Factor Validity and Reliability of the Revised Self-Leadership Questionnaire in a Greek sample

Proios Ioannis

Abstract

Purpose of the present study was to investigate the Revised Self-Leadership Questionnaire (RSLQ; Houghton & Neck, 2002) for the Greek population in physical activities contexts. 373 undergraduate students from a Greek university (Department Physical Education and Sport Science) participated in this study. The RSLQ factorial structure was accessed through exploratory and confirmatory factor analysis. The data failed to support the full nine-factor model 35-item. In contrast, a new factor model 8-factor 25-item was appeared to show evidence for being statistically significant and supports the scale's construct validity.

Keywords: self-leadership, factor structure, reliability, physical activities

1. Introduction

The measurement of persons’ socio-behavioral characteristics, that is, skills that can lead to better self-management of behaviors, is a rather attractive subject to researchers. Such an interest is due to the belief that knowledge of persons’ skills, such as self-leadership, can contribute to the design and implementation of programs aimed at the cover needs for self-direction and self-motivation to shape their behaviors in positive ways to enhance their overall performance (Houghton, Dawley, &DiLillo, 2012).

Self-leadership has been described as a process by which a person controls his/her own behaviors, creates influence and leads oneself using specific behavioral and cognitive strategies (Manz, 1986, 2015; Manz & Neck, 2004; Manz & Sims, 2001). Self-leadership strategies include three different dimensions: behavior-focused strategies, natural reward strategies and constructive thought pattern strategies (Anderson & Prussia, 1997; Houghton & Neck, 2002; Manz & Neck, 2004; Manz & Sims, 2001; Neck & Houghton, 2006; Prussia, Anderson, & Manz, 1998). The self-leadership construct has been measured by means of the Revised Self-Leadership Questionnaire (RSLQ; Houghton & Neck, 2002). The RSLQ consists 35 items in 9 subscales, representing three self-leadership dimensions: the behavior-focused which consists five subscales (self-goal setting; self-reward; self-punishment; self-observation and self-cueing), the natural reward which consists only one subscale, and the constructive thought which consists three subscales (visualizing successful performance; self-talk; and evaluating beliefs and assumptions. The RSLQ has indicated relative a good reliability and construct validity in several studies (e.g., Carmeli, Meitar, & Weisberg, 2006; Curral & Marques-Quinteiro, 2009; Houghton, Bonham, Neck, & Singh, 2004; Houghton & Jinkerson, 2007).

The RSLQ, except for its use in English language (e.g., Houghton & Neck, 2002; USA, and Mahembe, Engelbrecht, & De Kock, 2013, and Nel & Zyl, 2015; South Africa), it has been translated in another languages, as Chinese (Ho & Nesbit, 2009; Neubert & Wu, 2006), Portuguese (Curral & Marques-Quinteiro, 2009; Marques-Quinteiro, Curral, & Passos, 2012), Turkish (Dogan & Sahin, 2008), Hebrew (Carmeli et al., 2006), and German (Andressen & Konradt, 2007). The translated versions of the RSLQ further confirm the original findings by Houghton and Neck (2002) and give additional evidence of the measure’s cross-cultural validity (Houghton et al., 2012). Despite the fact that RSLQ has been used enough in self-leadership research, is support that more research is needed to assess the measurement properties of the RSLQ (Houghton & Neck, 2002), especially when it is applied in new cultural contexts (Houghton et al., 2012).

1 Department of Physical Education and Sport Science, Aristotle University of Thessaloniki, Greece.
E-mail: johnyproios@gmail.com
The significance of the current investigation lies in the examination of the factor structure of RSLQ in the Greek culture. Moreover, an examination of the RSLQ measure with that sample can contribute to the research on self-leadership skills in a different achievement domain, such as physical activities. To date, self-leadership has been examined in concepts, such as spirituality in the workplace (Neck & Milliman, 1994), performance appraisals (Neck, Stewart, & Manz, 1995), organizational change (Neck, 1996), entrepreneurship (Neck, Neck, & Manz, 1997), diversity management (Neck, Smith, & Godwin, 1997), job satisfaction (Houghton & Jinkerson, 2007; Robert & Foti, 1998), non-profit management (Neck, Ashcraft, & Vansandt, 1998), goal setting/goal performance (Godwin, Neck, & Houghton, 1999; Neck, Nouri, & Godwin, 2003), team performance (Stewart & Barrick, 2000), team sustainability (Houghton, Neck, & Manz, 2003), succession planning (Hardy, 2004) and ethics (Vansandt & Neck, 2003). Finally, this instrument can help to identify individuals with poor self-leadership skills in physical activities contexts. By doing so, support can be provided to such individuals to improve their self-influencing (self-leadership) skills which, in turn, could help them to perform more effectively. Hence, the purpose of the present study is to test the factor structure of the RLSQ instrument in a Greek sample.

2. Method

2.1 Participants

The study used undergraduate students from a Greek university (Department Physical Education and Sport Science), 385 questionnaires were distributed to participants and 373 completed questionnaires were returned. The resulting response rate was very high (96.9%) because research questionnaires were administered in class, with voluntary participation. The sample consisted of 197 male (52.8%) and 176 female (47.2%) participants. The majority (95%) fell in the age category between 18 and 24 years.

2.2 Instrument

Self-leadership. Self-leadership was measured with the Revised Self-Leadership Questionnaire (RSLQ) development by Houghton and Neck (2002). There are 35 items statements in the self-leadership scale measuring behavior-focused, natural reward, and cognitive thought strategies. Self-leadership strategies were scored using a 5-point Likert scale ranging from 1 = “totally disagree” to 5 = “totally agree”. There are nine subscales in the RSLQ. Behavior-focused self-leadership can be measured with five subscales identified as self-goal setting (5 items; $\alpha = 0.84$), self-reward (3 items; $\alpha = 0.93$), self-punishment (4 items; $\alpha = 0.86$), self-observation (4 items; $\alpha = 0.82$), and self-cueing (2 items; $\alpha = 0.91$). Natural reward self-leadership is measured with a single 5-item scale ($\alpha = 0.74$). Constructive thought self-leadership is measured with three subscales, including visualizing successful performance (5 items; $\alpha = 0.85$), self-talk (3 items; $\alpha = 0.92$), and evaluating beliefs and assumptions (4 items; $\alpha = 0.78$).

2.3 Procedure

Prior institutional permission was granted before conducting the research. Data were collected from undergraduate physical education classes after permission granted by the competent Professors and after the participants were informed of the nature of the study. Participation was voluntary, and no incentives were provided.

2.4 Data analysis

Preliminary analyses for violations of statistical assumptions were conducted. Then factor analyses (confirmatory and exploratory) were used to investigate the factor structure of the items from the nine RSLQ scales described above. Both analyses were used because the theoretical foundation of self-leadership is built upon in many theories. Exploratory factor analysis (EFA) is generally thought of as more of a theory-generating procedure, while the confirmatory factor analysis (CFA) is a theory-testing model (Stevens, 1996). Three procedures were used. First, all 35 items were forced to a nine-factor solution and CFA was used to assess the fit of original model of nine-factor 35-item in Greek language. Second, exploratory factor analyses were used because the correlation between the observed and latent variables was found to be uncertain (Byrne, 2001). Specifically, EFA were used to determine the extent to which the item measurement (the observed variables) were related to the nine latent constructs. Next, CFA was used to tested fit of the 8-factor 25-item model, based on the results of the EFA. The models (9-factor, 32-item and 8-factor, 25-item) were tested using CFA with maximum likelihood (ML) parameter estimates in AMOS 22.0. The ML algorithm was selected because it resulted in accurate fit indexes with ordered-categorical data that violate the assumption of multivariate normality (Hutchinson & Olmos, 1998; Olsson, Foss, Troye, & Howell, 2000). The sample size in the present study was adequate to estimate the various models based on two criteria: (a) the total sample size was larger than 300, and (b) the ratio of the total sample size to the number of freely estimated parameters should be greater than 10:1 (Kline, 2005). Several criteria were used to test the hypothesized factor structure of the
The traditional measure of fit for CFAs models is the $x^2$ goodness-of-fit test (statistic). Non-significant values suggest a good fit, since they indicate only a minor discrepancy between the observed and the estimated covariance matrix. However, now it is generally agreed that the $x^2$ statistic is sensitive to sample size and violations of multivariate normality (Bollen& Long, 1992). The goodness-of-fit index (GFI) represents the relative amount of variances reproduced by the specified model as compared to the saturated model. Fit indices exceeding .90 are considered evidence of adequate model fit (e.g., Bollen, 1989; Fan, Thompson, & Wang, 1999), although Hu and Bentler (1999) have proposed more conservative criteria, in which GFI at .95 or higher is deemed a good model fit. The comparative fit index (CFI) is an index in which values greater than .90 may indicate reasonably good fit of the data to the model (Hu &Bentler, 1999). Finally, the root mean square error of approximation (RMSEA) is an index least affected by the sample size (Fan et al., 1999). The RMSEA ≤ .05 indicates close approximate fit; values between .05 and .08 suggest reasonable error of approximation and RMSEA ≥ .10 suggests poor fit (Brown &Cudeck, 1993). These fit indices were selected as the best indices, acknowledging that there is not a single fit index that is considered to be the definitive marker of optimum model fit (Hu &Bentler, 1999; Byrne, 2001). To EFA two criteria were used to determine the factor structure: (a) retain items with a factor loading equal to or greater than .40, and (b) exclude items with double loadings.

3. Results

Preliminary analyses

Prior to analyses, data were examined for the presence of univariate and multivariate outliers (Tabachnick&Fidell, 2007), which might attenuate the results. The former was analyzed through standardized scores ($z ≥ 3.30$) and the latter thought Mahalanobis Distance ($p<.001$) and Studentized Deleted Residual (greater than ±4.00). Few univariate outliers were identified as having high $z$ scores while no multivariate outliers were found. Also, the normality of each of the 35 items was investigated in terms of its skewness (-.10 to 0.23, $M = 0.85$) and kurtosis (-.97 to 2.25, $M = 1.52$). These values were all within the level recommended for a CFA with maximum-likelihood estimation (skew > 2, Kurtosis > 7; West, Finch, & Curran, 1995), supporting the normality assumption for all variables. For this, no cases were deleted.

Factor analyses

Initially, a confirmatory factor analysis was applied because the RSLQ was developed based on theoretical grounds, specific hypotheses regarding which items should load significantly on which scale were suitable for confirmatory factor analysis (Stevens, 1996). Based on content validity the factorial structure of the RSLQ was tested making use a confirmatory factor analysis (9-factor correlated model) (Table 1). Results from the CFA analysis indicated lack-of-fit of the hypothesized model (chi-square = 1176.44, df = 526, $p<.001$, CFI = 834, GFI = .844, RMSEA = .058).

The next step in the analysis was to perform exploratory factor analysis (EFA) because the original factor structure could not be confirmed. The factor structure of the preliminary 35-item RSLQ was identified using EFA with principal component analysis as the extraction method. The Varimax with Kaiser Normalization method was used for rotation because the factors were thought to be uncorrelated. The Kaiser–Meyer–Olkin’s measure (KMO) of sampling adequacy (acceptable level >.50; Kaiser, 1970) and Bartlett’s test of sphericity (Bartlett, 1950) were calculated to verify the appropriateness of an EFA. The number of factors was determined using parallel analysis (PA) performed with SPSS syntax developed by O’Connor (2000). Only items with a strong loading (.40 or higher) on one factor and <.30 on any other factors were retained to form latent variables because large loadings on factors other than the primary factor could result in serious flaws in the factor structure (Costello & Osborne, 2005). However, these views vary slight within literature. For example, Tabachnick&Fidell (2007) suggest that the secondary loading (or cross-loading) should be no greater than .32. Costello and Osborne (2005) suggest that a loading of .50 is enough to be considered “strong,” while Guadagnoli and Velicer (1988) state that the loading should be .60 or greater.

The EFA with data from the 35-item RSLQ indicated that the sample and correlation matrix were appropriate for factor analysis (KMO Index = .800 and Bartlett’s test of sphericity was significant $x^2(595, N = 373) = 4367.10, p<.001$). Ten factors were extracted based on PA analysis that cumulatively accounted for 61.74% of the total variance. However, several items were found to have a poor loading (<.40) on all factors or had crossed loadings (loading ≥.30 in two or more factors). These items were then dropped and a repeated EFA was conducted with remaining items using the same method. This procedure was replicated until all retained items met the criteria. The
excluded items were: Item 34: “I write specific goals for my own performance”; Item 8: “I focus written notes to remind myself of what I need to accomplish”; Item 17: “I try to surround myself with the objects and the people that bring out my desirable behaviors”; Item 26: “When I have a choice, I try to do my work in ways that I enjoy rather than just trying to get it over with”; Item 1: “I use my imagination to picture myself performing well on important tasks”; and Item 23: “I openly articulate and evaluate my own assumptions when I have a disagreement with someone else”. The sub dimension of “self-observation” was also discarded given all items saturated in more than one factor. Finally, a 25-item scale was generated with all items having a strong loading on the primary factor but not on the other factors (Table 2).

The appropriateness of the EFA for the revised 25-item RSLQ was again justified, KMO = .787; Bartlett’s test: $x^2(300, N = 373) = 3.279.21, p < .001$. The number of factors was change on a repeated PA (Table 1). A clear 8-factor solution emerged. The initial eigenvalues of eight factors were 5.52, 2.32, 2.07, 1.82, 1.47, 1.34, 1.16, and 1.04. These latent variables explained 22.07%, 9.27%, 8.26%, 7.28%, 5.89%, 5.35%, 4.64%, and 4.17% of variance, respectively, and together 66.93% of the total variance. After carefully examining of the items, it was found that all items of the self-observation factor were excluded.

To examine the stability of the 8-factor solution derived from exploratory factor analyses, we conducted a CFA on the 25-item modified RLSQ using the same sample. As model postulated an 8-factor 25-item correlated latent factors structure, in which each item was constrained to load on its respective latent variable. Results indicated good model fit ($x^2 = 462.36, df = 247, p < .001, GFI = .91, CFI = .93$ and RMSEA = .048) (Table 1).

Table 1 Fit Indices for Models

<table>
<thead>
<tr>
<th>Model description</th>
<th>$x^2$</th>
<th>df</th>
<th>GFI</th>
<th>CFI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>9-factor</td>
<td>1176.44</td>
<td>526</td>
<td>.844</td>
<td>.834</td>
<td>.058</td>
</tr>
<tr>
<td>8-factor</td>
<td>432.36</td>
<td>247</td>
<td>.910</td>
<td>.930</td>
<td>.048</td>
</tr>
</tbody>
</table>

For the assessing of reliability internal consistency methods and Cronbach’s alpha were used (Anastasi & Ubrina, 1997; Cortina, 1993). Internal consistency method that used here was correlations with item-to-scale (average item-total correlation) (DeVellis, 2003). The results of the inter-total correlation indicated low to good relation among items (see Table 2). An average inter-item correlation of .30 or higher indicates acceptable reliability (Robinson, Shaver, & Wrightsman, 1991). Tabachnick and Fidell (2007) suggest that correlations exceeding .30 provide enough evidence to indicate that there is enough commonality to justify comprising factors. Values for an item-total correlation (point-biserial) between 0 and .19 may indicate that the question is not discriminating well, values between .20 and .39 indicate good discrimination and values .40 and above indicate very good discrimination. In item analysis, in order to protect the sumability aspect of the scale, it has to be higher than .30 (Kline, 1993), or at least .40 (Gliem & Gliem, 2003).

Table 2 Factor Loadings (IF), Communalities ($h^2$) and Reliability Characteristics ($\alpha$) and Item-total correlations (ITC) of the RSLQ-25

<table>
<thead>
<tr>
<th>Variable</th>
<th>IF</th>
<th>$h^2$</th>
<th>ITC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-Goal setting ($\alpha = .74$)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL02</td>
<td>.71</td>
<td>.55</td>
<td>.50</td>
</tr>
<tr>
<td>SL11</td>
<td>.70</td>
<td>.59</td>
<td>.56</td>
</tr>
<tr>
<td>SL20</td>
<td>.77</td>
<td>.64</td>
<td>.59</td>
</tr>
<tr>
<td>SL28</td>
<td>.57</td>
<td>.50</td>
<td>.49</td>
</tr>
<tr>
<td><strong>Self-Reward ($\alpha = .84$)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL04</td>
<td>.77</td>
<td>.65</td>
<td>.61</td>
</tr>
<tr>
<td>SL13</td>
<td>.88</td>
<td>.80</td>
<td>.73</td>
</tr>
<tr>
<td>SL22</td>
<td>.91</td>
<td>.85</td>
<td>.79</td>
</tr>
<tr>
<td><strong>Self-Punishment ($\alpha = .64$)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL06</td>
<td>.61</td>
<td>.54</td>
<td>.33</td>
</tr>
<tr>
<td>SL15</td>
<td>.76</td>
<td>.65</td>
<td>.49</td>
</tr>
<tr>
<td>SL24</td>
<td>.77</td>
<td>.60</td>
<td>.47</td>
</tr>
<tr>
<td>SL30</td>
<td>.58</td>
<td>.48</td>
<td>.40</td>
</tr>
<tr>
<td><strong>Self-Cue ($\alpha = .78$)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the assessing of reliability internal consistency methods and Cronbach’s alpha were used (Anastasi & Ubrina, 1997; Cortina, 1993). Internal consistency method that used here was correlations with item-to-scale (average item-total correlation) (DeVellis, 2003). The results of the inter-total correlation indicated low to good relation among items (see Table 2). An average inter-item correlation of .30 or higher indicates acceptable reliability (Robinson, Shaver, & Wrightsman, 1991). Tabachnick and Fidell (2007) suggest that correlations exceeding .30 provide enough evidence to indicate that there is enough commonality to justify comprising factors. Values for an item-total correlation (point-biserial) between 0 and .19 may indicate that the question is not discriminating well, values between .20 and .39 indicate good discrimination and values .40 and above indicate very good discrimination. In item analysis, in order to protect the sumability aspect of the scale, it has to be higher than .30 (Kline, 1993), or at least .40 (Gliem & Gliem, 2003).
Cronbach’s index of internal consistency ranged from 0.62 to 0.88 (see Table 2), with five subscales showing acceptable alpha coefficient ($\alpha > 0.70$; Kline, 2005) and three subscales showing moderately low alpha coefficient ($0.61 < \alpha < 0.70$; Kline, 2005). However, since alpha coefficient is affected by the number of items (Cortina, 1993), it is maintained that when the number of items is small, the measure should be considered reliable (Schmitt, 1996). In the present study, the three factors which exhibited alpha values below the acceptable limit comprised four, three and two items, respectively. According to Ntoumanis (2001) and Pallant (2010) the values 0.62, 0.64 and 0.65 of factors present in this study, can be considered satisfactory, since these factors comprise less than ten items (four, three and two items, respectively).

4. Discussion

The present study was designed to translate into Greek and to examine the construct validity of the Greek version of the RSLQ. Results showed that a 25-item instrument had adequate psychometric properties for measuring eight dimensions of self-leadership using Greek sample that participated on physical activities. This instrument can represents a first step towards the development of an individual self-leadership measure in the Greek physical activity contexts and for that several recommendations for future research are presented.

The results the current study is not provided support for the hypothesized nine-factor 35-item of original version of the RSLQ (Houghton & Neck, 2002). Factorial analysis has suggested that several items should be reviewed, as factorial structure showed that several of them simultaneously had low scores in more than one factor. This was the case of self-observation. Similar finding has been reported by Marques-Quinteiro et al. (2012), who examined the factorial structure of RSLQ in Portuguese context. Results imply that may be some systematic differences in the way the members separate cultures respond to and interpret item on the RSLQ, and it seems likely that these differences may reflect cultural differences relative to the U.S. sample of original instrument (Houghton, Carnes, & Ellison, 2014). Results other studies were suggested the exist important differences in self-leadership across cultures (Georgianna, 2007; Neubert & Wu, 2006). Based on of the existing cross-cultural self-leadership research suggests, we do expect that cultural differences may be reflected in the operationalization and measurement of the construct (Houghton et al., 2014).

This study has some limitations. First, it should be noted that the assessment of self-leadership dimensions was based on self-reports. Second, the findings cannot be generalized to the broader population that participate in physical activities without further replication and another ages, because the sample that was used consisted of mostly young (18-24 years). It would be of interest to examine whether the present measure could be generalized to other ages, and in other Greek population working adults. A systematic future research is necessary.

4.1 Conclusions
In conclusion, the current study failed to support the full nine-factor 35-item model. The psychometric evaluation of new factor model of the RSLQGr indicates good reliability and factorial validity. Therefore, this study makes an important contribution in advancing physical activity self-leadership research. Overall, the results of this study are in line with international research regarding the psychometric properties of the RSLQ. The RSLQ-Gr may prove to be an effective self-leadership measure with much potential to facilitate empirical self-leadership research and practice.

5. References


Houghton, J. D., Neck, C. P., & Manz, C. C. (2003). We think we can, we think we can, we think we can: The impact of thinking patterns and self-efficacy on work team sustainability. Team Performance Management, 9, 31–41. https://doi.org/10.1108/13527590310537963


