The standardized regression weights, or the standardized path coefficients, assess the relative predictive power of the model predictors. The standardized weights are also useful for making model comparisons (Kwan and Chan 2011) because they represent intra-variable relationships that have been standardized with a mean of zero and a standard deviation of 1.

The model in Figure 3 shows *height*, *weight*, and *rating* were specified as predictors of *attract*. The standardized regression weight for *rating* to *attract*, in Table 6 is .363, with a p-value less than .001 (as shown in Table 5, the unstandardized p-values are also applicable to the standardized coefficients). The strongest predictor of *attraction* is *rating*, which has the largest standardized path coefficient of .36, absolute value (S.E. = .027, C.R. = 6.444). However, the strongest path in the model, overall, is for the bi-directional predictive relationship between *academic* and *attract,* which has a standard weighted estimate of .525, as shown in Table 6, and in the path diagram (truncated to two decimal points).

3.3 Model Specification: Second Testing Phase

To find the model yielding the best fit, different model configurations are explored.

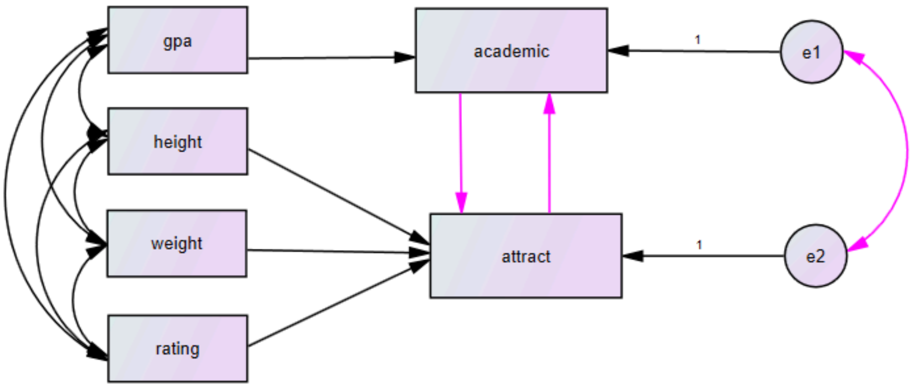
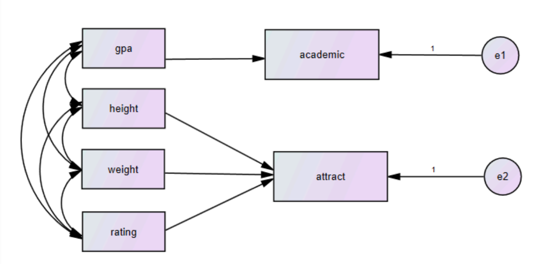


Figure 4. General model specification: Three possible optional paths.

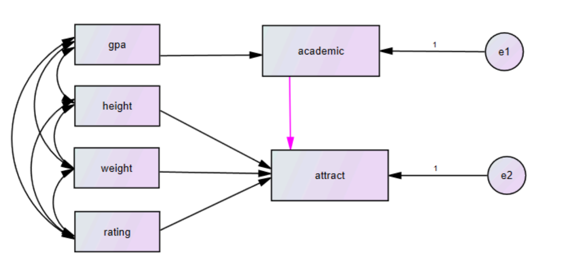
In this step, the correlated disturbances, and the paths between *academic* and *attract*, were treated as optional paths (or optional parameters). All three optional paths – in pink – are shown in Figure 4.

Figure 5. Model specification: Optional paths.

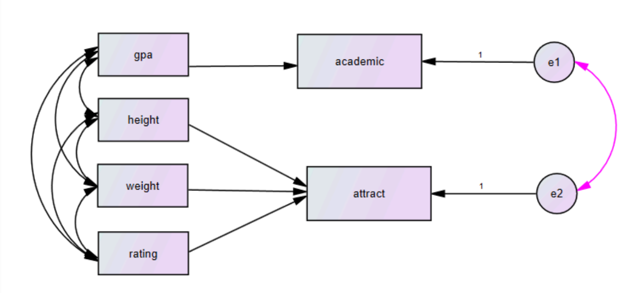
Default model 1



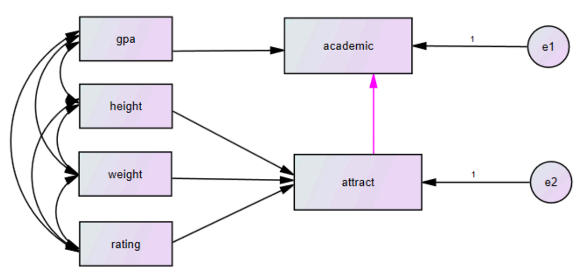
Default model 2



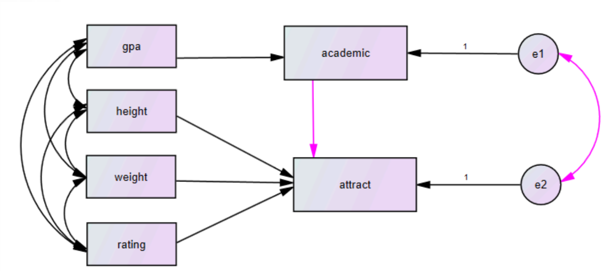
Default model 3



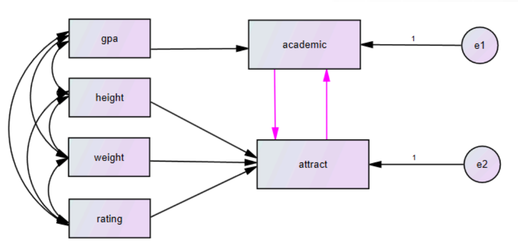
Default model 4



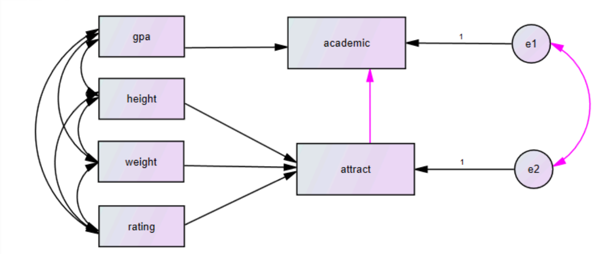
Default model 5



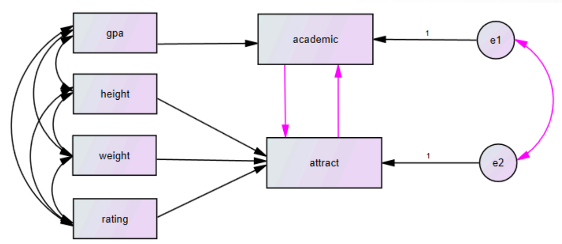
Default model 6



Default model 7



Default model 8



Model specification included/excluded optional paths. So, although the height and weight relationships to attract were insignificant in the initial run, both variables were left in the subsequent runs to see which model best fits when certain relationships are (in)excluded. Eight different model configurations have been specified (by AMOS) and the path diagrams for each new model generated are shown in Figure 5. The results of each new model specification shown in Figure 3 are summarized in Table 7, which shows that eight models have been generated, each with varying levels of fit.

Table 7. Model specification results.

Model	**Params**	**df**	**C**	**C-df**	**AIC**	**BCC**	**BIC**	**C/df**	**P**	**RMSEA**	**CFI 1**	**CFI 2**
Default 1	16	5	67.3	62.3	62.3	62.2	58.9	13.5	0.000	0.245	0.708	0.642
Default 2	17	4	3.07	**<u>-0.929</u>**	**<u>0.000</u>**	**<u>0.000</u>**	**<u>0.000</u>**	**<u>0.768</u>**	**<u>0.545</u>**	0.000	**<u>1.00</u>**	**<u>1.00</u>**
Default 3	17	4	19.2	15.2	16.1	16.1	16.1	4.80	0.001	0.135	0.929	0.913
Default 4	17	4	27.9	23.9	24.8	24.8	24.8	6.98	0.000	0.170	0.888	0.863
Default 5	18	3	2.76	-0.237	1.76	1.76	5.03	0.92	0.430	0.000	**<u>1.00</u>**	**<u>1.00</u>**
Default 6	18	3	2.90	-0.105	1.89	1.89	5.17	0.96	0.408	0.000	**<u>1.00</u>**	**<u>1.00</u>**
Default 7	18	3	18.2	16.2	18.2	18.2	21.4	6.35	0.000	0.161	0.924	0.907
Default 8	19	2	2.76	0.761	3.69	3.83	10.4	1.38	0.251	0.043	0.996	0.996
Sat.	21	0	0.00	0.000	4.93	5.21	18.3				**<u>1.00</u>**	**<u>1.00</u>**

For all the different model configurations shown, their parameters are being estimated in terms of correlations among the exogenous variables, but each specific default model varies in terms of which parameters are being included in the estimation process. Model 2 is the best fitting model (in bold and underlined), in terms of the AIC, BIC, BCC, and $\chi 2/df$ value. With respect to the chi-square value, model 8 is the best fitting. Model 2 (shown in Figure 5) only incorporates the path between academic and attract. Figures 6 and 7 show both the unstandardized and standardized path estimates for Model 2.

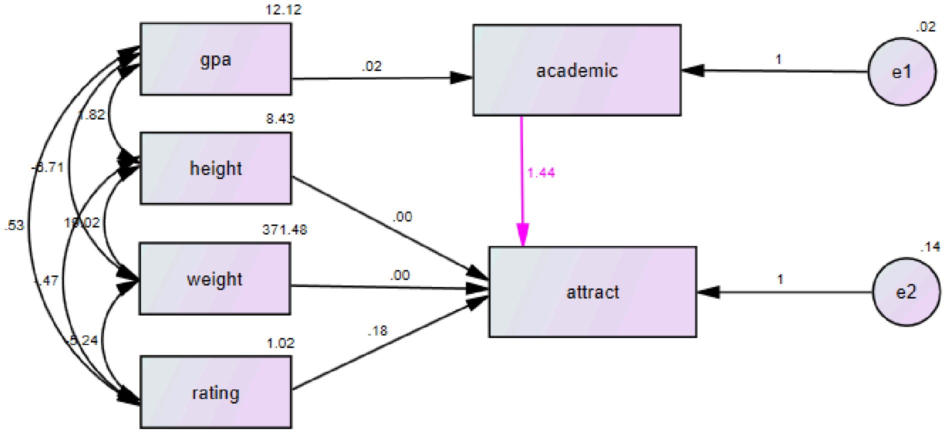


Figure 6. Best fitting model: Unstandardized path estimates for Model 2.

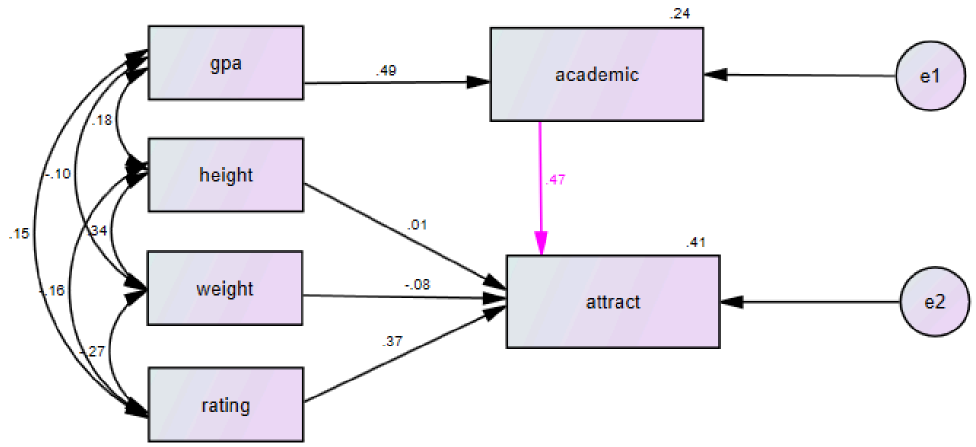


Figure 7. Best fitting model: Standardized path estimates for Model 2.

The results in the standardized solution shown in Figure 7 are consistent with the Felston and Bohrnstedt conclusion that perceptions of academic ability affect perceptions of attractiveness, but not the other way around. The effect size of the path from academic to attract is relatively large, with attract explaining roughly 41% of the variance in its associated predictors.

Re-visiting the study hypotheses:

H_1: *Female students' perceptions of physical attractiveness strongly affects their perceptions of academic ability.*

H_2: *A female student's perception of physical attractiveness can be predicted by a subject's height, weight, and physical appearance rating.*

We see that, overall, the hypothetical expectation that female students' perceptions of physical attractiveness strongly affects their perceptions of academic ability was not met. However, female student's perception of physical attractiveness were successfully predicted by a subject's height, weight, and physical appearance rating.

4 Limitations

The confirmatory technique used in SEM necessitates preliminary knowledge of all the relationships to be specified in the model (de Carvalho & Chima 2014). Knowing the number of parameters to be estimated – including covariances, path coefficients, and variances – is crucial, before data analyses can proceed (de Carvalho & Chima 2014). According to Kline (2005), SEM is a large sample technique, thus, conclusions extrapolated from a model based on a small sample size is unreliable as parameter estimation (variances, regression coefficients and covariances among variables) is often done by Maximum Likelihood (ML), which assumes normality among the indicator variables.

If a large model is tested with a small sample size – as is often the case in social sciences research where estimation involves less than a few hundred cases – this challenges the assumption of maximum likelihood estimation because it is based on large sample sizes. The sample size for this study, N = 209, is large, according to Hoogland and Boomsma 1998. However, the scope of the study is rather narrow – the focus was solely on females, which does not offer a full view of the behavior of the model across genders. However, the study does solidify the main hypothetical expectation that female perceptions of physical attractiveness affect perceptions of academic ability, which could be of importance to child psychologists studying sociocognitive issues related to adolescent females.

5 Conclusion

As aforementioned, the results are consistent with the Felston and Bohrnstedt conclusion that perceptions of ability affect perceptions of attractiveness, but not the other way around. This is opposite of the main hypothetical expectation of this study, which is that female perceptions of physical attractiveness affect perceptions of academic ability. Felston and Bohrnstedt believe that this result can be explained, in large part, by the notion that stimuli related to physical attractiveness is relatively ambiguous, as compared with stimuli related to academic ability (p. 386).

The Felston and Bohrnstedt study made an important contribution by explaining how perceptions of physical attractiveness affects students' perceptions of academic ability in their peers. Future research could assess how a child's own perceived sense of attractiveness affects their own academic performance.

References

Barrett, P. (2007). Structural equation modelling: adjudging model fit. Personality and Individual Differences, 42, 815–824.

Boyatzis, C., Baloff, P., & Durieux, C. (1998). Effects of Perceived Attractiveness and Academic Success on Early Adolescent Peer Popularity. The Journal of Genetic Psychology, 159(3), 337-344.

Byrne, D. O., O. London, and K. Reeves. (1968). The effects of physical attractiveness, sex and attitude similarity on interpersonal attraction. *Journal of Personality, 36*: 259-271.

Byrne, B. M. (2010). Structural equation modeling with AMOS, (2nd ed.). New York: Routledge.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Clifford, M. M., and E. Walster (1973). The effect of physical attractiveness on teacher expectation. *Sociology of Education, 46*: 248-258.

de Carvalho, J. & Chima, F. (2014). Applications of Structural Equation Modeling In Social Sciences Research. American International Journal of Contemporary Research. Vol. 4 No. 1 (January 31, 2014).

Davis, W. R. (1993). The FC1 rule of identification for confirmatory factor analysis: A general sufficient condition. Sociological Methods & Research, 21(4), 403-437.

Dion, K., E. Berscheid, and E. Walster. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*: 285-290.

Felson, R., & Bohrnstedt, G. (1979). "Are the Good Beautiful or the Beautiful Good?" The Relationship between Children's Perceptions of Ability and Perceptions of Physical Attractiveness. Social Psychology Quarterly, 42(4), 386-392.

George, D., & Mallery, M. (2010). SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update (10a ed.) Boston: Pearson.

Hayduk, L., Cummings, G., Boadu, K., Pazderka-Robinson, H., Boulianne, S. (2007). Testing! Testing! One, two, three--Testing the theory in structural equation models! Personality and Individual Differences, 2007, Vol.42 (5), p.841-851.

Hoyle, R. (1995). The structural equation modeling approach: Basic concepts and fundamental issues. In R. H. Hoyle (Ed.), Structural Equation modeling: Concepts, issues, and applications (pp. 1-15). Thousand Oaks CA: Sage.

Hox, J. J, & Bechger, T. M. (1998). An introduction to structural equation modeling. Retrieved from: http://joophox.net/publist/semfamre.pdf.

Joreskog, K. G. (1973). A general method for estimating a linear structural equation system. Pp. 85-112 in A. S. Goldberger and 0. D. Duncan (eds.), Structural Equation Models in the Social Sciences. New York: Seminar Press.

Kaplan, D. (2000). Structural equation modeling: Foundations and extensions. Thousand Oaks, CA: Sage Publications.

Kenny, D. A. (2015). Measuring Model Fit. Retrieved from: http://davidakenny.net/cm/fit.htm.

Kenny, D., & Mccoach, D. (2003). Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling. Structural Equation Modeling: A Multidisciplinary Journal, 10(3), 333-351.

Kwan, J.L.Y. & Chan, W. (2011). Comparing standardized coefficients in structural equation modeling: a model reparameterization approach. Behavior Research Methods *43*: 730. https://doi.org/10.3758/s13428-011-0088-6.

Schofield, J. W. (1981). Complementary and conflicting identities: Images and interaction in an interracial school. In S. R. Asher & J. M. Gottman (Eds.). *The development of children's friendships.* (pp. 53-90). New York: Cambridge University Press.

Scrabis-Fletcher, K. & Silverman, S. (2017). Student Perception of Competence and Attitude in Middle School Physical Education. Physical Educator, 74(1), 85-104.

Suhr, D. (2008). Step your way through Path Analysis. University of Northern Colorado. Retrieved: https://pdfs.semanticscholar.org/aeb5/4396b979b6351cc09a4d0a9109a882de6e67.pdf.

Wheaton, B., Muthen, B., Alwin, D., F., and Summers, G. (1977). Assessing Reliability and Stability in Panel Models. *Sociological Methodology, 8* (1), 84-136.