

A Confirmatory Factor-Analytic Evaluation of the Purpose in Life Test: Preliminary Psychometric Support for a Replicable Two-Factor Model

Stefan E. Schulenberg · Amanda M. A. Melton

Published online: 15 November 2008
© Springer Science+Business Media B.V. 2008

Abstract This study's purpose was to use confirmatory factor analysis to compare published factor-analytic models of the 20-item Purpose in Life test (PIL) to identify the one that provides the best fit to the data. To date many different models have been described, with limited evidence to support whether they are replicable. This study utilized data from undergraduates ($N = 620$) from a medium-sized university located in the southern United States. Ten different PIL models were tested, with support found for the two-factor model (exciting life, purposeful life) of Morgan and Farsides. Recommendations and implications for research are provided.

Keywords Logotherapy · Meaning · Purpose in life · PIL · Confirmatory factor analysis

1 Introduction

The Purpose in Life test (PIL; Crumbaugh and Maholick 1964, 1969) was originally grounded in the logotherapy paradigm. Logotherapy's focus is on the importance of perceived meaning and purpose in life and how these constructs enhance well-being (see for example, Frankl 1959/1985, 2004). Logotherapy emphasizes choice and responsibility, as well as the significance of deriving and living one's life by a personally meaningful hierarchy of values (Frankl 1959/1985; Schulenberg et al. *in press*). In other words, for meaning to be perceived, individuals should be aware of what life aspects are most vital and live their lives consistently with those values.

In logotherapy meaning is a broad construct that is conceptually and empirically related to many domains. For instance, meaning is positively related to constructs such as hope, faith, love, subjective well-being, health, and happiness, and negatively associated with

S. E. Schulenberg (✉) · A. M. A. Melton
Department of Psychology, The University of Mississippi, University, MS 38677, USA
e-mail: sschulen@olemiss.edu

constructs such as depression, anxiety, drug/alcohol use, boredom proneness, and psychological distress (Melton and Schulenberg 2007, 2008; Pöhlmann et al. 2006; Schulenberg et al. *in press*). Meaning is a construct central to the human condition (King et al. 2006; Schulenberg et al. *in press*; Wong and Fry 1998), and one that is a core facet of the Positive Psychology movement (Schulenberg et al. *in press*). Over the decades there has been much interest in how the meaning construct should be assessed, with the PIL being among the first and most influential of the quantitative measures.

The PIL was developed by James C. Crumbaugh to be a general tool to assess meaning (Crumbaugh and Maholick 1964, 1969), and it is commonly used and researched with regard to this purpose (Frazier et al. 2003; Reker 2000; Steger 2006). While the PIL is primarily regarded as a quantitative measure, it actually is composed of several parts, soliciting both quantitative (Part A) and qualitative (Parts B and C) data. Part A is of prime interest to researchers because it yields information that is easily quantified and compared across samples (Crumbaugh and Henrion 1988; Hutzell 1988). Parts B (13 incomplete sentences) and C (paragraph composition relating to future goals and past meaningful experiences) are of greater interest on a therapeutic level, given the focus on qualitative information (Crumbaugh and Henrion 1988; Hutzell 1988). Because Part A is often used in research and is easily quantified, it is the focus of the remainder of this article, and it is heretofore referred to as the PIL.

1.1 Reliability and Validity of the PIL

The PIL contains 20 items with a 7-point Likert-type response format. Each item uses different anchors, depending on the content of the item. Item content was designed to reflect the diversity of variables associated with the meaning construct. The content of each of the 20 PIL items are described in Table 1. Examples of item content include enthusiasm (versus boredom), excitement in living, contentment in daily tasks, and having a reason for being alive. A total score ranging from 20 to 140 is calculated based on the sum of the individual items, and higher scores are consistent with greater reports of personal meaning. The PIL does not require lengthy administration time and is easy to administer and score (Schulenberg and Henrion 2005). PIL scores are reliable, with internal consistency and split-half reliability coefficients commonly reported in the higher 0.70s to the lower 0.90s

Table 1 Purpose in Life test items

Item	General item content	Item	General item content
1	Enthusiasm (versus boredom)	11	Having a reason for being alive
2	Excitement in living	12	World being meaningful
3	Presence of clear life goals	13	Individual responsibility
4	Life being meaningful	14	Freedom in making decisions
5	Newness of each day	15	Being prepared for death
6	Wishing more lives	16	Suicidal thoughts
7	Activity after retirement	17	Capacity to discover meaning
8	Life goal completion	18	Life internally/externally determined
9	Good things in life	19	Contentment in daily tasks
10	Life lived having been worthwhile	20	Presence of goals/life purpose

(e.g., Crumbaugh and Henrion 1988; Hutzell 1988; McGregor and Little 1998; Melton and Schulenberg 2008; Reker 2000; Sink et al. 1998).

With regard to validity, growing support for the PIL has been noted in a number of studies or reviews (Chamberlain and Zika 1988; Crumbaugh and Henrion 1988; Crumbaugh and Maholick 1969; Hutzell 1988; Melton and Schulenberg 2007, 2008; Reker 2000; Reker and Cousins 1979). For example, scores on the PIL have been positively correlated with extroversion, life satisfaction, self-control, happiness, responsibility, self-acceptance, and emotional stability, among others, and they tend to be negatively correlated with depression and anxiety (Crumbaugh and Henrion 1988; Hutzell 1988, 1989; Reker 2000; Robak and Griffin 2000).¹ Such correlations are consistent with postulates of logotherapy and research studies that document the association between meaning and well-being (Melton and Schulenberg 2008; Schulenberg 2004; Schulenberg et al. *in press*).

The PIL has generated many research studies since its introduction in the 1960s. One particular area that continues to be of interest is whether the PIL is one-dimensional (Marsh et al. 2003; Reker 2000; Shek 1992), or whether there are theoretically valid and statistically viable subscales. Indeed, the PIL has been criticized on the grounds that many studies have reported different underlying factor structures for the measure, with limited evidence to suggest a consistently replicable model (e.g., Frazier et al. 2003; Steger 2006). Considering that such replication is essential to establishing structural validity, and that structural validity is an important, yet neglected aspect of a measure's psychometric viability (Steger 2006), it seems essential that the uncertainty regarding the PIL's dimensionality be addressed. If the PIL does have a consistently replicable factor structure, what would this mean for the PIL and for the meaning construct that it was designed to assess?

1.2 Factor-Analytic Investigations of the PIL

Many factor-analytic investigations of the PIL have been published over the years. They differ markedly in terms of the procedures employed and results. One useful way to organize them is by the number of factors proposed in the respective model. The current article focuses on one- and two-factor models of the PIL because they are more likely to be easier to interpret.²

1.2.1 One-Factor Models

There are two primary factor-analytic studies that propose one-factor models of the PIL, Marsh et al. (2003) and Steger (2006). In a study of the PIL using confirmatory factor-analytic methods, Marsh et al. investigated the PIL's structure with social and treatment drinkers ($N = 357$ and 137 , respectively), reporting a 17-item, one-dimensional model. Alternatively, Steger noted how decisions made when conducting exploratory factor analysis with respect to determining the number of factors to extract (i.e., the eigenvalue rule, the scree plot, parallel analysis) can play a major role in the results obtained from

¹ Some critics of the PIL have argued that certain items may be confounded too specifically with other constructs, such as depression (e.g., Dyck 1987; Steger 2006; Yalom 1980).

² Many authors, including Sato and Tanaka (1974), Reker and Cousins (1979), Harlow et al. (1987), and Chamberlain and Zika (1988), have reported PIL structures comprised of numerous (three or more) and varied factors, which is further evidence of how tangled this literature has become. These studies are problematic to interpret because they often report items patterning onto more than one factor (e.g., Sato and Tanaka 1974), or do not report how items pattern onto their respective factors (e.g., Harlow et al. 1987).

study to study. Steger focused on the PIL with a sample of undergraduate students ($N = 148$). Variations in models were obtained based on whether the eigenvalue rule was utilized, as compared to the scree plot and parallel analysis. In particular, the eigenvalue method supported five factors while the combination of scree plot and parallel analysis resulted in support for a 16 item, one-factor model. Given the parsimonious nature of Steger's one-factor model, it is of primary interest for the current study. The items comprising Steger's and Marsh et al.'s models are presented in Table 2.

1.2.2 Two-Factor Models

A variety of two-factor models of the PIL have appeared in the literature, most notably in reports by Walters and Klein (1980), Dufton and Perlman (1986), Molcar and Stuempfig (1988), Shek (1988), McGregor and Little (1998), Waisberg and Starr (1999), and Morgan and Farsides (*in press*). Similar to most studies of the PIL's factor structure, these investigations tend to involve exploratory techniques. The earliest of these studies, Walters and Klein (1980), reported on the factorial validity of the PIL with two random samples of high school students ($n = 349$ and 404) drawn from a larger sample ($N = 1,082$). The full 20-item PIL was not used. Items pertaining to retirement, death, suicide, responsibility, and freedom in making decisions were dropped. Such items, particularly those relating to death, suicide, and retirement, were considered to be inappropriate for large numbers of adolescents. Varimax rotation resulted in two factors, termed despair and enthusiasm. Despair (items 1, 3–4, 6, 8–9, 11–12, 20) was described as relating to a negative life approach, whereas enthusiasm (items 2, 5, 17–19) was interpreted as tapping life excitement.

Dufton and Perlman (1986) studied the factor structure of the PIL in a sample of college students ($N = 232$). Utilizing principal component factor analysis with varimax rotation, coupled with pattern coefficients in excess of 0.35 for items to pattern onto a factor and the eigenvalue rule to determine the number of factors to extract, the authors reported a two-factor solution consisting of life satisfaction (items 1–2, 5–6, 9–10, 19) and life purpose (items 3–4, 8, 11–12, 17, 20). Dufton and Perlman concluded that while there were two factors present, life satisfaction was the dominant factor. In a similar investigation of college students ($N = 201$), using principal components factor analysis with varimax rotation (0.40 was used as the cutoff for interpreting pattern coefficients), Molcar and Stuempfig (1988) reported a two-factor solution indicative of general meaning/purpose in life (items 3–4, 7–9, 11, 13, 17, 20) and excitement related to daily living (items 1–2, 5, 10, 12, 14, 18–19).

In a noteworthy factor-analytic study of a Chinese translation of the PIL, Shek (1988) utilized principal component analysis with varimax rotation with a sample of secondary school students ranging in age from 11 to 20 ($N = 2,140$). The author documented support for a two-factor model which was comprised of a quality of life factor (items 1–2, 5–6, 8–9, 11–12, 16, 19) and a meaning of existence factor (items 3–4, 13, 17–18, 20).³ This two-factor structure of the Chinese version of the PIL was further examined by Shek (1992, 1993) in studies of meaning and well-being with samples of Chinese college and secondary school students.

³ Shek et al. (1987) conducted a similar factor-analytic study of the Chinese PIL with 480 students (ages 18–25), with differences in how items pattern onto factors evident when the results are compared across studies. The 1988 model is of interest in the current study given its subsequent use in 1992 and 1993 investigations.

Table 2 Confirmatory factor analyses of one- and two-factor Purpose in Life test models ($N = 620$)

Model/sample	Sample	Factors	Items	χ^2/df	RMSR	GFI	AGFI	CFI	RMSEA	TLI	NFI
Crumbaugh and Maholick (1964, 1969)	225 participants from two non-clinical and three clinical samples	1	1–20	4.41	0.08	0.88	0.85	0.88	0.07	0.87	0.85
Marsh et al. (2003)	357 social drinkers, 137 treatment drinkers	1	1–6, 8–13, 16–20	7.57	0.13	0.83	0.78	0.83	0.10	0.81	0.81
Steger (2006)	148 college students	1	1–6, 8–13, 16–17, 19–20	5.78	0.08	0.88	0.84	0.89	0.09	0.87	0.87
Walters and Klein (1980) ^a	High school students: $n = 349$, $n = 404$	1	1, 3–4, 6, 8–9, 11–12, 20	6.11	0.07	0.90	0.86	0.90	0.09	0.88	0.88
		2	2, 5, 17–19								
Dufton and Perlman (1986)	232 college students	1	1–2, 5–6, 9–10, 19	10.31	0.28	0.87	0.82	0.83	0.12	0.80	0.82
		2	3–4, 8, 11–12, 17, 20								
Molcar and Stuempfig (1988)	201 college students	1	3–4, 7–9, 11, 13, 17, 20	4.50	0.08	0.90	0.87	0.90	0.08	0.89	0.88
		2	1–2, 5, 10, 12, 14, 18–19								
Shek (1988)	2,140 secondary school students	1	1–2, 5–6, 8–9, 11–12, 16, 19	5.16	0.08	0.89	0.86	0.90	0.08	0.88	0.88
		2	3–4, 13, 17–18, 20								
McGregor and Little (1998) ^b	146 university students	1	1–2, 5, 8–9, 19	7.60	0.07	0.93	0.87	0.93	0.10	0.90	0.92
		2	3, 17, 20								
Waisberg and Starr (1999)	146 substance abusers	1	3–4, 6, 8–13, 16–17, 20	4.47	0.08	0.90	0.87	0.91	0.08	0.90	0.89
		2	1–2, 5, 9, 18–19								
Morgan and Farsides (in press) ^c	University and community sample ($N = 200$)	1	2, 5, 7, 10, 17–19	4.23	0.06	0.96	0.93	0.95	0.07	0.93	0.93
		2	3, 8, 20								

Note. RMSR root mean square residual, GFI goodness of fit index, AGFI adjusted goodness of fit index, CFI comparative fit index, RMSEA root mean square error of approximation, TLI Tucker–Lewis Index, NFI normed fit index

^a Items 7, 10, and 13–16 were not included in the original model because they were deemed inappropriate for large groups of adolescents

^b Items 13–15 were not included in the original analysis due to low item-total correlations

^c Items 13–16 were not included in the original model because of low item-total correlations

In an often-cited investigation of happiness and meaning in relation to personal projects, McGregor and Little (1998) performed a principal-components analysis with the PIL with a sample of university students ($N = 146$). Given the nature of the study and previous research suggesting the possibility of a multifactor solution (referring to Chamberlain and Zika 1988), they were particularly interested in aspects of the PIL pertaining to happiness and meaning. They retained the first two factors of the solution, both of which had eigenvalues greater than one and both of which could be interpreted as happiness and meaning factors specifically. The happiness factor was comprised of items 1, 2, 5, 8, 9, and 19, while the meaning factor was comprised of items 3, 17, and 20 (items 13–15 were not included in the analysis due to low item-total correlations). Support was provided for each factor in a principal-components analysis of a variety of well-being measures.

McGregor and Little's (1998) research has generated much interest. For instance, Laura King, Joshua Hicks, and colleagues have utilized McGregor and Little's three-item meaning factor (items 3, 17, and 20), coupled with item 4 (personal existence being meaningful), in a series of studies relating to positive affect and meaning in relation to each other, as well as with other constructs, such as religious commitment, negative affect, and global/local focus (Hicks and King 2007, 2008; King et al. 2006), reporting coefficient alphas for these four PIL items ranging from the low 0.80s to 0.90 across multiple college student samples.

While most factor-analytic studies of the PIL focus on college students, Waisberg and Starr (1999) conducted their study with a sample of people who abused substances ($N = 146$), ultimately retaining a life is meaningful factor (items 3–4, 6, 8–13, 16–17, 20) and an interesting daily life factor (items 1–2, 5, 9, 18–19). Item 9 patterned onto both factors, and items 7, 14, and 15 were identified as possibilities for being dropped from the scale. These items did not pattern onto either factor and they had low item-total correlations in the inter-item reliability analysis.

Finally, in the most recent factor-analytic study of the PIL, and as part of their goal of developing a comprehensive measure of meaning based primarily on three, established measures of meaning, Morgan and Farsides (in press) factor analyzed the PIL with a sample of adults (a combined college and community sample) with a mean age of 32 ($SD = 12.35$, ranging in age from 15 to 75 years). Using principal-axis factoring with promax rotation, as well as the scree plot to determine the number of factors to extract, they retained two factors which they named exciting life (items 2, 5, 7, 10, 17–19) and purposeful life (items 3, 8, 20).

In order to better evaluate the two-factor models in relation to one another, as well as in comparison to the one-factor models discussed previously, they are also presented in Table 2. When examining both the one-factor and two-factor models, it becomes apparent that there is little consistency among the results in terms of the items that pattern onto the respective factors. Interpretation is further confounded by the fact that most of these models were derived from one sample, without having been validated on additional samples. Thus, whether the models are replicable becomes a question of central importance.

1.3 The Present Study

There are significant differences among the various factor-analytic studies available in the literature. This may be the result of different samples and variations in methodology, to name a couple of possibilities. Because of the substantial differences between factor-analytic studies, and due to some studies raising questions concerning the relevance of

specific PIL items, further research is warranted. The primary goal of the current investigation was to examine previously documented models of the underlying factor structure of the PIL, using confirmatory factor-analytic techniques to determine whether it is characterized by a replicable factor structure. Previously reported factor-analytic models that were examined included those proposed by Walters and Klein (1980), Dufton and Perlman (1986), Molcar and Stuempfig (1988), Shek (1988), McGregor and Little (1998), Waisberg and Starr (1999), Marsh et al. (2003), Steger (2006), and Morgan and Farsides (*in press*). These models were of primary interest given the extraction of one or two factors, limited to no items patterning onto more than one factor, and larger numbers of items per factor. Thus, these models were more parsimonious, easier to interpret, and more likely to be reliable. In addition, the original 20-item PIL (Crumbaugh and Maholick 1964, 1969) was examined as a one-factor model.

2 Method

2.1 Participants

Previously collected data from two college student samples were accessed for use in this study. While college students are often regarded as samples of convenience, they are important to study in such contexts given the increasing recognition of the significance of the meaning construct to this population (DeVogler and Ebersole 1980; Laverty et al. 2005). With respect to the current study, data from both student samples were collected in groups, with participants receiving class credit or extra credit for their involvement. The first dataset (Schulenberg 2004) consisted of 341 undergraduate students enrolled in psychology courses at a medium-sized university located in the southern United States (66% women, 34% men; 84% Caucasian, 14% African American; M age = 19.5 years, SD age = 1.6).

The second dataset (Melton and Schulenberg 2007) was comprised of 279 psychology undergraduate students drawn from the same population as sample 1 (64% women, 35% men; 77% Caucasian, 18% African American; M age = 19.8 years, SD age = 2.6). The purpose of the Schulenberg (2004) study (sample 1) was to collect psychometric data on various measures of meaning and to examine how they statistically relate to each other and to a measure of psychological distress. The goal of the Melton and Schulenberg (2007) study (sample 2) was to collect information regarding the association between meaning in life and boredom proneness in order to evaluate a fundamental logotherapy postulate. Given that both groups were drawn from the same population, with comparable demographic characteristics, and further considering the similarity of the data obtained from these samples [e.g., the PIL means and standard deviations were virtually identical for both samples (sample 1: $M = 107.8$, $SD = 15.4$; sample 2: $M = 108.3$, $SD = 15.3$)], the data from both samples were combined ($N = 620$) to enhance the subsequent statistical analyses.

2.2 Measures

2.2.1 Sample 1

In addition to the PIL, the measures of meaning employed with sample 1 included the Life Purpose Questionnaire (LPQ; Hablas and Hutzell 1982; Hutzell 1989), the Seeking of

Noetic Goals test (SONG; Crumbaugh 1977a, b), and the Meaning in Suffering Test (MIST; Starck 1983, 1985). The PIL and Life Purpose Questionnaire were designed to assess the general construct of meaning in life. For both measures, higher scores are suggestive of greater meaning. The Seeking of Noetic Goals test was developed to assess a person's motivation to find meaning in life, with higher scores interpreted to mean that the person has stronger motivation (in logotherapy terms noetic refers to the uniquely human aspects of the individual, such as humor). The Meaning in Suffering Test was developed to quantify the degree that a person has discovered meaning in unavoidable suffering, with higher scores suggestive of greater meaning (Starck 1983, 1985). For point of context for the current investigation, correlations among the PIL, Life Purpose Questionnaire, and Meaning in Suffering Test are generally significant and positive, whereas correlations between the Seeking of Noetic Goals test and the PIL, Life Purpose Questionnaire, and Meaning in Suffering Test are significant and negative (Schulenberg 2004). The Seeking of Noetic Goals pattern of correlation with these other measures is often interpreted as suggesting that higher reports of perceived meaning are associated with less of a need to find additional meaning. As reported originally in Schulenberg (2004), using the present data the coefficient alphas for the measures administered ranged from 0.82 (Life Purpose Questionnaire) to 0.91 (Purpose in Life test).

In addition to the aforementioned measures of meaning, the Outcome Questionnaire (OQ-45.2; Lambert et al. 1996) was included with sample 1 as a measure of psychological distress. The Outcome Questionnaire assesses an array of mental health problems, yielding a total score and three subscale scores: symptomatic functioning (items that tap depression, anxiety, and substance abuse), interpersonal problems (items that tap quality of interpersonal relationships), and social role adjustment (items that tap problems with family, work, and leisure).⁴ Higher Outcome Questionnaire total and subscale scores are suggestive of greater psychological distress. Considering the logotherapy framework and the associated research described in the introduction section of this article, it would be expected that Outcome Questionnaire scores would correlate significantly and negatively with the PIL, Life Purpose Questionnaire, and Meaning in Suffering Test, and significantly and positively with the Seeking of Noetic Goals test. These findings are anticipated given that psychological distress is related to meaninglessness and the need to find meaning or life purpose (Schulenberg 2004; Schulenberg et al. *in press*). As reported in Schulenberg (2004), based on the present data coefficient alphas for the Outcome Questionnaire total and subscale scores ranged from 0.68 (social role adjustment subscale) to 0.95 (total score).

2.2.2 Sample 2

As for measures used with sample 2, given that the purpose of the Melton and Schulenberg (2007) study was to better understand the relationship between perceived meaning in life and boredom proneness, both the PIL and the Boredom Proneness Scale (BPS; Farmer and Sundberg 1986) were administered to a college student sample. Higher scores on the BPS are associated with greater boredom proneness.⁵ With respect to providing context for the current article, considering the logotherapy paradigm (e.g., Frankl 1959/1985), meaning as assessed by the PIL should be associated with lower scores on the Boredom Proneness

⁴ For a more in-depth discussion of the measures used with sample 1, including a description of their items, format, and psychometric properties, the reader is referred to Schulenberg (2004).

⁵ For a more extensive description of the Boredom Proneness Scale and its psychometric properties, the reader is referred to Melton and Schulenberg (2007).

Scale, such that the correlation is negative and statistically significant. In other words, meaninglessness should be associated with boredom proneness. Melton and Schulenberg (2007) reported coefficient alphas for the PIL and the Boredom Proneness Scale to be 0.90 and 0.85, respectively.

3 Results

3.1 Confirmatory Factor Analyses

The previously mentioned factor-analytic models were imposed on the combined data from samples 1 and 2 ($N = 620$). Maximum likelihood estimation, the most popular and justified method (Bryant and Yarnold 1995; Hoyle 2000; Thompson 2004), was used in the analyses, which were run on AMOS 6 software. Fit indices were chosen a priori to determine whether the particular model tested was a reasonable fit to the data, and included the comparative fit index (CFI; Bentler 1990), the goodness of fit index (GFI; Jöreskog and Sörbom 1981), the adjusted goodness of fit index (AGFI; Jöreskog and Sörbom 1989), the normed fit index (Bentler and Bonett 1980), and the Tucker–Lewis index (Bentler and Bonett 1980).⁶ Values for these indices tend to range from 0 to 1.0, with larger values (some assert 0.90 as a lower-bound threshold, while others advocate values of at least 0.95) interpreted to suggest reasonable model fit (Brown 2006; Bryant and Yarnold 1995; Hair et al. 2006; Hoyle 2000; Kline 2005; Thompson 2004).

Other indices, in which lower values are interpreted as suggestive of reasonable model fit, were calculated and include the root mean square residual (RMSR; Jöreskog and Sörbom 1981), the root mean square error of approximation (RMSEA; Steiger 1990), and the ratio of χ^2 to degrees of freedom (χ^2/df ; Hoelter 1983). While there is some difference of opinion regarding acceptable values for RMSR and RMSEA, values of 0.06 and smaller were sought as being suggestive of reasonable model fit based on consultation with multiple sources (Brown 2006; Browne and Cudeck 1993; Bryant and Yarnold 1995; Thompson 2004). Smaller χ^2/df ratios (i.e., generally smaller than three and those approaching zero) are suggestive of reasonable model fit, but similar to other fit indices, specific interpretive cut-offs are often a point of debate (Bollen 1989; Brown 2006; Bryant and Yarnold 1995; Hair et al. 2006; Hoelter 1983).

The fit indices for the combined data are presented in Table 2, organized by the proposed models. Based on the aforementioned criteria, χ^2/df , RMSR, and RMSEA values tended to be too high, while GFI, AGFI, CFI, TLI, and NFI values tended to be too low. One interesting aspect to note about the models presented in Table 1 relates to the original 20-item PIL. While many indices of the model are not within the range that would be considered to suggest reasonable fit, it is intriguing that many of the indicators are quite comparable, and in some cases better, than models that were developed as a refinement of the original measure. Ultimately, however, the 10-item, two-factor model (exciting life, purposeful life) proposed by Morgan and Farsides (in press) stands out as the strongest model, with all indices except χ^2/df and RMSEA falling within acceptable standards. χ^2/df may be less of a concern in this case given it is the lowest of all the tested models. This particular indicator is not as widely regarded as it once was given a particular lack of consensus regarding interpretive guidelines, and it also hasn't experienced the same degree

⁶ For a thorough discussion of fit indices, the reader is referred to Hoyle (2000), Thompson (2004), Kline (2005), Hair et al. (2006), and Brown (2006).

of rigorous empirical study as has some of the other indices, such as RMSEA (Brown 2006). With regard to RMSEA, which is often further interpreted with a 90% confidence interval (Brown 2006; Kline 2005), although the value itself is slightly high, it is within an acceptable range based on some standards (e.g., Browne and Cudeck 1993). Moreover, the confidence interval is narrow, from an acceptable low of 0.060 to a high of 0.085.

3.2 Reliability, Descriptive, and Correlational Analyses with Morgan and Farsides' Model

Internal consistency reliability coefficient alphas were calculated for each factor using the current data (factor 1 = 0.79, factor 2 = 0.75), and are acceptable by a variety of standards (DeVellis 2003; Nunnally and Bernstein 1994; Vogt 2005). In the original Morgan and Farsides article, the coefficient alphas for factors 1 and 2 were 0.88 and 0.77, respectively.

Descriptive data were not reported in the original study considering the authors' focus on developing a larger, more comprehensive measure of meaning. Theoretically, given that each item uses a 7-point, Likert-type response format, scores may range from 7 to 49 for factor 1 (seven items) and 3 to 21 for factor 2 (three items). Using the current data ($N = 620$), the descriptive statistics are reported here for each factor as context for future research: factor 1 ($M = 37.68$, $SD = 6.02$, minimum = 10, maximum = 49) and factor 2 ($M = 16.69$, $SD = 2.62$, minimum = 6, maximum = 21).

With regard to correlations, in the Morgan and Farsides (in press) study, the association between PIL factors 1 (exciting life) and 2 (purposeful life) was 0.58 (significant at 0.01, 1-tailed). The correlation between these factors using the current data ($N = 620$) was calculated to be 0.65 (significant at 0.01, 1-tailed). These factors share approximately 42% of the variance.

To further assess the validity of this two-factor model of the PIL, correlations were calculated between these factors and the measures of meaning and psychological distress used originally with samples 1 and 2. These data are presented in Table 3, with all correlations statistically significant at 0.01 (1-tailed) and in the expected directions given what these measures purport to assess. However, the pattern of correlations between the PIL factors and the measures of meaning and psychological distress do little to further discriminate among the factors.

Table 3 Correlations of Morgan and Farsides' model of the PIL with measures of meaning and psychological distress

	LPQ	SONG	MIST	OQ-45.2	OQ-SD	OQ-IR	OQ-SR	BPS
Factor 1 (exciting life)	0.70**	-0.35**	0.39**	-0.64**	-0.63**	-0.57**	-0.50**	-0.67**
Factor 2 (purposeful life)	0.65**	-0.33**	0.41**	-0.52**	-0.51**	-0.47**	-0.41**	-0.54**

Note. PIL Purpose in Life test, LPQ Life Purpose Questionnaire, SONG Seeking of Noetic Goals test, MIST Meaning in Suffering Test, OQ-45.2 Outcome Questionnaire, OQ-SD symptomatic functioning subscale, OQ-IR interpersonal problems subscale, OQ-SR social role adjustment subscale, BPS Boredom Proneness Scale. The correlations between the PIL factors and the LPQ, SONG, MIST, and OQ-45.2 scores were calculated using data from sample 1 ($N = 341$), while the correlations between the PIL factors and the BPS were calculated using data from sample 2 ($N = 279$)

**Correlation significant at $p < 0.01$ (1-tailed)

4 Discussion

Crumbaugh originally developed the PIL to be a general measure of meaning, and the item content contained in the PIL covers an array of topics within and related to the meaning construct as initially theorized in the logotherapy paradigm. Thus, in many respects, given diverse item content, it is not surprising that studies seeking to reorganize the measure into subscales are able to use exploratory factor analysis to locate smaller clusters of items within the overall framework of the instrument. However, researchers often conduct such analyses on the PIL without thorough explanations of the applicability of their techniques, often reporting subscales that have little theoretical value or utility, and have not been replicated in other samples. Such problems are evident in the larger exploratory factor-analytic literature as well (see Reise et al. 2000 for a discussion of factor analysis use in scale revision).

The primary purpose of the current research was to examine various factor-analytic models of the PIL that have appeared in the literature. Focus was placed on models that were more likely to be parsimonious, interpretable, and reliable. The primary impetus behind this empirical investigation was the knowledge that the factor-analytic literature of the PIL is characterized by inconsistency. Such inconsistency is likely due to a variety of factors, such as differences in samples and variations in methodology across studies. For example, studies employing exploratory factor analysis use variations in methods of extraction (e.g., principal axis, principal components) and rotation (e.g., orthogonal versus oblique), as well as differing guidelines on how many factors to extract, such as the eigenvalue rule (Guttman 1954; Kaiser 1960, 1970), Cattell's (1966) scree test, and/or parallel analysis (Horn 1965).

The Kaiser–Guttman rule of retaining eigenvalues larger than 1.0 is often not inter-pretively useful because it tends to result in the retention of too many factors (Nunnally and Bernstein 1994; Reise et al. 2000; Zwick and Velicer 1986). Cattell's (1966) scree test provides a graphical presentation of the number of factors to extract by an examination of plotted eigenvalue magnitudes (for more on this procedure see Comrey and Lee 1992; Gorsuch 1983; Thompson 2004); however, it is often characterized by subjectivity in terms of interpretation. There is growing consensus that an alternative approach, Horn's (1965) parallel analysis, may be particularly useful with regard to determining the number of factors to extract (Hayton et al. 2004; Thompson 2004; Zwick and Velicer 1986), and this method may become increasingly prominent in factor-analytic investigations of the PIL. These are but a few of the factor extraction methods available, with each having the potential to yield different results with regard to the number of factors retained. With specific regard for the PIL, researchers cannot seem to agree on whether and how factors should be extracted, or which items pattern together to compose a factor. A complicating aspect of this literature is that many researchers propose a factor structure for the PIL without validation with an independent sample. All of these issues likely play a role in the inconsistency evident in the results of factor-analytic studies of the PIL.⁷

With these issues in mind, the current investigation employed confirmatory factor analysis to investigate which of the factor-analytic models proposed in the literature was the best fit to the data. Previously collected undergraduate student data ($N = 620$) were used to examine models of interest (those consisting of one or two factors, few or no items patterning onto more than one factor, and many items per factor). The Walters and Klein (1980), Dufton and Perlman (1986), Molcar and Stuempfig (1988), Shek (1988), McGregor

⁷ For further discussion on this topic relating to the PIL, the reader is referred to Steger (2006).

and Little (1998), Waisberg and Starr (1999), Marsh et al. (2003), Steger (2006), and Morgan and Farsides (*in press*) models were tested, along with the original PIL established by Crumbaugh and Maholick (1964, 1969) (Table 2). Overall, the data support the two-factor model of Morgan and Farsides, independently and in relation to the other models tested. Support for the Morgan and Farsides model is an important finding. Any factor structure that is derived must be replicable across samples within the same population, generalizable from one population to the next (Reise et al. 2000), and theoretically sound and psychometrically stable. Based on the current data, the PIL does appear to have a replicable, theoretically useful, and psychometrically viable subscale structure, referring to Morgan and Farsides' model comprised of two factors (exciting life, purposeful life).

The two PIL factors are related, sharing approximately 42% of the variance. To better understand each factor's unique relationship to other measures, patterns of correlation were calculated between other indices of meaning, as well as psychological distress. While the correlational analyses (Table 3) were of the expected significance and direction, they did not shed much additional light regarding the nature of these factors due to the relatively similar patterns of correlation between each factor and the other measures employed. This finding is not entirely unexpected, however, given the narrow range of external criteria that were available and which were used for purposes of comparison with the PIL factors. For instance, all of the measures displayed in Table 3 are self-report instruments, and most of them assess concepts similar to the PIL. Thus, the number of variables should be expanded in future research to include those that would theoretically be related, although not necessarily similar to, the PIL. External criteria can be further expanded to include not only self-report questionnaires, but specific behaviors potentially relevant to the meaning construct as well (e.g., volunteer activity participation). Such studies have implications for theory, and appear to touch on what Fiske (1971) termed *extrinsic convergent validity*, which was further discussed by Lubinski (2004, 2006) in recent publications.⁸

With respect to the PIL as a 20-item, one-factor model, while it was not found to be a reasonable fit given the current data, it is interesting to note that it was comparable to many of the models that were developed as a means of refining the measure. This finding is not interpreted as a condemnation of the PIL per se. The PIL has been widely researched for over 40 years since its publication. The measure yields reliable scores, and correlates with numerous constructs across many studies, significantly and in the expected direction depending on the construct of interest (see Hutzell 1988 and Melton and Schulenberg 2008 for reviews). However, from the standpoint of structural validity, the two-factor model of Morgan and Farsides is an improvement over the original 20-item PIL, based on these college-student data. In all likelihood, despite the results of the current study supporting Morgan and Farsides' two-factor model, the content and subscale organization of the measure will continue to be debated. Some researchers and clinicians will continue to prefer the 20-item PIL because of the 40-plus year history of research on the measure, while others will prefer to revise or otherwise use variations of the PIL, perhaps due to the criticism of items being too closely confounded with other factors. To this end, factor-analytic investigations are useful in parceling out items that may be too closely related to other constructs, and remain an important line of future empirical inquiry.

⁸ Extrinsic convergent validity is the concept that measures "may be considered conceptually equivalent and empirically interchangeable if they display corresponding correlational profiles across a heterogeneous collection of external criteria" (Lubinski 2006, p. 109). Lubinski (2004) provided a specific example relating to different measures of verbal content.

4.1 Additional Directions for Research

With regard to the 20-item PIL, replication of these data with multiple samples drawn from other populations is warranted to further examine the measure's factor structure. While the current results are viewed as preliminary support for a two-factor model, factor structure may vary by population. Samples should be drawn from clinical and community populations, and should be increasingly diverse in terms of sex, age, and racial/ethnic background. The respondents who participated in this research were predominantly female, Caucasian undergraduates. If researchers choose to examine previously developed models, confirmatory factor-analytic methods are useful in this respect. However, if researchers choose to use exploratory factor analysis to develop a new model of the PIL, these findings need to be replicated in additional samples, and confirmatory factor-analytic procedures with these additional samples is one means of replication (Reise et al. 2000). Moreover, given the sheer number of models developed with exploratory factor analysis, resulting in a convoluted literature with respect to factorial validity, new models should not be formulated without good reason.

Another interesting area that appears to be useful to explore empirically is that of short form development. One observation from the current data is that Morgan and Farsides' (in press) purposeful life factor (items 3, 8, and 20) is quite similar to the three-item meaning factor (items 3, 17, and 20) of McGregor and Little (1998). If Morgan and Farsides' purposeful life factor is isolated, and the logic of King, Hicks, and colleagues as relates to the McGregor and Little factor is applied, that is, adding item 4 because it asks specifically about meaning in life, the reliability of the purposeful life factor improves from 0.75 to 0.81. Moreover, the cluster of these four items is not only supported strongly by confirmatory factor analysis with the current data ($\chi^2/\text{df} = 0.30$, $\text{RMSR} = 0.01$, $\text{GFI} = 1.00$, $\text{AGFI} = 1.00$, $\text{NFI} = 1.00$, $\text{TLI} = 1.01$, $\text{CFI} = 1.00$, and $\text{RMSEA} = 0.00$ with a 90% confidence interval ranging from 0.00 to 0.06), this specific clustering of items is consistent with the previously noted research of Walters and Klein (1980), Dufton and Perlman (1986), Molcar and Stuempfig (1988), Waisberg and Starr (1999), Marsh et al. (2003), and Steger (2006), not to mention Yalom's (1980) observation that these four items are among those that deal explicitly with life meaning/purpose (p. 456).

Clearly there is interest in examining smaller clusters of PIL items as to their psychometric utility. Such empirical endeavors should not be undertaken lightly, however. Rigorous procedures should be followed to ensure that such a form is a reliable and valid construct and a strong representation of the larger PIL, not to mention the need to consider salient issues such as bandwidth and fidelity (see for example, Aiken and Groth-Marnat 2006; Cronbach 1984; Reise et al. 2000; Smith and McCarthy 1995; Smith et al. 2000; Wei et al. 2007 for discussions of relevant psychometric issues with respect to scale validation and short-form development). One could not assume the psychometric properties of a short form based on the reasoning that items were drawn from a well-studied parent measure. Short forms need to be validated independently of their respective larger forms.

Additionally, the advent of promising measures such as the Meaning in Life Questionnaire must be considered in this context (MLQ; Steger et al. 2006). While the Meaning in Life Questionnaire is a recent measure, it has been used in a growing number of investigations (Kashdan and Breen 2007; Steger and Kashdan 2007; Strack and Schulenberg 2008). The Meaning in Life Questionnaire contains a five-item Presence subscale which measures meaning, and a specific avenue of empirical investigation would be to determine whether a brief form of the PIL (e.g., comprised of items 3, 4, 8, and 20) measures meaning similarly to this subscale (in which case the question as to whether a

PIL short form is needed arises), or whether it assesses a different, perhaps complementary, aspect of meaning (in which case there would be utility in validating such a short form).

4.2 Concluding Comments

As for enhancing the understanding of the PIL's dimensionality, this study was successful in providing data that replicates a new model, namely the two-factor model proposed by Morgan and Farsides ([in press](#)). The current study serves as an initial step in clarifying this tangled literature, and points to avenues for research with this often-used measure of meaning.

Acknowledgments The authors would like to acknowledge the assistance of Dr. Schulenberg's research team with data collection and data entry, and Dr. William T. Hoyt for helpful comments and suggestions on a previous draft of this manuscript.

References

- Aiken, L. W., & Groth-Marnat, G. (2006). *Psychological testing and assessment* (12th ed.). Boston, MA: Pearson.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238–246. doi:[10.1037/0033-2909.107.2.238](#).
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588–606. doi:[10.1037/0033-2909.88.3.588](#).
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
- Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99–135). Washington DC: American Psychological Association.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245–276. doi:[10.1207/s15327906mbr0102_10](#).
- Chamberlain, K., & Zika, S. (1988). Measuring meaning in life: An examination of three scales. *Personality and Individual Differences*, 9, 589–596. doi:[10.1016/0191-8869\(88\)90157-2](#).
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cronbach, L. J. (1984). *Essentials of psychological testing* (4th ed.). New York, NY: Harper and Row.
- Crumbaugh, J. C. (1977a). *Manual of instructions: The Seeking of Noetic Goals test (SONG)*. Abilene, TX: Viktor Frankl Institute of Logotherapy.
- Crumbaugh, J. C. (1977b). The Seeking of Noetic Goals test (SONG): A complementary scale to the purpose in life test (PIL). *Journal of Clinical Psychology*, 33, 900–907. doi:[10.1002/1097-4679\(197707\)33:3<900::AID-JCLP2270330362>3.0.CO;2-8](#).
- Crumbaugh, J. C., & Henrion, R. (1988). The PIL test: Administration, interpretation, uses, theory and critique. *International Forum for Logotherapy*, 11, 76–88.
- Crumbaugh, J. C., & Maholick, L. T. (1964). An experimental study in existentialism: The psychometric approach to Frankl's concept of noogenic neurosis. *Journal of Clinical Psychology*, 20, 200–207. doi:[10.1002/1097-4679\(196404\)20:2<200::AID-JCLP2270200203>3.0.CO;2-U](#).
- Crumbaugh, J. C., & Maholick, L. T. (1969). *Manual of instructions for the Purpose in Life test*. Abilene, TX: Viktor Frankl Institute of Logotherapy.
- DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.
- DeVogler, K. L., & Ebersole, P. (1980). Categorization of college students' meaning of life. *Psychological Reports*, 46, 387–390.
- Duften, B. D., & Perlman, D. (1986). The association between religiosity and the purpose-in-life test: Does it reflect purpose or satisfaction? *Journal of Psychology and Theology*, 14, 42–48.
- Dyck, M. J. (1987). Assessing logotherapeutic constructs: Conceptual and psychometric status of the purpose in life and seeking of noetic goals tests. *Clinical Psychology Review*, 7, 439–447. doi:[10.1016/0272-7358\(87\)90021-3](#).

- Farmer, R., & Sundberg, N. D. (1986). Boredom proneness—The development and correlates of a new scale. *Journal of Personality Assessment*, 50, 4–17. doi:[10.1207/s15327752jpa5001_2](https://doi.org/10.1207/s15327752jpa5001_2).
- Fiske, D. W. (1971). *Measuring the concepts of personality*. Chicago, IL: Aldine Publishing Company.
- Frankl, V. E. (1959/1985). *Man's search for meaning* (Rev. ed.). New York, NY: Washington Square Press.
- Frankl, V. E. (2004). *On the theory and therapy of mental disorders: An introduction to logotherapy and existential analysis* (J. M. DuBois, Trans.). New York, NY: Brunner-Routledge (Original work, 8th ed., published 1999).
- Frazier, P., Oishi, S., & Steger, M. (2003). Assessing optimal human functioning. In W. B. Walsh (Ed.), *Counseling psychology and optimal human functioning* (pp. 251–278). Mahwah, NJ: Erlbaum.
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, 19, 149–161. doi:[10.1007/BF02289162](https://doi.org/10.1007/BF02289162).
- Hablas, R., & Hutzell, R. (1982). The Life Purpose Questionnaire: An alternative to the Purpose-in-Life test for geriatric, neuropsychiatric patients. In S. A. Wawrytko (Ed.), *Analecta Frankliana: The proceedings of the First World Congress of Logotherapy: 1980* (pp. 211–215). Berkeley, CA: Strawberry Hill.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Harlow, L. L., Newcomb, M. D., & Bentler, P. M. (1987). Purpose in life test assessment using latent variable methods. *The British Journal of Clinical Psychology*, 26, 235–236.
- Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, 7, 191–205. doi:[10.1177/1094428104263675](https://doi.org/10.1177/1094428104263675).
- Hicks, J. A., & King, L. A. (2007). Meaning in life and seeing the big picture: Positive affect and global focus. *Cognition and Emotion*, 21, 1577–1584. doi:[10.1080/02699930701347304](https://doi.org/10.1080/02699930701347304).
- Hicks, J. A., & King, L. A. (2008). Religious commitment and positive mood as information about meaning in life. *Journal of Research in Personality*, 42, 43–57. doi:[10.1016/j.jrp.2007.04.003](https://doi.org/10.1016/j.jrp.2007.04.003).
- Hoelter, J. W. (1983). The analysis of covariance structures: Goodness of fit indices. *Sociological Methods & Research*, 11, 325–344. doi:[10.1177/0049124183011003003](https://doi.org/10.1177/0049124183011003003).
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179–185. doi:[10.1007/BF02289447](https://doi.org/10.1007/BF02289447).
- Hoyle, R. H. (2000). Confirmatory factor analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 465–497). San Diego, CA: Academic Press.
- Hutzell, R. R. (1988). A review of the Purpose in Life test. *International Forum for Logotherapy*, 11, 89–101.
- Hutzell, R. R. (1989). *Life Purpose Questionnaire overview sheet*. Berkeley, CA: Institute of Logotherapy Press.
- Jöreskog, K. G., & Sörbom, D. (1981). *LISREL V: Analysis of linear structural relationships by the method of maximum likelihood*. Chicago, IL: National Educational Resources.
- Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7 user's reference guide*. Chicago, IL: Scientific Software.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141–151. doi:[10.1177/001316446002000116](https://doi.org/10.1177/001316446002000116).
- Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, 35, 401–415. doi:[10.1007/BF02291817](https://doi.org/10.1007/BF02291817).
- Kashdan, T. B., & Breen, W. E. (2007). Materialism and diminished well-being: Experiential avoidance as a mediating mechanism. *Journal of Social and Clinical Psychology*, 26, 521–539. doi:[10.1521/jscp.2007.26.5.521](https://doi.org/10.1521/jscp.2007.26.5.521).
- King, L. A., Hicks, J. A., Krull, J. L., & Del Gaiso, A. K. (2006). Positive affect and the experience of meaning in life. *Journal of Personality and Social Psychology*, 90, 179–196. doi:[10.1037/0022-3514.90.1.179](https://doi.org/10.1037/0022-3514.90.1.179).
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.
- Lambert, M. J., Hansen, N. V., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., et al. (1996). *Administration and scoring manual for the OQ-45.2*. Stephenson, MD: American Professional Credentialing Services LLC.
- Laverty, W. H., Pringle-Nelson, C., Kelly, I. W., Miket, M. J., & Janzen, B. L. (2005). Expressions of life meaning among college students. *Psychological Reports*, 97, 945–954. doi:[10.2466/PRO.97.7.945-954](https://doi.org/10.2466/PRO.97.7.945-954).
- Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) “‘General intelligence’, objectively determined and measured”. *Journal of Personality and Social Psychology*, 86, 96–111. doi:[10.1037/0022-3514.86.1.96](https://doi.org/10.1037/0022-3514.86.1.96).
- Lubinski, D. (2006). Ability tests. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 101–114). Washington, DC: American Psychological Association.

- Marsh, A., Smith, L., Piek, J., & Saunders, B. (2003). The purpose in life scale: Psychometric properties for social drinkers and drinkers in alcohol treatment. *Educational and Psychological Measurement*, 63, 859–871.
- McGregor, I., & Little, B. R. (1998). Personal projects, happiness, and meaning: On doing well and being yourself. *Journal of Personality and Social Psychology*, 74, 494–512. doi:[10.1037/0022-3514.74.2.494](https://doi.org/10.1037/0022-3514.74.2.494).
- Melton, A. M. A., & Schulenberg, S. E. (2007). On the relationship between meaning in life and boredom proneness: Examining a logotherapy postulate. *Psychological Reports*, 101, 1016–1022. doi:[10.2466/PRO.101.7.1016-1022](https://doi.org/10.2466/PRO.101.7.1016-1022).
- Melton, A. M. A., & Schulenberg, S. E. (2008). On the measurement of meaning: Logotherapy's empirical contributions to Humanistic psychology. *The Humanistic Psychologist*, 36, 31–44. doi:[10.1080/08873260701828870](https://doi.org/10.1080/08873260701828870).
- Molcar, C. C., & Stuempfig, D. W. (1988). Effects of world view on purpose in life. *The Journal of Psychology*, 122, 365–371.
- Morgan, J., & Farsides, T. (2007). Measuring meaning in life. *Journal of Happiness Studies*. doi:[10.1007/s10902-007-9075-0](https://doi.org/10.1007/s10902-007-9075-0).
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Pöhlmann, K., Gruss, B., & Joraschky, P. (2006). Structural properties of personal meaning systems: A new approach to measuring meaning of life. *Journal of Positive Psychology*, 1, 109–117. doi:[10.1080/17439760600566008](https://doi.org/10.1080/17439760600566008).
- Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, 12, 287–297. doi:[10.1037/1040-3590.12.3.287](https://doi.org/10.1037/1040-3590.12.3.287).
- Reker, G. T. (2000). Theoretical perspective, dimensions, and measurement of existential meaning. In G. T. Reker & K. Chamberlain (Eds.), *Exploring existential meaning: Optimizing human development across the life span* (pp. 39–55). Thousand Oaks, CA: Sage.
- Reker, G. T., & Cousins, J. B. (1979). Factor structure, construct validity and reliability of the seeking of noetic goals (SONG) and purpose in life (PIL) tests. *Journal of Clinical Psychology*, 35, 85–91. doi:[10.1002/1097-4679\(197901\)35:1<85::AID-JCLP2270350110>3.0.CO;2-R](https://doi.org/10.1002/1097-4679(197901)35:1<85::AID-JCLP2270350110>3.0.CO;2-R).
- Robak, R. W., & Griffin, P. W. (2000). Purpose in life: What is its relationship to happiness, depression, and grieving? *North American Journal of Psychology*, 2, 113–120.
- Sato, F., & Tanaka, H. (1974). An experimental study on the existential aspect of life: I. The cross-cultural approach to purpose in life. *Tohoku Psychologica Folia*, 33, 20–46.
- Schulenberg, S. E. (2004). A psychometric investigation of logotherapy measures and the Outcome Questionnaire (OQ-45.2). *North American Journal of Psychology*, 6, 477–492.
- Schulenberg, S. E., & Henrion, R. P. (2005). Logotherapy past, present, and future: A conversation with James C. Crumbaugh. *International Forum for Logotherapy*, 28, 65–71.
- Schulenberg, S. E., Hutzell, R. R., Nassif, C., & Rogina, J. M. Logotherapy for clinical practice. *Psychotherapy: Theory, Research, Practice, Training* (in press).
- Shek, D. T. L. (1988). Reliability and factorial structure of the Chinese version of the purpose in life questionnaire. *Journal of Clinical Psychology*, 44, 384–392. doi:[10.1002/1097-4679\(198805\)44:3<384::AID-JCLP2270440312>3.0.CO;2-I](https://doi.org/10.1002/1097-4679(198805)44:3<384::AID-JCLP2270440312>3.0.CO;2-I).
- Shek, D. T. L. (1992). Meaning in life and psychological well-being: An empirical study using the Chinese version of the purpose in life questionnaire. *The Journal of Genetic Psychology*, 153, 185–200.
- Shek, D. T. L. (1993). The Chinese purpose-in-life test and psychological well-being in Chinese college students. *International Forum for Logotherapy*, 16, 35–42.
- Shek, D. T. L., Hong, E. W., & Cheung, M. Y. P. (1987). The purpose in life questionnaire in a Chinese context. *The Journal of Psychology*, 121, 77–83.
- Sink, C. A., Van Keppel, J., & Purcell, M. (1998). Reliability estimates of the purpose in life and seeking of noetic goals tests with rural and metropolitan area adolescents. *Perceptual and Motor Skills*, 86, 362.
- Smith, G. T., & McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment*, 7, 300–308. doi:[10.1037/1040-3590.7.3.300](https://doi.org/10.1037/1040-3590.7.3.300).
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12, 102–111. doi:[10.1037/1040-3590.12.1.102](https://doi.org/10.1037/1040-3590.12.1.102).
- Starck, P. L. (1983). Patients' perceptions of the meaning of suffering. *International Forum for Logotherapy*, 6, 110–116.
- Starck, P. L. (1985). *Guidelines-Meaning in Suffering Test*. Abilene, TX: Viktor Frankl Institute of Logotherapy.
- Steger, M. F. (2006). An illustration of issues in factor extraction and identification of dimensionality in psychological assessment data. *Journal of Personality Assessment*, 86, 263–272. doi:[10.1207/s15327752jpa8603_03](https://doi.org/10.1207/s15327752jpa8603_03).

- Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, 53, 80–93. doi:[10.1037/0022-0167.53.1.80](https://doi.org/10.1037/0022-0167.53.1.80).
- Steger, M. F., & Kashdan, T. B. (2007). Stability and specificity of meaning in life and life satisfaction over one year. *Journal of Happiness Studies*, 8, 161–179. doi:[10.1007/s10902-006-9011-8](https://doi.org/10.1007/s10902-006-9011-8).
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173–180. doi:[10.1207/s15327906mbr2502_4](https://doi.org/10.1207/s15327906mbr2502_4).
- Strack, K. M., & Schulenberg, S. E. (2008). *The meaning in life questionnaire: A psychometric study with individuals with serious mental illness* (submitted).
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis*. Washington, DC: American Psychological Association.
- Vogt, W. P. (2005). *Dictionary of statistics & methodology: A nontechnical guide for the social sciences* (3rd ed.). Thousand Oaks, CA: Sage.
- Waisberg, J. L., & Starr, M. W. (1999). Psychometric properties of the purpose in life test with a sample of substance abusers. *International Forum for Logotherapy*, 22, 22–26.
- Walters, L. H., & Klein, A. E. (1980). A cross-validated investigation of the Crumbaugh purpose-in-life test. *Educational and Psychological Measurement*, 40, 1065–1071. doi:[10.1177/001316448004000434](https://doi.org/10.1177/001316448004000434).
- Wei, M., Russell, D. W., Mallinckrodt, B., & Vogel, D. L. (2007). The experiences in close relationship scale (ECR)-short form: Reliability, validity, and factor structure. *Journal of Personality Assessment*, 88, 187–204.
- Wong, P. T. P., & Fry, P. S. (1998). *The human quest for meaning: A handbook of psychological research and clinical applications*. Mahwah, NJ: Erlbaum.
- Yalom, I. D. (1980). *Existential psychotherapy*. New York, NY: Basic Books.
- Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99, 432–442. doi:[10.1037/0033-2909.99.3.432](https://doi.org/10.1037/0033-2909.99.3.432).

Copyright of Journal of Happiness Studies is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.