Validation of Questionnaire Estimating Predictors of Behavior Intention to Engage in First Sexual Intercourse Among Eighth-grade Pupils

**Aim** To use structural equation modeling in testing the construct validity of the questionnaire on the first sexual intercourse among young adolescents.

**Methods** Previously created questionnaire for the estimation of different factors influencing the intention to engage in the first sexual intercourse was validated. The data were gathered anonymously from a nationally representative sample of 1217 elementary school pupils aged 14-15 in Slovenia. The construct validity was determined by the structural equation modeling, LISREL 8.7.

**Results** The reliability of the questionnaire was satisfactory (Cronbach α = 0.73). Using the structural equation modeling analyses, the fit between empirical and theoretical models was confirmed, with 18 variables (17 independent and 1 dependent) for 6 latent indicators (5 external and 1 internal). The measures of goodness of fit were $\chi^2_{(80)} = 57.040; P = 0.976; \text{RMSEA} < 0.001; \text{RMR} = 0.0249; \text{GFI} = 0.994; \text{AGFI} = 0.988; \text{NFI} = 0.993; \text{RFI} = 0.986$. The instrument explained 34% of the variance in behavior intention.

**Conclusion** The construct validity of the questionnaire on the intention to engage in the first sexual intercourse among young adolescents was satisfactory. This questionnaire could be a useful tool in health promotion programs.

Polonca Brčar¹, Mojca Zvezdna Dernovšek²
¹Community Health Center Ljubljana, Ljubljana, Slovenia
²University Psychiatric Hospital, Ljubljana Polje, Slovenia

Received: July 30, 2009
Accepted: October 2, 2009

**Correspondence to:**
Polonica Brčar
Community Health Center
Ljubljana
Metelkova 9
1000 Ljubljana, Slovenia
brcarp@voljo.net
A person’s sexual behavior, including behavior related to sexual initiation, is influenced by physical and psychological characteristics of an individual, written and unwritten moral norms and rules, and social and cultural beliefs and behaviors (1,2). Therefore, it is important to identify the factors that influence such behavior and incorporate the obtained knowledge in preventive activities.

A number of theories exist to explain behavior and influences on behavior. The Theory of Planned Behavior (TPB) is the most popular socio-cognitive model, developed by Ajzen in 1985 (3). It is based on the theory of reasoned action of Ajzen and Fishbein (3-5), which proposes that people are normally reasonable and use available information in a predictable way. According to TPB, behavior is directly determined by intention to engage in behavior and perception of control over performance of behavior (3-5). The behavioral intention can be mathematically predicted from a linear combination of 3 variables (Figure 1): 1) attitudes toward a given behavior, which are a function of salient beliefs about consequences of this behavior, weighted by evaluation of the consequences; 2) subjective norms, which refer to the persons’ willingness to accept beliefs of important others and their motivation to comply with the expectations of important others; and 3) perceived behavioral control, which indicates how effective a person is in controlling his or her behavior (3-5).

Ajzen’s model proved to be suitable for evaluation of determinants of a healthy life-style and is widely used in studying different types of healthy behavior: condom use, use of contraceptives, legal and illegal drug use, physical activity, dietary behaviors, sun protective behaviors, screening attendances, breast/testicular self-examination, and adherence to medications (5).

Studying the determinants of behavioral intention to engage in the first sexual intercourse is important for understanding sexual behavior of youth and can be used in the creation of healthy behavior promotion programs (6-9). In our previous study, we used TPB model to develop a questionnaire for this purpose (10). The aim of the present study is to perform validity analysis of the questionnaire – to determine whether the questionnaire assesses what it was designed to assess. For validation of the questionnaire, we used structural equation modeling (SEM) analysis (11-13), which is one of multivariate techniques for the evaluation of construct validity of a questionnaire (6,13-18). SEM directly incorporates measurement error in the estimation process and simultaneously estimates several interrelated dependence relationships (19).

The aim of the study was to use SEM in testing the construct validity of the questionnaire on engaging into the first sexual intercourse among young adolescents.

METHODS

Questionnaire

In the previous study a questionnaire was constructed according to the TPB model to measure the predictors of the intention to engage in the first sexual intercourse (10).

The first part of the questionnaire consists of a short story depicting a situation from adolescents’ life, with the purpose to evoke empathy in participants (web-extra materi-
After the story, there is a set of questions on participants’ attitude toward engaging into the first sexual intercourse; subjective norms; perceived behavioral control, and behavioral intention to engage in the first sexual intercourse. Questions are closed-ended and answers are given on a 6-point rating scale. The last question is about participants’ own sexual experience (web-extra material).

**Population**

The studied population was that of eighth-grade Slovene pupils. Their mean age ± standard deviation was 14.4 ± 0.3 in 1996/97 school year (n = 26,406 pupils; 50.2% girls and 49.8% boys). The Institute for Employment of the Republic of Slovenia selected every 446th pupil from the register of eighth-graders. The entire classes of the selected 59 pupils were included in the sample. In this way, 59 eighth-grade classes from 59 schools throughout Slovenia, ie, 1395 pupils, were chosen (corresponding to our target sample of 1200). The sample is representative of Slovenia and comprises 5.3% of the observed eighth-grade population. The response rate was 87.2%. The final number of usable units of the sample was 1217, ie, 4.6%) of the total eighth-grade population (N = 26,406). Hundred and seventy eight questionnaires were not included in the analysis because pupils did not agree to participate or did not complete the study procedure.

**Ethical issues and procedure**

The Commission for Medical Ethics at the Ministry of Health of the Republic of Slovenia approved the research (permission No. 17/05/97). The majority of participants in the survey were younger than 15, so according to statutory requirements, written consent from their parents was obtained. Participants had the right to leave the study at any time.

The questionnaire was distributed by school nurses and physicians, who were instructed in the procedure. They informed the participants about the purpose, content, procedure, and anonymous and voluntary nature of the questionnaire. After filling out the questionnaire, each participant sealed the envelope containing the questionnaire and returned it to the data collector. Questionnaires were completed during one school hour by all the pupils present in the class at the time of the survey. Polling took place over two days, one month prior to the end of school year, ie, conclusion of the adolescents’ compulsory education.

**Data analysis**

In descriptive statistical analysis, mean, median, standard deviation, skewness, kurtosis, and percentiles (25th and 75th) were calculated to exclude dichotomous variables. In further statistical analysis, the association between individual variables was tested using Pearson matrix. The probability level lower than 0.05 was considered statistically significant. The whole Pearson Matrix was tested by Bartlett test of sphericity, and the probability level lower than 0.001 was considered significant. The anti-image correlation was used to analyze the partial correlations between variables and all variables with correlation higher than 0.50 had to be eliminated. In this way, we excluded 2 variables because the number of missing data exceeded 30%, 1 variable had no correlation in the matrix, and 26 variables were merged into 13 variables. The 41 variables were tested with SEM to determine whether the data collected with our instrument showed similarity with the theoretical model on which we constructed the instrument. The statistical procedure is presented in Figure 2.

The final instrument was constructed with the use of step-down procedure (ie, by entering all of the 41 variables at the beginning and later eliminating non-significant variables in each iteration).

The SEM is a multivariate technique combining aspects of multiple regression (examining dependence relationships).
and factor analysis (representing unmeasured concepts – factors – with multiple variables) to estimate a series of interrelated dependence relationships simultaneously. The SEM has 2 advantages: the ability to directly incorporate measurement error in the estimation process and the simultaneous estimation of several interrelated dependence relationships (19).

To determine if empirical and structural model fitted well, we used 3 measures. The absolute fit measures allow the determination of the degree to which the overall model (structural and measurement model) predicts the observed covariance and correlation matrix. Thus, low χ² values, resulting in significance levels greater than 0.2, indicate that the actual and predicted input matrices are not statistically different.

The direct and indirect relationships significance in the model was evaluated by testing the statistical and practical significance. One-way correlations between indicators and corresponding latent variables of endogenous and exogenous construct were determined with loadings (lambda x and lambda y). Latent variable is the operationalization of a construct in structural equation modeling. A latent variable cannot be measured directly but can be represented or measured by one or more variables.

Standardized parameters of latent variables in the structure were used to estimate the fit of empirical and structural model, and this is represented by calculation of the percent of explained variance from standardized parameters of latent variables. SEM demands the complete data in each unit (19). Reliability was measured by Cronbach α.

The data were analyzed using SPSS, version 11.5 (SPSS Inc., Chicago, IL, USA) and LISREL, version 8.7 (Jöreskog&Sörbom, 2004), Windows application for structural equation modeling and other multilevel techniques.

RESULTS

Table in the web extra material presents variables’ mean, median, standard deviation, skewness, kurtosis, percentiles (25th and 75th), and the percentage of missing data. There were no variables with normal distribution of data but the deviation of skewness and kurtosis allowed us to include all of them in further analysis. Two variables were excluded due to over 30% of missing data (V19prist – “probability of priest’s consent to sexual intercourse” and V20prist – “motivation to comply with the priest’s opinion on decision to engage in sexual intercourse”).

All 57 variables were analyzed by linear regression. In a correlation matrix, 62.5% of correlations were significant at the 0.05 level. The whole Pearson matrix was significant (χ² = 16966.46, P < 0.001). All variables had at least one correlation coefficient higher than 0.30. The correlation coefficient between some variables were not high enough to select variables for the multivariate analysis using only Pearson correlation. An additional analysis of the correlation between individual variables gave us the basis for merging them. Partial correlation coefficient in highly correlated variables ranged between 0.656 and 0.807. According to these results, 26 variables were merged and the following 13 were created: evaluation of the memory of the first intimate sexual experience – good/pleasant; evaluation of the deepened understanding between partners – good/pleasant; evaluation of the end of tension before the first sexual intercourse – good/pleasant; evaluation of the maturity feeling – good/pleasant; evaluation of collapse of friendship bonds – good/pleasant; evaluation of having guilty conscience because of premature sexual intercourse – good/pleasant, evaluation of the fear of the reaction of parents – good/pleasant; evaluation of feeling of shame because of sinful behavior – good/pleasant; evaluation of feeling of membership to the peer group – good/pleasant; evaluation of the deepened understanding between partners – good/pleasant; likelihood of father/mother’s agreement with the decision to engage in the first sexual intercourse; and motivation to respect father/mother’s decision on the first sexual intercourse. We also introduced a new dependent variable: “How possible is it that you would behave in a manner described in the story now or in the next 12 months?”.

Structural equation modeling approach for predicting behavior intention

For 41 variables, validity and reliability values were calculated and 16 variables with significant impact on dependant variable were chosen, the variables with 5% and 10% of error probability were selected. “Past sexual experience” was included in the multivariate model as an independent variable too, because this dimension in SEM model significantly influences the variability of behavioral intention to engage in the first sexual intercourse. For all 18 variables (17 independent [predictive] and 1 dependent variable), the practical significant correlations inside their own di-
mension were calculated (≥0.30) and included in the SEM. For the 18 items for 6 latent variables (5 external and 1 internal), the number of degrees of freedom in the structural model was 80 (n = 836).

Determinants of fit between empirical and structural model

Determinants of fit between empirical and structural model were assessed in 3 ways:

1. Measures of absolute fit. The results of the likelihood-ratio χ² (χ² = 57.040; P = 0.976) confirmed that differences between models were not significant. The difference between predicted structure and data was minimal; the root mean square residual (0.025) and the root mean square error of approximation (<0.001) were low (values under 0.08 were considered acceptable). The goodness of fit index, which represents the overall degree of fit (the squared residuals from prediction compared with the actual data) had the value near one, which indicated perfect fit (0.994), but was not adjusted for the degrees of freedom. The next 3 indices were incremental fit measures that compared the proposed model with the null model (the ideal model). The adjusted goodness of fit index (0.988), an extension of the goodness of fit index, exceeded the recommended level of 0.90, adjusted by the ratio of degrees of freedom for the proposed model to the degrees of freedom for the null model. The normed fit index (0.993) and the relative fit index (0.986) both had large values (near one), indicating a high level of goodness of fit.

2. Evaluation of direct and indirect relationships significance in the model. The correlation coefficients between latent variables with both the “behavioral intention” and variables of exogenous construct reached practical significance (≥0.30). The “subjective norm,” “perceived behavioral control,” and “past sexual behavior” were weakly correlated (<0.30). The “perceived behavioral control” was negatively correlated with all latent variables. The “belief in pleasant outcome of behavior” and “evaluation of pleasant outcome of behavior” had strong correlation (>0.50); the “belief in pleasant outcome of behavior” and the “perceived behavioral control” had the lowest correlation value in the matrix.

On the basis of the determinants of latent variables and indicators, we concluded that the practical significance of direct and indirect relationships logically imitated the theoretical model.

3. Calculating the standardized parameters of latent variables in structural model. The independent latent variables explained 34% of variability of dependent variable “behavioral intention.” The results are graphically presented as a full structural model with all indicators and relationships between them shown using covariant and asymptotic covariant matrix as input (due to ordinal variables and asymmetric distribution of data).

The relationship between variables in structural model was significant (P < 0.005). “Perceived behavioral control” was the strongest (negative) predictor of “behavioral intention” (r = -0.329). These results could be explained by the fact that participants with higher behavioral control had lower intention to engage in the first sexual intercourse then or during the next 12 months. “Subjective norm” was the second strongest predictor of “behavioral intention” (r = 0.183). This means that participants with higher agreement of important others with their decision to engage in the first sexual intercourse had higher behavioral intention to do it. Similarly, besides “subjective norm,” also the “past sexual behavior” had influence on “behavioral intention” (r = 0.174). An explanation may be that participants with more sexual experience had greater intention to engage in sexual intercourse. “Attitudes towards behavior” was the weakest predictor of “behavioral intention.” “Evaluation of pleasant outcome of behavior” (r = 0.100) and “belief in pleasant outcome of behavior” (r = 0.058) had slightly more predictive power. Participants who found the behavior more pleasurable and had good and pleasant attitude toward behavior outcome had greater intention to engage in the first sexual intercourse then or during the next 12 months. The unexplained variance was standard error as well. Like independent latent variable, it represents the powerful impact on “behavioral intention.” This means that there were additional predictors not included and measured in this study and this finding should receive considerable attention in future research (Figure 3).

Reliability

The reliability of the 17 predictive variables in the SEM model was satisfactory. Internal consistency was evaluated using Cronbach α coefficient, first calculated for each dimension of model: Cronbach α was 0.77 for attitudes toward the behavior; 0.79 for subjective norms; 0.74 for perceived behavioral control; and 0.73 for all 17 independent variables together. Corrected item-to-total correlations were greater than 0.50 and individual correlations in the inter-item-correlation matrix were greater than 0.30.
The questionnaire on the first sexual intercourse among young adolescents demonstrated good construct validity and enabled us to assess the determinants of behavior intention in the sexual life of youths. This was the first step in developing a reliable and valid measuring instrument, as the result of the analysis was a shorter version of the questionnaire, after exclusion of 17 out of 57 variables.

With 17 variables, we could assess the 5 dimensions of the behavioral intention predictors: past sexual behavior, belief in the pleasant outcome of behavior, evaluation of pleasant outcome of behavior, subjective norms, and perceived behavior control.

Sexual behavior represents important part of adolescents’ identity. The decision to engage in the first sexual intercourse is not only the consequence of hormonal changes but is also determined by social influence (1,2). Research on factors influencing decisions concerning sexual life allow us to develop evidence-based interventions with the aim to promote healthy behavior.

The aim of our study was test the construct validity of questionnaire for estimating the predictors of first sexual intercourse of youngsters. The instrument was constructed using Ajzen's theory of planned behavior in a previous study (10). The research was focused on the assessment of degrees of fit between our instrument and theoretical model. Our main point of interest was to construct the question-
naire which would separately assess behavior intention and behavior, and which would also assess the influence of important others on behavior intention. Our approach took into account that adolescents assessed consequences of their behavior on the basis of their own experiences and experiences of important others, especially parents. It is easier for youngsters to assess what is pleasant or good than what is probable or not probable (1, 2).

We are aware of all of the drawbacks of the Ajzen's TPB model (5, 20). One of the most frequently mentioned drawbacks is that the data are based on self-assessment, but there was no objective measure possible in this case (5, 20). Also, a meta-analysis found that the explained variance was 11% higher when behavior was self-assessed than when it was objectively measured (5, 20). In our research, the only way to get the data was self-assessment. We are aware that the behavior intention is only one of the steps before actual behavior, as it has been shown that TPB underestimates the influence of irrational and emotional factors on behavior. However, it was found that emotions influence the behavior intention in a more complex cognitive process so that TPB model remained as a pure cognitive construct (4, 15, 21-23). The actual decision to engage in the first sexual intercourse is realized under emotional (being in love, passion, fear) and social pressure (the influence of partner, the influence of peer group) and there is high probability of expected actual behavior (5, 20). Our instrument was designed to interview participants in an environment without previous emotional or social influences. The questionnaire was designed to selectively analyze the behavior intention to engage in the first sexual intercourse. The variables for instrument were collected from the literature and our previous research on these topics (10, 24), adding new data for the revision of the theoretical construct (5, 15, 21, 25).

Our data were submitted to structural equation analysis to determine the impact of variables on behavior intention. The SEM model was constructed from univariately statistically significant variables (attitudes to the behavior, subjective norms, and perceived behavior control) and past sexual behavior. In the SEM model, attitudes on the behavior were divided in 2 dimensions: beliefs of pleasant outcomes of the behavior and evaluation of the pleasant outcome of the behavior. In the final SEM model, there were 5 independent dimensions. The 18 variables were included in the model. All calculated items showed good fit between empirical and theoretical model and good reliability. Results excluded random error and its impact on measurements.

This model explained 34% of the variability of behavioral intention. On the basis of multivariate analysis, we conclude that the questionnaire can be used for assessing the impact of predictors (factors included in the questionnaire) of behavioral intention to engage in the first sexual intercourse in youth. The perceived behavioral control was a stronger predictor ($r = -0.33$) than the subjective norm ($r = 0.18$), and past sexual behavior ($r = 0.17$) was stronger predictor than attitudes toward behavior (belief in pleasant outcome of behavior $r = 0.06$; evaluation of pleasant outcome of behavior $r = 0.10$), with the weakest prediction on behavioral intention. Our findings are contrary to the results of a recent meta-analysis of perspective applications of TPB (5), in which the perceived behavioral control was the weakest predictor and attitude toward behavior was the strongest one. There are some explanations for this controversy. First, it is possible that the attitudes toward behavior in our study were not well determined and thorough enough. Second, the decision for engaging into the first sexual intercourse is influenced by strong emotions, and self-efficacy or self-control could play the main role and influence behavior directly. Third, unexplained variance ($r = 0.67$) or only 34% explained variability of behavioral intention indicates that some determinants were overlooked. In further research, we have to assess the predictors of sexual initiation that other researchers found relevant (refusal skills, self-standards, etc) (5-9, 16, 26-31). For this purpose, nominal or focus group technique could be used.

There are several limitations of our questionnaire and analysis. First, the theory of planned behavior is just one of the theories for explaining the predictors and intention. Second, the statistical procedure might be too complicated for a brief evaluation and might need a randomly selected sample of at least 400 participants. Third, we explained only 34% variability of behavioral intention and need to increase the strength of the model by adding new determinants, found important in other studies (5).

However, there is also a strong point of our approach. The instrument was validated (content validity) and was proven to have good psychometric characteristics. Our instrument is available free of charge for researchers and we strongly encourage other researchers to use it.

Health promotion should begin already in early childhood since attitudes toward healthy behavior are formed early in life. Instruments for measuring the influence of norms, attitudes, and beliefs are very important for determining health-promoting targets and for evaluation of
the effectiveness of intervention. The questionnaire on intention to engage in the first sexual intercourse in youth might be a useful tool for assessing the outcomes and effectiveness of health promotion measures.

References


