# COMPARISON OF INTERNET-BASED VERSUS PAPER-AND-PENCIL ADMINISTERED ASSESSMENT OF POSITIVE DEVELOPMENT INDICATORS IN ADOLESCENTS' SAMPLE*

**Rimantas Vosylis**

Mykolas Romeris University
Faculty of Social Policy
Department of Psychology
Ateities Str. 20, LT-08303 Vilnius
Tel. (85) 2714620
E-mail: rvosylis@yahoo.co.uk

**Oksana Malinauskienė**

Mykolas Romeris University
Faculty of Social Policy
Department of Psychology
Ateities Str. 20, LT-08303 Vilnius
Tel. (85) 2714620
E-mail: oksana@mruni.eu

**Prof. Dr. Rita Žukauskienė**

Mykolas Romeris University
Faculty of Social Policy
Department of Psychology
Ateities Str. 20, LT-08303 Vilnius
Tel. (85) 2714620
E-mail: rzukausk@mruni.eu

*The aim of this study was to evaluate the use of the online data collection method to survey adolescents about their psychological characteristics in a follow up-longitudinal study on positive youth development in order to test the psychometric equivalence of two assessment methods. 1030 participants (17–19 years old) completed paper-and-pencil questionnaires in schools (505 boys and 525 girls), 132 (28 boys and 104 girls) completed Internet-based questionnaires, and 47 (15 boys 32 girls) completed both, measuring positive development indicators.*

*The findings suggest that adolescents report less socially desirable behaviour and active citizenship in Internet-based questionnaires, but generally Internet-based administration does not have any differences in the means values of positive development indicators as compared to paper-and-pencil administration. Internet-based questionnaires have higher or similar internal consistencies as compared with paper-and-pencil questionnaires and are highly correlated with each other when administered using Internet-based and paper-and-pencil assessment. There is no interaction effect of the Internet versus paper-and-pencil assessment and the sex of adolescents on the positive development indicators. Limitations of this study are discussed.*

***Key words**: Internet-based assessment, paper-and-pencil self-administered assessment**,** positive development indicators, adolescents.*

* The authors would like to thank M. Richardson, BASIC, Crawley, England for reading and improving the English of this article.

## Introduction

Students and researchers have become increasingly comfortable with the Internet, and many of them are interested in using the Internet-based questionnaires to collect data. The use of the Internet reduces many of the costs associated with collecting data on human behaviour. However, with advantages of using the Internet for data collection, there are challenges that should be addressed. This paper discusses the advantages and limitations of online data collection as an alternative to paper-and-pencil assessment, with particular reference to the conduct of a longitudinal study on the positive development indicators, involving upper secondary school students in Lithuania.

Collecting research data through traditional paper-and-pencil methods can be costly and time-consuming. This becomes extremely difficult in longitudinal studies focused on transitions from adolescence to early adulthood for follow-up, as participants move both from school or their parents' house and to other cities in the same country or abroad. Conducting Internet-based surveys is an alternative that appears to have the potential, and indeed is already used world-wide (Yun and Trumbo, 2000) to collect large amounts of data efficiently and economically within relatively short time frames.

The advantages of web-based research techniques have been extensively documented. Many researchers support their cost-effectiveness, flexibility and control over format, large samples, lower cost, efficiency of data management, rapid access to participants, increased participation, and ability to follow-up with participants, and popularity among certain populations such as adolescents and young adults (Van Selm and Jankowski, 2006). Internet surveys are more accurate than paper-and-pencil surveys, and data collection and processing is automatic and faster (Wright et al., 1998; Barbeite and Weis, 2004), guarantees a rather short time frame for the collection of responses and are time-and cost-saving (Mertler, 2003), protects against the loss of data and makes transferring data into a database for analysis simpler (Ilieva et al., 2002). The quality of the data is improved as people can be reminded to go back to an item that was missed, and manual data entry from a paper-based survey is not necessary (Barbeite and Weis, 2004).

However, there are some concerns regarding data quality in web-based surveys. As few of potential disadvantages of web-based data collection researchers include the time or costs of initial development, technical difficulties experienced by users, data integrity and data security (Ahern, 2005; Jones et al., 2008; Van Selm and Jankowski, 2006). Another disadvantage is the experimenter's inability to control the environmental conditions in which the Internet participants' complete a survey. For example, it is difficult for researchers to control the order in which participants complete online surveys (Nosek et al., 2002).

There continues to be some uncertainty about the reliability and validity of the data collected on the Internet because of the sampling biases (Kraut et al., 2004), participant dropout and attrition (O'Neill and Penrod, 2001), and incomplete data. Some research reports response rates to be generally lower for online surveys than for mail or telephone surveys (Kraut et al., 2004). Some

studies (Bates and Cox, 2008) indicate that more incomplete questionnaires have been found in the Internet conditions, whereas in other studies paper-and-pencil yielded more missing data (e.g., Denscombe, 2006), whilst yet other studies found no difference between these two conditions (Wu and Newfield, 2007). Thus, collecting data via the Internet has its own set of challenges that make it different from more traditional methods of data collection; but overall, the disadvantages of online data collection are found to be much lower.

As the method of data collection can affect the answers that are obtained, it is important to determine whether responses to web-based questionnaires are comparable to those obtained via self-assessment in the classroom. To date, there is no conclusive evidence to indicate a difference in responses between paper-and-pencil surveys and online surveys (Ilieva et al., 2002). Several studies did not find major differences between data gathered via Internet-based and paper-and-pencil questionnaires. For example, T. Joubert and H. J. Kriek (2009) conducted two studies in which scores obtained online were compared with scores obtained by paper-and-pencil methods. In their study, the psychometric properties of the paper-and-pencil and Internet-based applications were very similar. S. Hays and R. S. McCallum (2005) administered a computer-administered version and a paper-and-pencil version and found that relative rankings were similar across administration formats of the Minnesota Multiphasic Personality Inventory–Adolescent version. Some researchers found no difference in adolescent reports of sensitive information given online and in paper-and-pencil version. For example, no significant differences in the perceived level of privacy and confidentiality between web-based and paper-and-pencil questionnaires were found, and this did not differ by gender in the study by P. M. Van De Looij-Jansen and E. J. De Wilde (2008).

Other studies report widely divergent inconsistency rates when two assessment formats are compared. Findings from some studies show that adolescents disclose some sensitive information in computerized questionnaires more often than in paper-and-pencil conditions. For example, in the study by N. D. Brener et al. (2006), students who completed questionnaires on the computer were more likely to report the prevalence of risk behaviours compared to students who completed paper-and-pencil questionnaires. Significant, but small, differences between the two modes of data collection were found for the Strengths and Difficulties Questionnaire (SDQ) subscales "emotional symptoms" (paper-and-pencil > web-based) and "pro-social behavior" (paper-and-pencil > web-based), and carrying a weapon (web > paper-and-pencil) (Van De Looij-Jansen, and De Wilde, 2008), but for other sensitive topics like the use of alcohol or marijuana, vandalism, and stealing no differences were found (Van De Looij-Jansen and De Wilde, 2008).

A fundamental assumption of an Internet research is that the results obtained are comparable to in-person (off-line) research (Meyerson and Tryon, 2003). Thus, before using an electronic version, it is necessary to ensure that the psychometric characteristics are identical to those of the traditional test form. L. M. Honaker (1988) has stated that "psychometrically, two forms of a test are considered to be equivalent if it has been

demonstrated that the two forms are parallel" (p. 562). T. Buchanan, J. A. Johnson and L. Goldberg (2005) argue that the characteristics of the testing medium or the samples used (often differing from those used in the development and validation of the off-line measure) may impact on a measure's psychometric properties and ultimately its power to reliably and validly measure the construct(s) of interest. According to D. Bartram (1994), for the electronic version to be equivalent to traditional, both forms must have equal reliabilities, intercorrelations at the level expected from their reliability, have comparable correlations with other variables as well as equal means and standard deviations. Also, the factor structure of the two forms of an instrument should be identical in order to be two forms considered as equivalent.

Some studies have demonstrated that on-line versions of tests are equivalent to traditional paper-and-pencil versions of the same instruments. J. M. Stanton (1998) reported a similar factor structure for an organizational justice scale when the Internet and in-person data were compared. In P. Meyerson and W. W. Tryon (2003) study, an on-line version of a sexual boredom scale correlated with other scales that mirrored those of an original off-line version and had almost identical reliability coefficients. Researchers concluded that the two versions of the tests were essentially parallel and, thus, psychometrically equivalent. They have concluded that data collection using web-based questionnaires is reliable, valid, reasonably representative, cost-effective, and efficient (Meyerson and Tryon, 2003). T. Buchanan and J. L. Smith (1999) reported comparable Cronbach alpha coefficient and confirmatory factor structures for the Internet and in-person administration of the Self-Monitoring Scale–Revised. R. N. Davis (1999) found a slightly lower internal consistency in a web-based version of the Ruminative Responses Scale than in the paper-and-pencil version of the scale. In K. A. Pasveer and J. H. Ellard (1998) study, data collected electronically online in two samples were compared with traditional paper-and-pencil measure data from two university samples in three psychometric studies of a new measure of self-trust, the Self-Trust Questionnaire (STQ). Measures of internal consistency for the STQ were very similar for online and student samples, although they were slightly higher for the web-based version (0.86 and 0.88 vs. 0.84 and 0.86, respectively). The factor structure of the STQ was also very similar in factor analyses of the scale in each sample. Their findings indicate that the advantages of the online as a data source, including large heterogeneous samples, outweigh problems with data accuracy and generalizability, making the online an attractive source of data for researchers developing self-report personality inventories. Furthermore, J. H. Krantz and R. Dalal (2000) state that off-line and on-line research data from the same study can "essentially replace each other" (p. 56). While there is evidence that online tests can be reliable and valid, there is also evidence that psychometric properties may change subtly when a test is placed on the internet. Differences are found mainly in factor structures of questionnaires which measure more than one construct e.g. T. Buchanan et al. (2005), in an evaluation of a web-based version of a Five-factor personality inventory, found that a small number

of the items loaded on the different factors to those they had loaded on in the offline development sample.

Thus, as some discrepancies in findings still exist, the further exploration of psychometric equivalence of the two assessment methods is an important part of empirical research by documenting mean differences, also as differences in variation and conducting a multivariate comparison of the two correlation matrices. This study seeks to explore the comparability of paper-and-pencil versus online Internet versions of the important and widely used instruments (such as Subjective Well-being, Self-efficacy, School Burnout and some others) to assess psychological adjustment and functioning.

The aim of this study was to evaluate the use of the online data collection method to survey adolescents about their psychological characteristics in a follow-up longitudinal study on positive youth development in order to test the psychometric equivalence of the two assessment methods.

## Method

### *Study Design and Procedure*

The data used for this particular study are from an ongoing longitudinal Positive Youth Development (PYD) study that examines the mechanisms and processes through which young people develop their competences from adolescence to young adulthood.

The first data collection took place in the spring of 2008 and included four cohorts of students aged 15–19, followed by the second assessment in 2009 and the third in 2010. Student participants were drawn from eight high schools in the administrative region of Klaipėda, Lithuania. For this particular study, data from the third assessment, which took place in 2010 when children still enrolled in the schooling system (two youngest cohorts) were asked to complete paper-and-pencil questionnaires at school or to choose an online-based questionnaire to fill in at home, were used. E-mail messages were sent in advance to participants, offering to choose the mode of assessment, e.g., the pen-and-pencil form or the on-line questionnaire. Paper-and-pencil assessment was conducted in schools by the researchers and several trained research assistants upon obtaining the consent of school authorities and parents. Participants who choose to fill the online form of questionnaire were provided with passwords in order to access the online form; also, they were asked to give personal details, such as name and family name, e-mail address and other details which were also asked from the participants that completed the paper-and-pencil version. The online questionnaire was based at www.manoapklausa.lt, which provides free-of-charge service for conducting Internet-based surveys. Three weeks later, in order to access children who had been absent from school during the data collection or living in other cities or abroad, another invitation to participate in the study was sent to the whole sample via e-mail. In addition to those who had completed the questionnaire for the first time, because they had been absent during the initial data collection, there were 50 subjects that completed the questionnaire using paper and pencil in school and then completed the online version. This provided a possibility to analyze the differences of positive development indicators between the two forms of administration in the two independent samples (those who filled

either paper-and-pencil or Internet-based questionnaire) and two dependent samples (those who completed both forms of questionnaires within 4 to 6 weeks between the measurements).

### *Subjects*

In this particular study, all cases that had any missing data were excluded. Overall, 1030 participants (17–19 years old) completed paper-and-pencil questionnaires in schools (505 boys and 525 girls), 132 (28 boys and 104 girls) completed the Internet-based questionnaire (independent samples), and 47 (15 boys 32 girls) completed both (dependent samples). Differences between the two independent samples were evaluated for the following demographic variables: parent / caregiver currently living with, parent / caregiver employment status, age, and city currently living in, using the Chi-square test. Significant differences were found for age and city currently living in. There were younger participants and living in three biggest cities in Lithuania among those that filled the questionnaire using the online form versus the paper-and-pencil sample ($p < 0.001$). The distribution in the parent / caregiver employment status and parent / caregiver currently lived with did not differ between the two groups ($p > 0.05$).

### *Measures*

The two versions of the questionnaire (paper-and-pencil and Internet-based) were identical in terms of the questions asked, their wording, and the order of presentation in the survey.

*Life satisfaction*. The Satisfaction with Life Scale (SWLS) is a measure of life satisfaction, developed by E. Diener and colleagues (Diener et al., 1995).

*Social well-being*. Social well-being was measured using the short-form version of the scale that consists of five items (Keyes, 2005).

*Self-efficacy*. The General Self-efficacy Scale (GSE) (Schwarzer and Jerusalem, 1995). The GSE is a 10-item scale designed to assess optimistic self-beliefs used to cope with a variety of demands in life.

*Pro-social tendencies*. Pro-social Tendencies Measure (PTM, Carlo and Randall, 2002). The 23-item version of the PTM was composed of 6 sub-scales: public (4 items), anonymous (5 items), dire (3 items), emotional (4 items), compliant (2 items), and altruism (5 items).

*Closeness to others*. Other questions were developed for the Positive Youth Development (PYD) study. Similar questions are in the C. L. Keyes (2006) study, and consist of eight items, e.g. "How many people there are (a) you feel close to?"… to (mother; father; brother and (or) sister; classmates, etc.). Responses were given on a 4-point scale ranging from (1) "especially close" to (6) "not very close".

*School burnout*. School Burnout Inventory (SBI, Salmela-Aro et al., 2009). The inventory consists of 10 items measuring three factors of school burnout: (a) exhaustion at school (4 items), (b) cynicism toward the meaning of school (3 items), and (c) sense of inadequacy at school (3 items).

*Voluntary work*. Questions were developed for the Positive Youth Development (PYD) study. Questions' measuring which voluntary work is popular among young people (e.g., *How often do you do these activities? Helping elderly people?)*. Responses

to the question were: (1) never, (2) approximately once a year, (3) approximately onie a month, and (4) more than once a month.

*Community and Neighborhood*. Community and Neighborhood Measure (Tolan et al., 2001) assesses the degree to which youths perceive problems in their neighborhood, evaluation of relationships with neighbours. The scale consists of 5 items.

*Active citizenship*. Questions developed for the Positive Youth Development (PYD) study, which evaluate the activities that could interest youth as citizens, e.g. *when you will be grown-up, are you going to join an environment protection organization?* Responses to the question were: (1) definitely not, (2) unlikely, (3) might be, (4) likely, (5) definitely yes.

*Socially Desirable Behaviour*. Questions were developed for the Positive Youth Development (PYD) study and consist of questions measuring the perceived civic behaviour in the future, e.g. help to policemen or policewomen to keep public order. Responses to the question were: (1) definitely not, (2) unlikely, (3) might be, (4) likely, (5) definitely yes.

### *Data analysis*

In order to examine the differences in positive development indicators between two modes of administration, several analyses were carried out. For the two independent samples (1027 who answered the questionnaire using paper and pencil and 132 who used the online questionnaire), the mean differences using Student's t criteria for independent samples the, equality of variances using Levens' test, and the equality of Cronbach α coefficients using the Feldt test were evaluated. For the two dependent samples (47 who answered the paper-and-pencil form of the questionnaire and then the online questionnaire), the mean differences using Student's t criteria for dependent samples, the intra-class correlation coefficient (ICC) (3.1) between the two measurements was evaluated. ICCs instead of Pearson's correlation were used, because Pearson's r shows to what extent two repeated measures fit on a straight line, and does not evaluate the possible systematic differences (e.g., increase in score means re-test), while ICC assesses whether the measures on the subject are identical and have no systematic differences (e.g., Brouwer et al., 2004). The two-factor analysis of variance for a mixed design to evaluate the interaction between the mode of administration (conditions: online / paper-and-pencil) and sex (conditions: boy/girl) was also utilized.

## Results

Results of the analysis of two independent samples are presented in Table 1. Significant differences in the mean values of 4 out of 17 scales were found. Subjects that completed the online questionnaire scored higher on satisfaction with life and on a scale measuring *public* pro-social tendency. Subjects that completed the questionnaire using paper and pencil scored higher on *active citizenship* and *socially desirable behaviour*. While testing the equality of variances in the two samples, three scales were found to have different variances. Score means for the *dire* and *altruism* of pro-social tendencies scales and the score mean for the *self-efficacy* scale were slightly higher among those who completed the questionnaire using paper and pencil. Four scales out of 17 differed in their

*Table 1.* ***Mean, variance and internal consistency differences in questionnaires administered using paper-and-pencil and Internet in two independent adolescent samples***

| Measures | | Mean (st. dev.) | | Independent sample test for equality of mean values | | Leven's test for equality of variances | | Cronbach α | | The Feldt test for equality of two Cronbach α coefficients | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Paper-and-pencil (N=1027) | **Online** (N=132) | **T** | ***p* (T)** | **F** | ***p* (F)** | Paper-and-pencil (N=1027) | **Online** (N=132) | **W** | ***p* (W)** |
| *Satisfaction with life* | | **4.32 (1.15)** | **4.65 (1.12)** | **–3.134** | **0.002** | 0.030 | 0.862 | 0.826 | 0.793 | 0.841 | 0.083 |
| *Scales measuring pro-social tendencies:* | *compliant* | 3.75 (0.83) | 3.70 (0.70) | 0.752 | 0.452 | 0.867 | 0.352 | 0.796 | 0.868 | 0.647 | 0.001 |
| | *public* | **2.18 (0.75)** | **2.36 (0.63)** | **–3.019** | **0.003** | 2.578 | 0.109 | 0.737 | 0.751 | 0.947 | 0.352 |
| | *anonymous* | 2.88 (0.69) | 2.85 (0.68) | 0.474 | 0.636 | 0.095 | 0.757 | 0.707 | 0.759 | 0.823 | 0.078 |
| | *dire* | 3.00 ***(0.77)*** | 2.98 ***(0.65)*** | 0.226 | 0.822 | 5.885 | 0.015 | 0.572 | 0.644 | 0.832 | 0.091 |
| | *emotional* | 3.31 (0.75) | 3.19 (0.69) | 1.783 | 0.075 | 1.658 | 0.198 | 0.507 | 0.709 | 0.590 | < 0.001 |
| | altruism | 3.12 **(0.60)** | 3.10 **(0.51)** | 0.354 | 0.724 | **7.139** | **0.008** | *0.413* | *0.528* | 0.804 | 0.057 |
| *Self-efficacy scale* | | 2.91 **(0.52)** | 2.93 **(0.39)** | –0.341 | 0.734 | **7.610** | **0.006** | 0.837 | 0.904 | **0.589** | **< 0.001** |
| *Closeness to others* | | 4.03 (1.71) | 4.20 (1.75) | –0.956 | 0.339 | 0.288 | 0.592 | 0.631 | 0.579 | 0.876 | 0.146 |
| *School burnout:* | *exhaustion* | 4.11 (1.15) | 3.94 (0.97) | 1.944 | 0.052 | 3.248 | 0.072 | 0.718 | 0.691 | 0.913 | 0,230 |
| | *cynicism* | 3.31 (1.17) | 3.42 (1.13) | –0.987 | 0.324 | 0.369 | 0.544 | 0.743 | 0.732 | 0.960 | 0.361 |
| | inadequacy | 3.56 (1.26) | 3.56 (1.09) | –0.056 | 0.955 | 3.352 | 0.067 | *0.484* | *0.561* | 0.851 | 0.121 |
| *Voluntary work* | | 1.85 (0.55) | 1.83 (0.50) | 0.325 | 0.745 | 1.363 | 0.243 | 0.567 | 0.571 | 0.991 | 0.485 |
| *Community involvement* | | 1.61 (0.85) | 1.62 (0.81) | –0.126 | 0.899 | 0.148 | 0.700 | ***0.699*** | ***0.627*** | ***0.810*** | ***0.043*** |
| *Social well-being* | | 2.29 (0.90) | 2.31 (0.94) | –0.256 | 0.798 | 1.460 | 0.227 | 0.694 | 0.653 | 0.882 | 0.157 |
| *Active citizenship* | | **2.39 (0.80)** | **2.12 (0.75)** | **3.953** | **<0.001** | 0.035 | 0.852 | 0.818 | 0.832 | 0.923 | 0.285 |
| *Socially desirable behaviour* | | 2.46 (0.79) | 2.15 (0.76) | **4.387** | **<0.001** | 0.020 | 0.887 | 0.802 | 0.763 | 0.835 | 0.076 |

*Table 2.* ***Mean differences and intra-class correlations between positive development measures administered using paper-and-pencil and Internet and paper-and-pencil-online and sex interaction effects on these measures***

| Measures | | Mean (st. dev.) (N=49) | | Dependent sample t test for equality of mean values | | Intra-class correlation coefficient | | **Two-factor analysis of variance** (paper-and-pencil -online*sex interaction effect) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **T1: paper-and-pencil** | **T2: online** | **T** | ***p* (T)** | **ICC (3.1)** | ***p* (ICC)** | **F** | ***p* (F)** | Partial eta squared | Observed power |
| *Satisfaction with life* | | ***4.91 (0.99)*** | ***4.58 (1.22)*** | ***2.396*** | ***0.021*** | **0.783** | **< 0.001** | 0.391 | 0.535 | 0.009 | 0.094 |
| *Scales measuring pro-social tendencies:* | *compliant* | ***3.75 (0.65)*** | ***3.96 (0.62)*** | ***–2.165*** | ***0.036*** | **0.612** | **0.001** | 0.020 | 0.888 | < 0.001 | 0.052 |
| | *public* | 2.26 (0.63) | 2.34 (0.24) | –0.946 | 0.349 | **0.669** | **< 0.001** | 1.572 | 0.216 | 0.034 | 0.233 |
| | *anonymous* | 2.91 (0.65) | 2.94 (0.62) | –0.368 | 0.714 | **0.783** | **< 0.001** | 0.109 | 0.743 | 0.002 | 0.062 |
| | *dire* | 3.09 (0.61) | 3.18 (0.64) | –1.159 | 0.252 | **0.808** | **< 0.001** | 1.657 | 0.205 | 0.036 | 0.243 |
| | *emotional* | 3.22 (0.56) | 3.31 (0.58) | –1.493 | 0.142 | **0.849** | **< 0.001** | 3.427 | 0.071 | 0.071 | 0.441 |
| | *altruism* | 3.11 (0.40) | 3.11 (0.45) | –0.076 | 0.940 | **0.535** | **0.005** | 0.191 | 0.664 | 0.004 | 0.071 |
| *Self-efficacy scale* | | 2.95 (0.30) | 2.89 (0.38) | 1.586 | 0.120 | **0.821** | **< 0.001** | 0.041 | 0.840 | 0.001 | 0.055 |
| *Closeness to others* | | 3.86 (1.40) | 4.22 (1.90) | –1.430 | 0.161 | **0.740** | **< 0.001** | 0.291 | 0.593 | 0.008 | 0.082 |
| *School burnout:* | *exhaustion* | 3.86 (1.13) | 4.04 (0.85) | –1.206 | 0.234 | **0.642** | **< 0.001** | 0.056 | 0.814 | 0.001 | 0.056 |
| | *cynicism* | 3.30 (1.30) | 3.51 (1.06) | –1.217 | 0.230 | **0.654** | **< 0.001** | 2.664 | 0.110 | 0.056 | 0.359 |
| | *inadequacy* | 3.54 (1.12) | 3.56 (1.20) | –0.140 | 0.889 | **0.746** | **< 0.001** | 3.037 | 0.088 | 0.063 | 0.400 |
| *Voluntary work* | | 1.87 (0.41) | 1.85 (0.45) | 0.297 | 0.767 | **0.719** | **< 0.001** | 0.470 | 0.496 | 0.010 | 0.103 |
| *Community involvement* | | 1.65 (0.76) | 1.60 (0.76) | 0.454 | 0.652 | **0.647** | **< 0.001** | 2.927 | 0.094 | 0.061 | 0.388 |
| *Social well-being* | | 2.37 (0.96) | 2.37 (0.85) | 0.000 | 1.000 | **0.760** | **< 0.001** | 0.096 | 0.758 | 0.002 | 0.061 |
| *Active citizenship* | | 2.25 (0.69) | 2.22 (0.74) | 0.391 | 0.698 | **0.863** | **< 0.001** | 0.136 | 0.714 | 0.003 | 0.065 |
| *Socially desirable behaviour* | | 2.28 (0.82) | 2.39 (0.80) | –1.585 | 0.120 | **0.906** | **< 0.001** | 0.199 | 0.657 | 0.004 | 0.072 |

internal consistency between the two samples. Scales for *compliant* and *emotional* pro-social tendencies and the scale for *self-efficacy* had higher internal consistencies in the sample that completed the questionnaire online (at $p < 0.001$) and Cronbach $\alpha$ for scale of *community involvement* was higher in a sample that filled questionnaire using paper and pencil (at $p < 0.5$).

Results of the analysis of two dependent samples are presented in Table 2. Significant differences in the mean values were found in two scales. Subjects scored higher on the satisfaction with life when using paper and pencil to complete the questionnaire, and significantly lower on *complaint* scale measuring pro-social tendency. Intra-class correlations (ICC 3.1) indicated that all scales were significantly correlated to the same measurements on the second administration at $p < 0.001$, apart one scale (*altruism*) which was significantly correlated at $p < 0.005$. The correlation between sex and the form of administration was not significant. Effect size did not exceed 0.071, in most cases being < 0.05, and the observed power did not exceed 0.441, in most cases being < 0.1.

## Discussion

In this study, the possible use of Internet-based tools in a psychological research was explored by comparing two modes of administration – paper and pencil and web-based questionnaires – measuring indicators of positive development. No differences for most indicators of positive development were found; this is in agreement with the previous research reporting only a few differences between assessments using online versus paper-and-pencil types of questionnaires (Horswill and Coster, 2001; Valejo et al., 2008; Cronk and West, 2002). When evaluating the differences between the two forms of administration in independent samples, differences in the mean values and variances in several scales were found; however, most of differences could be attributed to differences in the samples rather than to the condition of administration. Participants that choose to complete the questionnaire online, scored higher on satisfaction with life, and the decision to use the Internet could be related to the fact that they had personal computers at home and thus to a higher quality of life. Moreover, a difference in mean values while measuring satisfaction with life in two dependent samples was found, which was opposite to the independent sample result: participants scored higher on satisfaction with life while answering the questionnaire using the paper-and-pencil form in comparison with the Internet-based questionnaire. No reasonable explanation as to why participants scored higher on satisfaction with life and lower on pro-social tendency of *compliant* while answering the questionnaire using the paper-and-pencil form was deduced, but it is worth mentioning that the difference was statistically significant only at $p < 0.5$. With the *compliant* scale, it might be the case that this form of administration had an impact on mean differences, since we found that the Cronbach α coefficient was higher in the independent online sample. Another possible explanation lies in the retroactive history events specific to the sample that was used for the repeated analysis. These 49 participants were aware of the fact that they had already filled this questionnaire at school, but still chose to spend an hour at home by the computer and answer the

questions for the second time. This decision might be related to the fact that something had happened to them during the period between the two measurements that resulted in a decreased satisfaction with life and an increase in compliance (e.g., they were "home arrested" by their parents) and influenced their decision to fill the questionnaire for the second time.

The other two scales that showed a significant difference in the mean values between two independent samples were active citizenship and socially desirable behaviour: both scales had higher mean values in a sample that completed the questionnaire using the paper-and-pencil form. These differences are more likely to result from the differences in the form of administration, since several studies had reported rather similar results, e.g., P. M. Van De Looij-Jansen and E. J. De Wilde have found that adolescents tend to disclose sensitive information about themselves, like carrying a weapon, while using the Internet rather than paper-and-pencil questionnaires (Van De Looij-Jansen and De Wilde, 2008), although strict conclusions cannot be made since the same difference between modes the of administration were not found in dependent samples.

Three scales differed in their variance between independent samples which differed in the form of questionnaire administration: (the *dire* and *altruism* scales, measuring pro-social tendencies, and the self-efficacy scale). Since in all three cases the variance was higher in the sample that completed the questionnaire using paper and pencil, and the sample that had filled the questionnaire using paper-and-pencil was considerably larger, these differences are more likely to be a result of differences between the samples rather than the form of administration.

Four out of 17 scales differed in their internal consistency measured by the Cronbach α coefficient. The *compliant* and *emotional* scales, measuring pro-social tendencies, and the *self-efficacy* scale had a higher internal consistency in the sample that completed questionnaire using the Internet, and the scale for community involvement had a higher internal consistency while measured using the paper-and-pencil form, although only at $p < 0.05$. It is also worth mentioning that even though the scale for pro-social tendency of *altruism* and the scale for *inadequacy* (one of the school burnout indicators) did not have significantly different Cronbach α between the two samples, their internal consistency was below the acceptable level (0.413 and 0.484, respectively) in the sample that filled the questionnaire using the paper-and-pencil form and would have been a subject for reviewing and modification (George and Mallery, 2003), while the Cronbach α levels (0.528 and 0.561, respectively) in the sample that completed the questionnaire online, would have been considered acceptable for large-scale studies. To summarize, the results suggest that the form of administration can have an impact on the internal consistency of the instruments measuring the positive development indicators, but the impact of the Internet-based administration is rather positive than negative in terms of using the Internet to conduct research studies in adolescent samples.

All indicators of positive development, measured using the paper-and-pencil form, positively correlated with the same indica-

tors at the second administration via the Internet with ICCs ranging from 0.53 to 0.91. No effect of the form of administration and the sex of participants was found, which indicates that the results obtained by measuring positive development indicators using web-based questionnaires may be compared with the ones obtained by the traditional paper-and-pencil method.

To sum up, findings of this study suggest that Internet-based assessment of positive development indicators can be used as a comparable alternative to paper-and-pencil assessment, although when using two forms to measure indicators such as socially desirable behaviour and active citizenship, some precautions should be taken into account, since adolescents report less socially desirable behaviour and active citizenship in Internet-based questionnaires.

## Limitations

Despite the findings suggesting that Internet-based assessment of positive development indicators can be a suitable alternative to employing the traditional paper-and-pencil form, there were some important limitations of this study, which need to be addressed. The first one is related to the samples we used for this analysis. The two groups we used to test for mean differences and other parameters were not equal in size and somewhat biased. The second limitation was related to the differences in group size in dependent and independent samples. Since in both cases the sample size was very different, very firm conclusions concerning the Internet-based administration effect on the positive development indicators cannot be drawn. It is also important that our findings are limited to adolescence, school context and positive development indicators: the effects of Internet-based administration can be very different in other samples, e.g., adult or senior, and on different measures they can be different as well.

## Conclusions

1. The Internet-based administration of the questionnaire decreased adolescents' reactivity to the participation in this study while reporting less socially desirable behaviour and active citizenship, but in general the mean values of positive development indicators did not differ from the mean values obtained from paper-and-pencil form of questionnaires.
2. Internet-based questionnaires on positive development had a higher or similar internal consistencies as compared with paper-and-pencil questionnaires.
3. There is no interaction effect of Internet versus paper-and-pencil assessment and the sex of adolescent on positive development indicators.
4. Measures of positive development were highly correlated in Internet-based and paper-and-pencil assessments.

**REFERENCES**


Ahern N. R. Using the Internet to conduct research // Nurse Researcher. 2005, vol. 13 (2), p. 55–70.

Barbeite F. G., Weiss E. M. Computer self-efficacy and anxiety scales for an Internet sample: Testing measurement equivalence of existing measures and development of new scales // Computers in Human Behavior. 2004, vol. 20 (1), p. 1–15.

Bartram D. Computer-based assessment // International Review of Industrial and Organizational Psychology / Ed. by C. L. Cooper, I. T. Robertson. Chicester: John Wiley, 1994, 9. P. 31–69.

Bates S. C., Cox J. M. The impact of computer versus paper-pencil survey, and individual versus group administration, on self-reports of sensitive behaviors // Computers in Human Behavior. 2008, vol. 24, p. 903–916.

Brener N. D., Eaton D., Kann L., Grunbaum J. A., Gross L. A., Kyle T. M., Rodd J. G. The association of survery setting and mode with self-reported health risk behaviors among high school students // Public Opinion Quarterly. 2006, vol. 70, p. 354–374.

Brouwer S., Kuijer W., Dijkstra P. U., Göeken L. N. H., Groothoff J. W., Geertzen J. H. B. Reliability and stability of the Roland Morris Disability Questionnaire: Intra class correlation and limits of agreement // Disability and Rehabilitation. 2004, vol. 26 (3), p. 162–165.

Buchanan T., Johnson J. A., Goldberg L. R. Implementing a Five-Factor Personality Inventory for use on the Internet // European Journal of Psychological Assessment. 2005, vol. 21, p. 115–127.

Buchanan T., Smith J. L. Using the Internet for psychological research: Personality testing on the World Wide Web // British Journal of Psychology. 1999, vol. 90, p. 125–144.

Carlo G., Randall B. A. The development of a measure of prosocial behaviors for late adolescents // Journal of Youth and Adolescence. 2002, vol. 31 (1), p. 31–44.

Cronk B. C., West J. L. Personality research on the Internet: A comparison of Web-based and traditional instruments in take-home and in-class settings // Behavior Research Methods, Instruments, & Computers. 2002, vol. 34 (2), vol. 177–180.

Davis R. N. Web-based administration of a personality questionnaire: Comparison with traditional methods // Behavior Research Methods, Instruments, & Computers. 1999, vol. 31, p. 572–577.

Denscombe M. Web-based questionnaires: An assessment of the mode effect on the validity of data // Social Science Computer Review. 2006, vol. 24 (2), p. 246–254.

Diener E., Emmons R. A., Larsen R. J., Griffin S. The satisfaction with Life Scale // Journal of Personality Assessment. 1995, vol. 49, p. 71–75.

George D., Mallery P. SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn & Bacon, 2003.

Hays S., McCallum R. S. A comparison of the pencil-and-paper and computer-administered Minnesota Multiphasic Personality Inventory-Adolescent // Psychology in the Schools. 2005, vol. 42 (6), p. 605–13.

Honaker L. M. The equivalency of computerized and conventional MMPI administration: A critical review // Clinical Psychology Review. 1988, vol. 8, p. 561–577.

Horswill M. S., Coster M. E. User-controlled photographic animations, photograph-based questions, and questionnaires: three Internet-based instruments for measuring drivers' risk-taking behavior // Behavior Research Methods, Instruments, & Computers. 2001, vol. 33 (1), p. 46–58.

Ilieva J., Baron S., Healey N. M. Online surveys in marketing research: Pros and cons // International Journal of Market Research. 2002, vol. 44, p. 361–376.

Jones S., Murphy F., Edwards M., James J. Doing things differently: Advantages and disadvantages of Web questionnaires // Nurse Researcher. 2008, vol. 15 (4), p. 15–26.

Joubert T., Kriek H. J. Psychometric comparison of paper and-pencil and Online personality assessments in selection setting // SA Journal of Industrial Psychology / SA Tydskrif vir Bedryfsielkunde. 2009, vol. 35 (1), p. 78–88.

Keyes C. L. M. The subjective well-being of America's youth. Toward a comprehensive assessment // Adolescent and Family Health. 2005, vol. 4, p. 3–11.

Keyes C. L. Mental health in adolescents: Is America's youth flourishing? // American Journal of Orthopsychiatry. 2006, vol. 76 (3), p. 695–402.

Krantz J. H., Dalal R. Validity of Web-Based psychological research // Psychological Experiments on the Internet / Ed. by M. H. Birnbaum. San Diego: Academic Press, 2000. P. 35–60.

Kraut R., Olson J., Banaji M. R., Bruckman A., Cohen J., Couper M. Psychological research online: Opportunities and challenges // American Psychologist. 2004, vol. 59, p. 105–117.

Mertler C. A. Patterns of response and nonresponse from teachers to traditional and Web surveys // Practical Assessment, Research & Evaluation. 2003, vol. 8 (22) (retrieved http://pareonline.net/getvn.asp?v=8&n=22, 2012-01-14).

Meyerson P., Tryon W. W. Validating Internet research: A test of the psychometric equivalence of Internet and in-person samples // Behavior Research Methods Instruments & Computers. 2003, vol. 35, p. 614–620.

Nosek B. A., Banaji M., Greenwald A. G. Harvesting implicit group attitudes and beliefs from a demonstration Web site // Group Dynamics. 2002, vol. 6, p. 101–115.

O'Neill K. M., Penrod S. D. Methodological variables in Web-based research that may affect results: Sample type, monetary incentives, and personal informatikon // Behavior Research Methods, Instruments, & Computers. 2001, vol. 33 (2), p. 226–233.

Pasveer K. A., Ellard J. H. The making of a personality inventory: Help from the WWW // Behavior Research Methods, Instruments, & Computers. 1998, vol. 30, p. 309–313.

Salmela-Aro K., Kiuru N., Leskinen E., Nurmi J-E. School Burnout Inventory (SBI): Reliability and validity // European Journal of Psychological Assessment. 2009, vol. 25 (1), p. 48–57.

Schwarzer R., Jerusalem M. Generalized Self-Efficacy scale // Measures in health psychology: A user's portfolio. Causal and control Beliefs / Ed. by J. Weinman, S. Wright, M. Johnston. Windsor, UK: NFER-NELSON, 1995. P. 35–37.

Stanton J. M. An empirical assessment of data collection using the Internet // Personnel Psychology. 1998, vol. 51 (3), p. 709–725.

Tolan P. H., Gorman-Smith D., Henry D. B. Chicago youth development study community and neighborhood measure: Construction and reliability technical report. Families and Communities Research Group, Department of Psychiatry, The University of Illinois at Chicago, 2001.

Vallejo M. A., Mañanes G., Isabel Comeche M. A., Díaz M. I. Comparison between administration via Internet and paper-and-pencil aministration of to cinical istruments: SCL-90-R and GHQ-28 // Journal of Behavior Therapy and Experimental Psychiatry. 2008, vol. 39 (3), p. 201–208.

Van De Looij-Jansen P. M., De Wilde E. J. Comparison of Web-based versus paper-and-pencil self-administered questionnaire: Effects on health indicators in dutch adolescents // Health Research and Educational Trust. 2008, vol. 43 (5), p. 1708–1721.

Van Selm M., Jankowski W. J. Conducting online surveys // Quality and Quantity. 2006, vol. 40, p. 435–456.

Wright D. L., Aquilino W. S., Supple A. J. A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a Survey on smoking, alcohol, and drug use // Public Opinion Quarterly. 1998, vol. 62, p. 331–353.

Wu Y., Newfield S. Comparing data collected by computerized and written surveys for adolescence health research // The Journal of School Health. 2007, vol. 77 (1), p. 23–28.

Yun G. W., Trumbo C. W. Comparative response to a survey executed by post, e- mail and Web form // Journal of Computer-Mediated Communication. 2000, vol. 6 (1) (retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2000.tb00112.x/full, 2012-02-03).

## POZITYVIOS RAIDOS INDIKATORIŲ, VERTINTŲ NAUDOJANT INTERNETINĘ APKLAUSĄ IR SPAUSDINTUS KLAUSIMYNUS, PALYGINIMAS PAAUGLIŲ IMTYJE

**Rimantas Vosylis, Rita Žukauskienė, Oksana Malinauskienė**

S a n t r a u k a

Internetinių klausimynų naudojimas tampa vis populiaresnė duomenų rinkimo priemonė, nes taip mažinamos tyrimo sąnaudos, tyrėjai gali susisiekti su sunkiau pasiekiamais tiriamaisiais. Tai labai patogu atliekant tęstinius, ypač jaunimo raidos, tyrimus, nes, baigę mokyklą, daug mokinių keičia savo gyvenamąją vietą, kai kurie jų išvyksta į užsienį. Tačiau internetinių apklausų taikymas, be minėtų pranašumų, turi ir trūkumų, į kuriuos būtina atsižvelgti. Pavyzdžiui, rezultatų, gautų taikant internetinius ir spausdintus klausimynus, derinimas, ir klausimyno formos poveikis tyrimui naudojamų skalių psichometriniams rodikliams.

Šio tyrimo tikslas – nustatyti internetinės apklausos metodo tinkamumą paauglių psichologinėms charakteristikoms vertinti naudojant apklausą internetu ir

palyginti dviejų klausimyno administravimo formų – elektroninės ir spausdintos – įtaką pozityvios jaunimo raidos rodikliams tęstiniame pozityvios jaunimo raidos tyrime. Tyrimo metu 1 030 dalyvių (17–19 metų), 505 berniukai ir 525 mergaitės, užpildė spausdintus klausimynus mokyklose, 132 (28 berniukai ir 104 mergaitės) užpildė tik internetinį klausimyno variantą, o 47 (15 berniukų ir 32 mergaitės) užpildė abi klausimyno versijas.

Tyrimo rezultatai atskleidė, jog paaugliai, apklausti internetinės apklausos būdu, nurodė mažiau socialiai pageidaujamo elgesio, bet apskritai internetinis apklausos būdas neturėjo poveikio pozityvios jaunimo raidos indikatorių vidurkiams, palyginti su tradiciniais, spausdintais, klausimynais. Internetinių klausimynų skalių vidinis suderinamumas, lyginant jį su spausdintų klausimynų skalių vidiniu suderinamumu, buvo geresnis arba panašus ir pozityvios raidos rodikliai, įvertinti abejomis klausimyno versijomis, buvo stipriai tarpusavyje susiję. Nerasta lyties ir klausimyno tipo sąveikos efekto, susijusio su paauglių pozityvios raidos vertinimu. Straipsnyje aptariami ir tyrimo ribotumai.

*Pagrindiniai žodžiai:* internetinis anketavimas, tyrimas naudojant spausdintinę klausimynų versiją, pozityvios raidos indikatoriai, paauglystė.

*Gauta: 2012 02 21*