The aim of this study was to examine the psychometric properties of the Polish version of the Short Hopelessness Scale in an undergraduate student sample. The scale was originally developed as a brief measure of hopelessness for large scale epidemiological surveys and is based on two commonly used measures of hopelessness.

A total of 4098 students from different Polish universities took part in the study. Confirmatory factor analysis (CFA) was used to test the model fit. Pearson correlation coefficients were used to investigate the relationship between hopelessness scale and criterion variables.

Due to the lack of fit of the initial congeneric, unidimensional model, the lowest loading item was removed, and a tau-equivalent model with three items was tested. The tau-equivalent model showed a good fit to the data. The scale had adequate reliability. Hopelessness was positively related to anxiety, depression, and stress and negatively related to the quality of life, health, and self-esteem.

The study provides evidence of good psychometric properties in terms of factorial structure, reliability and criterion validity of the shortened 3-item Polish version of a hopelessness scale. Future studies should investigate its predictive value, especially in terms of suicidal ideation and behaviour, as well as other health outcomes.

key words
hopelessness; quality of life; scale; tau-equivalence; well-being

Validity and reliability of the Polish version of the Short Hopelessness Scale
BACKGROUND

Lack of hope is identified as having a negative impact on well-being. Hopelessness, defined as an individual’s tendency to perceive themselves, the world, and their future in a pessimistic way (Beck, Weissman, Lester, & Trexler, 1974), is closely related to psychosocial functioning. Hopelessness is most frequently mentioned as a major component of depression (Liu, Kleiman, Nestor, & Cheek, 2015), and a predictor of suicide ideation (Ribeiro, Huang, Fox, & Franklin, 2018; Chioqueta & Stiles, 2005; Kuo, Gallo, & Eaton, 2004). However, it shows much broader implications. It may affect the treatment of mood disorder patients, having an impact on future health and social functioning (Pompili et al., 2013). It has an effect on substance use and abuse (Baines, Jones, & Christiansen, 2016). Furthermore, the sense of hopelessness plays a significant role in the course of somatic diseases. It affects the progression of cardiovascular disease by exacerbating the atherosclerotic process (Everson, Kaplan, Goldberg, & Salonen, 2000) or the treatment process of oncological patients (Price et al., 2016). Among patients with a lower sense of hopelessness, there was a higher survival rate of the disease. Hopelessness is also a predictor of psychological functioning among palliative care patients (Mystakidou et al., 2008). The meaning of this concept is significant both for psychological and physical health.

A positive association was found between hopelessness and anxiety (Ahookhosh, Bahmani, Asgari, & Moghaddam, 2017; Thompson, Mazza, Herting, Randell, & Eggert, 2005; Wilson & Deane, 2010), depression (Liu et al., 2015; Beck, 1963, 1967; Melges & Bowlby, 1969; Chioqueta & Stiles, 2005; Thompson et al., 2005; Wilson & Deane, 2010) and perceived stress (Hjemdal, Friborg, & Stiles, 2012; Rice, Leever, Christopher, & Porter, 2006). Hopelessness is related negatively with self-esteem (Crocker, Luhtanen, Blaine, & Broadnax, 1994; Dori & Overholser, 1999; Petrie & Brook, 1992; Tarrier, Barrowclough, Andrews, & Gregg, 2004) and quality of life (Gustafsson-Liljus, Julkunen, & Hietanen, 2007; Mystakidou et al., 2008).

The most commonly used measure of hopelessness is the Beck Hopelessness Scale (BHS; Beck, Epstein, Brown, & Steer, 1988). It is a 20-item self-report inventory which reflects negative expectancies in the respondent. The response format is dichotomous (true/false) and the items consist of describing the respondent’s attitude over the past week. To control for acquiescence, nine items are keyed false and 11 are keyed true. Previous studies showed good validity and reliability of the scale, including its shorter version in students, and its significance has been consistently showed as a predictor of suicide ideation, suicide attempts, and suicide completion (Hanna et al., 2011). The cross-cultural validations of this scale showed its good psychometric properties, including predictive validity (Rueda-Jaimes et al., 2018), and strict measurement invariance regarding gender and depression status in a representative national sample (Kliem, Lohmann, Mößle, & Brähler, 2018). Everson, Kaplan, Goldberg, Salonen, and Salonen (1997) used two items as indicators of hopelessness in their epidemiological survey. Responses were on a 5-point Likert scale. Hopelessness measured by these items proved to be a good predictor of cardiovascular outcomes in longitudinal studies.

Taking into account the important role of hopelessness in general well-being, it seems that it should be investigated in any epidemiological study of health, at least as a potential confounding variable. Using shorter versions of psychometric tools makes it possible to control many different variables in large surveys without imposing an unnecessary burden on the respondents. This may improve research quality by reducing the bias related to random responses by tired participants. Moreover, short scales often show equally satisfactory psychometric qualities as their lengthier counterparts, when assessed with proper statistical methods within specific measurement contexts (Rammstedt & Beierlein, 2014; Ziegler, Kemper, & Kruyten, 2014). It is important to note that although brief measurement tools are an appropriate fit for large scale studies, they are not adequate for individual assessment (Kemper, Trapp, Kammann, Samuel, & Ziegler, 2018). In order to provide a short and reliable method of measuring hopelessness, the Short Hopelessness Scale (SHS) was developed (Clarke, Fisher, House, Smith, & Weir, 2008). It consists of four items. Two of them were taken from the scale developed by Everson, Kaplan, Goldberg, and Salonen (1997), and two from the scale developed by Beck, Epstein, Brown, and Steer (1988). Also, a slightly modified 6-point response scale from I totally disagree to I totally agreed was used. This scale was used because it is a short tool already in use in a large epidemiological longitudinal study (the Health and Retirement Study) available from a Public Data Resource (Sonnega & Weir, 2014). The aim of this study was to examine the validity and reliability of the Polish version of the SHS.

Currently, a trend of decline in mental health is being observed in young generations (Twenge, 2000; Twenge et al., 2010), and there is some indication that these trends are present in Poland (Höfer, Rockett, Värnık, Etzersdorfer, & Kapusta, 2012). Therefore, there is a high need for monitoring the mental health of young adults and investigating the factors contributing to its decline. In recent years more attention has been devoted to the potential role of studying in the occurrence and persistence of psychopathological symptoms. School and academic failure are well-recognized risk factors for developing behavioural and psychiatric problems (Gustafsson et al., 2010).
over, excessive studying, including such phenomena as compulsive overstudying (see Atroszko, 2018; Lescalzo & Giannini, 2018), has been identified as a potential risk factor for depression and anxiety among students (Atroszko, 2015). More than 15% of students may be at risk of it (Lawendowski, Bereznowski, Wróbel, Kierzkowski, & Atroszko, 2019), which coincides with a high prevalence of depression among students, with recent estimates suggesting up to 21% experiencing mild depression (Czerwiński, Mackiewicz, Mytlewska, & Atroszko, 2020). In some countries, more than half of the undergraduate students may experience moderate to extremely severe levels of depression, anxiety, and stress (see Mamun, Hossain, & Griffiths, 2019). For these reasons, large scale rigorous epidemiological studies among students are of high importance.

On the basis of previous research and theoretical frameworks, it was hypothesized that: the Polish version of the SHS has satisfactory factorial validity and reliability (H1); hopelessness is positively related to neuroticism, anxiety, depression and perceived stress (H2); hopelessness is negatively related to self-esteem and quality of life (H3).

PARTICIPANTS AND PROCEDURE

PARTICIPANTS

Four thousand and ninety-eight students took part in the study: 2471 women (60.3%), 1575 men (38.4%), and 52 individuals (1.3%) who did not report gender. Their mean age was $M = 21.06$ years ($SD = 2.87$). Participants were studying at different faculties, courses, modes, and years of study at various universities in Poland in the Pomerania region.

MEASURE

Hopelessness. Hopelessness was measured with the SHS used in the Health and Retirement Study (Clarke et al., 2008). The participants were presented the following statements: (a) "I feel it is impossible for me to reach the goals that I would like to strive for", (b) "The future seems hopeless to me and I can’t believe that things are changing for the better", (c) "I don’t expect to get what I really want" and (d) "There’s no use in really trying to get something I want because I probably won’t get it". Respondents provided answers on a six-point Likert-type scale from 1 (I strongly disagree) to 6 (I totally agree). The Polish version was developed, including the back-translation procedure. Translations were prepared by psychologists and lay bilingual persons working separately. Experts in the fields of translations and psychometrician also participated in the adaptation process.

Anxiety and depression. The Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983) was used to measure anxiety and depression. The scale consists of seven items that measure the level of anxiety and seven that measure the level of depression. Respondents provided answers on a four-point scale with response categories from 0 to 3, whose labels vary depending on the items. It showed good validity and reliability in previous research (Bjelland, Dahl, Haug, & Neckelmann, 2002; Nezlek, Rusanowska, Holas, & Krejtz, 2019), including among Polish undergraduate students (Czerwiński et al., 2020). In the present sample, the Cronbach’s $\alpha$ reliability coefficients were .85 for anxiety and .74 for depression.

General quality of life and general health. General quality of life and general health were measured with single-item measures developed on the basis of the WHOQOL-BREF questionnaire (Skevington, Lotfy, & O’Connell, 2004). The general quality of life was measured with the question: “How satisfied are you with your health?” with a nine-point Likert scale from 1 (very poor) to 9 (very good). General health was measured with the question: “How satisfied are you with your health?” with a nine-point Likert scale, from 1 (very dissatisfied) to 9 (very satisfied). Both measures showed good validity and test-retest reliability (intraclass correlation coefficients were .86 for general quality of life and .72 for general health) in previous research (Atroszko, Bagińska, Mokosińska, Sawicki, & Atroszko, 2015b).

Perceived stress. The sort version of the Perceived Stress Scale (PSS-4; Cohen, Kamarck, & Merrelstein, 1983) was used to measure perceived stress. The scale consists of four items according to feelings and thoughts during the last month. Respondents provided answers on a five-point response scale from 0 (never) to 4 (very often). It showed good validity and reliability in previous studies (Lee, 2012), including samples of university students (Atroszko, Andreassen, Griffiths, & Pallesen, 2015a; Atroszko et al., 2015b). The Cronbach’s $\alpha$ reliability coefficient in the present sample was .72.

Global self-esteem. The Rosenberg Self-Esteem Scale (R-SES; Rosenberg, 1965) was used to measure global self-esteem in a subsample. The subsample consisted of 885 participants: 569 women (64.3%), 305 men (34.5%), and 11 individuals (1.2%) who did not report gender. Their mean age was $M = 20.38$ years ($SD = 2.64$). The scale consists of 10 items that apply to one’s global self-esteem. Respondents provided answers on a four-point scale from 1 (strongly agree) to 4 (strongly disagree). It showed good validity and reliability in previous studies (Schmitt & Allik, 2005). In the present sample, the Cronbach’s $\alpha$ was .87.

The Single-Item Self-Esteem scale (SISE; Robins, Hendin, & Trzesniewski, 2001) was another tool used to measure global self-esteem in a subsample. The subsample consisted of 1799 participants: 1179 wom-
en (65.5%), 603 men (33.5%), and 17 individuals (1%) who did not report gender. Their mean age was $M = 20.80$ years ($SD = 2.60$). The scale consists of one statement “I have high self-esteem”. Respondents provided answers on a five-point scale from 1 (not very true of me) to 5 (very true of me). It showed good validity and reliability in previous studies (Robins et al., 2001).

Global self-esteem was also measured with a single-item measure developed on the basis of the WHOQOL-BREF questionnaire (Skevington et al., 2004). The question was “How satisfied are you with yourself?” with a nine-point Likert scale from 1 (very dissatisfied) to 9 (very satisfied). It showed good validity and test-retest reliability (intra-class correlation coefficient was .79) in previous studies (Atroszko, Sendal, & Atroszko, 2017).

**PROCEDURE**

Students were invited to participate anonymously in the study during lectures or classes. Over 90% agreed to fill in paper and pencil questionnaire. No material rewards were offered to the participants. The respondents received information about the possibility of ceasing participation in the study at any time.

**STATISTICAL ANALYSES**

A confirmatory factor analysis (CFA) for the congeneric model was carried out in Mplus 6.11 (Muthén & Muthén, 1998-2010), using the weighted least square mean and variance adjusted (WLSMV) estimator due to non-normality of distributions of items. The CFA models differ depending on how restrictive the assumptions they make concerning the measurement of a latent variable are. The congeneric model is the most commonly used and the least restrictive CFA model. It assumes that each item measures the same latent variable, with possibly different scales, different degrees of precision, and with different amounts of error. On the other hand, the parallel model is the most restrictive and it requires that all items must measure the same latent variable, on the same scale, with the same degree of precision, and with the same amount of error. The tau-equivalent model loosens these assumptions and allows individual item error variances to differ from one another, and the essentially tau-equivalent model further allows for different degrees of precision of measurement (Raykov, 1997a, 1997b; Graham, 2006). However, even the latter two models are rarely used as their restrictions make it difficult to achieve model fit. Because the model did not show satisfactory fit indices, the lowest loading item was removed ("I feel it is impossible for me to reach the goals that I would like to strive for"). Testing for a congeneric model in the case of three-item scales does not yield meaningful results due to having zero degrees of freedom. For this reason a tau-equivalent was tested instead. Although this procedure is normally considered too rigorous in its assumptions, in the case of three-item scales due to lack of better alternatives and in the case of short scales in general, it can be a good solution for testing the factorial structure of an instrument. This is because having only a few items with equal factor loadings is not as difficult to obtain as it is in the case of larger scales, and demonstrating adequate fit of a tau-equivalent model is an indicator of excellent psychometric properties of a measure. Considering that the chi-squared test is highly sensitive to sample size (Bentler & Bonett, 1980), fit measures used in the CFA were the comparative fit index (CFI; Bentler, 1990), the Tucker-Lewis index (TLI; Tucker & Lewis, 1973) and the root mean square error of approximation (RMSEA; Steiger, 1990). TLI and CFI values above .95 are indicative of good fit (Hu & Bentler, 1999), while values above .90 are considered as adequate (Kline, 2004). RMSEA scores of .08 and lower are acceptable (Brown & Cudeck, 1993). After testing factorial validity, means, standard deviations, percentages and correlation coefficients were calculated. These statistical analyses were conducted in IBM SPSS 25.

**RESULTS**

The one-factor model of the Short Hopelessness Scale showed the following fit indices: $\chi^2/df = 93.74/2$, CFI = .99, TLI = .98, RMSEA = .106. Standardized factor loadings on items were: .76, .86, .90, .84, respectively. Due to the inadequate value of RMSEA, a tau-equivalent one-factor model with three items was tested. The modified model had a good fit: $\chi^2/df = 26.36/2$, CFI = .99, TLI = .99, RMSEA = .056, with each item loading of .87. Cronbach’s $\alpha$ coefficient was .86.

Table 1 presents mean scores, standard deviations and percentages of study variables as well as their correlation coefficients with hopelessness.

**DISCUSSION**

Confirmatory factor analysis showed that the original 4-item one-factor solution did not have a satisfactory fit. Instead, a three-item solution was found to have better psychometric properties via testing a tau-equivalent model. All factor loadings were significant, with standardized values of .87. The level of internal consistency was very good (H1 substantiated). The scale has better psychometric properties than a comparable short BHS scale, which showed both problems with item loadings and internal consistency (Hanna et al., 2017).
The results showed that hopelessness was positively related to anxiety, depression, and perceived stress, and negatively related to self-esteem and subjective quality of life (H2 and H3 substantiated). It was most strongly related to the general self-esteem measured with Rosenberg’s scale. These findings are congruent with previous studies.

CONCLUSIONS

Hopelessness plays a significant role in individuals’ psychosocial functioning and has been thoroughly studied for decades. The three-item version of the Polish adaptation of the SHS is a promising measure which showed good psychometric properties in terms of factorial structure, reliability, and criterion validity; however, more data are needed on its predictive validity. Due to its short form, the tool is convenient for researchers to use as a screening instrument and can be included in large scale psychological and epidemiological studies.

STRENGTHS AND LIMITATIONS OF THE STUDY

The main strengths of this study include the large sample and valid and reliable psychological questionnaires. Also, some of the constructs were measured by more than one scale, which ensures higher validity of the results. The results are highly consistent with previous studies and the theoretical underpinnings of the construct of hopelessness. This study constitutes a valuable input to the research on hopelessness. In terms of limitations, the sample was a homogeneous student sample and predominantly female, which means that the results cannot be generalized without some restrictions. Therefore, this study shows that SHS can be a valid measure in a non-clinical student population (age around 18-25), and more studies are needed to investigate its validity in other age groups and clinical populations. Furthermore, all data in the present study were self-reported, which increases the risk of common method bias and poses other limitations related to such data. Most of the instruments were short, including single-item measures, which imposes restrictions on the variance of results, and may result in the underestimation of the real strength of association between the variables (which could be seen for example in differences of magnitude of the relationship between the SHS and self-esteem, with the highest correlation values for the ten-item Rosenberg scale).

Future research should focus on obtaining and analyzing further data in various populations; especially, cross-cultural validations of this scale are required, including measurement invariance regarding gender and depression status in representative national samples (see Kliem et al., 2018). The predictive validity of the SHS should be examined in longitudinal designs, especially in relation to suicide ideation.

REFERENCES

Atroszko, P. A. (2015). The structure of study addiction: Selected risk factors and the relationship with stress, stress coping and psychosocial functioning (Un-
published doctoral thesis). University of Gdansk, Gdansk, Poland.


**VALIDITY AND RELIABILITY OF THE POLISH VERSION OF THE SHORT HOPELESSNESS SCALE**


