Psihologia Resurselor Umane, **16** (2018), 40–50

http://dx.doi.org/10.24837/pru.2018.1.485

RESEARCH ARTICLE

Insights into the Criterion Validity of the Toxic Leadership Scale: A Bayesian Comparative Analysis Between Good and Poor Managers

STEFAN LIȚĂ
Psychological Services Section of the Romanian Gendarmerie,
Institute of Philosophy and Psychology of the Romanian Academy

Abstract

This paper presents data about the use of the Toxic Leadership Scale - TLS (Schmidt, 2008; Popa et al., 2013) in Romanian military organizations and analyzes its criterion validity. The study involved 243 participants from two military units with different organizational climate (normal vs. tense) who rated their leaders (Head and First Deputy) using the TLS (which consists in 30 items with five subscales and a global score). According to a Bayesian t-test for independent samples, we found compelling evidence that leaders of the high conflict unit received higher scores on the TLS than leaders of the low conflict unit, a fact that supports the criterion validity of the scale. The study provides additional evidence regarding the psychometric characteristics of the scale and indicates that it can be used in real organizational contexts for assessing senior management, even in high-stakes situations.

Keywords

toxic leadership, criterion validity, managers, military organizations, high-stake setting

Introduction

The topic of leadership has been extensively researched since the middle of the last century, and investigations have used different approaches and have focused on a variety of aspects of this phenomenon (*traits, behaviors, activities, relationships, interactions, processes, context, and consequences*). Most of the theory and research on leadership was concerned with what makes effective leaders and the underlying supposition was that ineffective leaders are those who fail to display the behaviors associated with effective leadership.

The research has tended to focus on positive rather than negative aspects and this trend helped us understand why some leaders are successful. However, far too little attention has been paid to negative facets and this line of questioning might help us understand not only what leaders are doing wrong, but also why people follow destructive bosses or corrupt politicians. During the past 25 years, much more information has become available on negative dimensions of leadership in organizations or societies and the purpose of this introduction is to present a short review of the literature concerning the dark side of leadership.

Correspondence regarding this article should be addressed to Stefan Liță, Ph.D., Institute of Philosophy and Psychology, Romanian Academy, Casa Academiei, Calea 13 Septembrie nr.13, Bucharest, Romania. E-mail: csl@ifpar.org

Negative approaches to leadership

Following the above mention logic, leadership has begun to be approached from a negative perspective, where accent has been put on the flaws and dysfunctions of leaders and managers from both public and private institutions. Thus, this literature generated notions such as "tyrannical", "abusive", "toxic" "narcissistic*"* or "destructive" leadership. For instance, Ashforth (1994) developed the concept of *petty tyranny* and identified six major characteristics of tyrants: arbitrariness and self-aggrandizement, belittlement of subordinates, lack of consideration, a forcing style of conflict resolution, discouragement of initiative, and noncontingent punishment. He also found that the effects of petty tyranny on employees included lower leader endorsement, higher frustration, stress, and reactance, greater helplessness and alienation.

Tepper (2000) described the concept of *abusive supervision* and defined it as workers' perceptions of their supervisors' sustained display of verbal and nonverbal hostile behaviors. He found that abusive supervision lowered subordinates' satisfaction and organizational commitment, and increased work-family conflict and psychological distress.

Lipman-Blumen (2005) has defined *toxic leaders* as those who engage in many destructive behaviors and who exhibit certain dysfunctional personal characteristics. To count as toxic, these behaviors and qualities of character must inflict some reasonably severe and enduring harm on their followers and their organizations. The intent to harm others or to enhance the self at the expense of others distinguishes severely toxic leaders from the careless or unintentional toxic leaders, who also cause adverse effects.

Rosenthal and Pittinsky (2006) discussed the concept of *narcissistic leadership* and suggested a series of features associated with ineffective leadership, such as arrogance, self-absorption, amorality, a lack of sensitivity to others, hypersensitivity and anger, irrationality, inflexibility, feelings of inferiority and hostility, need for recognition and superiority, and paranoia.

Einarsen, Aasland and Skogstad (2007) proposed the concept of *destructive leadership* and conceptualized it as the systematic and repeated behavior that violates the legitimate interest of the organization by undermining and/or sabotaging the organization's goals, tasks, resources and effectiveness, as well as the motivation, well-being or job satisfaction of employees. It has been associated with several workplace problems such as bullying, job dissatisfaction, job demands, work stressors, interpersonal issues and job insecurity.

Negative leadership has been framed under many labels and researchers talked about aversive, despotic and unethical leadership, derailed, harassing and intolerable bosses, or health-endangering leaders and unsupportive managerial behaviors. Among these ideas, the concept of toxic leadership has gained more popularity and became the more recognizable terms related to negative leadership – "a kind of global label for leaders who engage in numerous destructive behaviors and who exhibit certain dysfunctional personal characteristics" (Lipman-Blumen, 2005).

The problem of toxic leadership was explained in different ways from the dark triad of personality (Paulhus & Williams, 2002) to the toxic triad of leadership (Padilla, Hogan, Kaiser, 2007). While the dark triad emphasized only the negative characteristics of the leader, the toxic triad also highlighted the interaction between the destructive leader, submissive subordinates and favoring environment. On the one hand, Kellerman (2004) defined it as "bad leadership" resulting from negative personality traits and suggested that there are seven types of toxic leaders: incompetent, rigid, intemperate, callous, corrupt, insular, and evil. On the other hand, Price (2006) has explained toxic leadership as a mechanism of self-justification by which leaders see themselves in an exceptional situation in which moral limitations can be put on hold. While some of the characteristics are inherent to the personality of the leaders, others seem to come from extraordinary situations and circumstances in which the moral code of individuals can be adjusted to suit the moment, leading to unethical behaviors and poor results for the group.

Toxic leadership in military units

The toxic leadership was observed and investigated in military organizations (Popa, 2012). Even politicians (Secretary of the Army Thomas E. White) asked the U.S. Army War College to address how the Army could effectively assess leaders to detect those who might have "destructive leadership styles" (Reed, 2004). Since then, studies conducted in the army found that toxic commanders were considered to be: amoral, arrogant, cowardly, deceiving, dissatisfied, greedy, incompetent, insecure, irresponsible, unapproved, dysfunctional, intentional, narcissistic, egocentric, narrow-minded and unworthy of confidence (Doty & Fenlason, 2013; Williams, 2005). It has also been argued that toxic bosses engage in different dysfunctional behaviors to achieve what they want, such as self-promotion at the expense of subordinates, penalizing honest mistakes, restricting communication with subordinates, solving surface problems and losing time (Steele, 2011). Furthermore, a former CEO of the Center for Creative Leadership, Lieutenant General Walter F. Ulmer, concluded that toxic leaders are individuals whose behavior appears driven by self-centered careerism at the expense of subordinates and unit, and whose style is characterized by abusive and dictatorial behavior that promotes an unhealthy organizational climate (Ulmer, 2012).

Some researchers tried to group the characteristics, behaviors and attitudes of toxic leaders into categories or dimensions. For example, Reed (2004) described three major elements: disinterest in the well-being of subordinates, personality characteristics and behaviors that affect the organizational climate and the subordinates' belief that the primordial motivation of the leader is his interest. Also, Steele (2011) proposed four significant categories that toxic leaders exemplify: micromanagement, despicable soul and aggressiveness, narrow mind / rigid mind, poor attitude / bad example.

Other studies have examined the incidence of toxic leadership in the army. For instance, the presence of toxic leaders has been estimated between 5% and 20% according to the Annual Survey of Army Leadership (Riley et al., 2013). Moreover, it was stated that more than half of the military personnel experienced a toxic leader and 57% of the military personnel considered leaving their service due to their treatment by a senior officer (Reed & Bullis, 2009; Walter, 2012).

Using two samples of US Marines and Navy officer candidates, Schmidt (2008) developed the Toxic Leadership Scale (TLS) which has become the most used instrument for measuring toxic management. Schmidt (2008) identified negative relationships between TLS scores and satisfaction aspects: satisfaction with the workplace, satisfaction with colleagues, satisfaction with the supervisor and satisfaction with salary. Also, Schmidt (2014) found that TLS scores were negatively correlated with group cohesion, professional satisfaction, productivity, organizational confidence and organizational commitment.

Similar results have been obtained by other authors. Gallus and colleagues (2013) have found a negative relationship between toxic leadership and professional satisfaction, organizational commitment and politeness within the unit. Also, Truhon and McDonald (2013) pointed out that TLS is correlated in the expected direction with more measurements of equal opportunities: sexual harassment and discrimination, differentiated behavior towards women and minorities, positive behavior of equal opportunities, racist behavior, age discrimination, religious discrimination and discrimination against people with disabilities.

Eventually, the US Army recognized the impact of such behavior and incorporated a description of toxic leadership even in their formal doctrine. Therefore, toxic leadership has been described (Lindsay, Watola, & Lovelace, 2016) as a "combination of self-centered attitudes, motivations, and behaviors that have adverse effects on subordinates, the organization, and mission performance. This leader lacks concern for others and the climate of the organization, which leads to short- and long-term negative effects. The toxic leader operates with an inflated sense of self-worth and from acute self-interest. Toxic leaders consistently use dysfunctional behaviors to deceive, intimidate, coerce, or unfairly punish others to get what they want for themselves.

The negative leader completes short-term requirements by operating at the bottom of the continuum of commitment, where followers respond to the positional power of their leader to fulfill requests. This may achieve results in the short term, but ignores the other leader competency categories of leads and develops. Prolonged use of negative leadership to influence followers undermines the followers' will, initiative, and potential and destroys unit morale"

The Present Study

TLS was used more to evaluate former managers or leaders, not actual chiefs or heads. For example, the adaptation study of the TLS scale for Romanian language included the requirement to respond having as reference "*the most difficult leader they ever had*" (Popa et al., 2013). In this context, it might be important to investigate how TLS operates in real settings when employees have to evaluate their current leader.

Taking into account the results obtained in other countries, the purpose of this study is to present data on the use of the Toxic Leadership Scale (Schmidt, 2008; Popa et al., 2013) in Romanian military organizations and to analyze its validity regarding external criterion. The analysis continues the series of organizational studies in a military context, trying to integrate research directions: one centered on the positive aspects of leadership, such as managerial competences (Liţă, 2016) and leadership efficiency (Liţă, 2014a), and the other one focused on the negative aspects of leaders and employees, such as dysfunctional behavior (Liţă, 2017), interpersonal deviance (Liţă, 2014b) or personal weaknesses which affect the police chiefs' leadership styles (Liţă & Grigoraş, 2007).

Given the limitations of the few studies which have been reported on the use of TLS in real contexts, the present study contributes to the literature by (1) collecting data in high stake contexts, (2) investigating the psychometric properties in organizational assessment, and (3) presenting empirical evidence regarding the criterion-related validity.

Considering the general evidence reviewed above, we advance the general hypothesis that the Toxic Leadership Scale can detect dysfunctional leaders. If this hypothesis is correct, we expect to notice the following two facts: (a) the analysis between units will prove that leaders of the high conflict unit obtain higher scores at TLS than leaders of the low conflict unit, (b) the analysis within units will indicate the scores of the managers at TLS are quite similar.

Method

Procedure

To establish the difference between the good and the poor managers we used two general indicators.

First, both units have been subject to an organizational audit. In the first unit, this verification was planned in the previous year within the annual framework of periodical control and there were no warning signs of organizational or individual maladjustment. In the second unit, this verification was decided as a result of different written complaints made by the deputy head against the head of the unit, focused on dysfunctional behavior. Therefore, organizational resources, regarding time, people and energy, had been additionally spent to investigate several issues which might suggest that the second unit has more problems than the first one.

Second, the employees of both units have completed a questionnaire regarding the organizational climate and managerial efficiency. The psychometric properties of this questionnaire have been previously analyzed and presented at a scientific conference (Liţă, Scrieciu, & Ţoi, 2014) and in a scientific publication (Liţă, Scrieciu, & Ţintea, 2018). The reliability coefficients are between 0.72 and 0.94 for the scales, and the results can be grouped in three general indicators of psychosocial climate (*activity, relationship and the global index*) and three general indicators of managerial efficiency (*task oriented, people oriented and global index*). Based on these results, the units might be classified in five situations (bad, critical, vulnerable, good and very good). The

outcomes of these questionnaires indicated qualitative differences between the two units:

- The dimensions of organizational climate were perceived as Good and Very Good in unit 1 and as Critical or Bad in unit 2.
- The dimensions of managerial efficiency were perceived as Good and Very Good in unit 1 and as Vulnerable or Critical in unit 2.

Participants

The data collection was carried out in a high stake situation, a psychosocial survey that was part of an organizational audit. The sample consisted of 243 participants from two military units. In the first unit, the situation was normal and there was no information about possible conflicts. The Head of unit and First Deputy were evaluated by 163 employees. In the second unit, the situation was tense, with open conflict between the head of unit and the first deputy. Both managers have been evaluated by 80 employees.

Measures

The Romanian version of the Toxic Leadership Scale, adapted by Popa and colleagues (2013), was used experimentally. The scale contains 30 Likert-type items measured on 6 levels and evaluates five dimensions: self-promotion, abusive leadership, unpredictability, narcissism, authoritarian leadership.

The employees anonymously filled in the TLS and evaluate their managers in a high stake context of an organizational audit, knowing that the results will be presented to the managers of the unit and the higher level management team.

Data Analysis

The analyses were focused on the reliability, the factor structure and the criterion validity of the TLS. The five subclasses and the overall toxicity index were subjected to reliability analysis. The 30 items were included in a factor analysis with the *principal axis factoring* method and varimax orthogonal rotation technique, similar to the adaptation study (Popa et al., 2013). To evaluate the criterion validity, comparisons were made between the means obtained by the managers of the two units (normal situation vs. conflict situation).

To compute the statistical analysis, we used both SPSS and JASP software.

Results

Preliminary analyses

Table 1 presents a descriptive summary of the variables included in the study. It can be noticed that the means of TLS obtained from the evaluation of the two heads in the two units are relatively small, the highest scores being in narcissism and authoritarian leadership. In this context, it should be emphasized that the low figures are quite normal because the employees evaluated their own managers in a high stake context, represented by the realization of an organizational audit.

To test the assumption of normality, we choose the Shapiro-Wilk because a recent Monte Carlo simulation, which compares the Shapiro-Wilk, Anderson-Darling, Kolmogorov-Smirnov and Lilliefors tests, has found that it has the best power for a given significance (Razali & Wah, 2011).

Table 1 shows the entire variables have a deviation from the assumption of normality according to Shapiro–Wilk Normality Test. The null hypothesis of this test is that the population is normally distributed, therefore if the *p*-value is less than the .05 alpha level, then the null hypothesis is rejected and there is evidence that the data tested are not coming from a normally distributed population (Shapiro & *Wilk*, 1965).

Table 1. *Descriptive statistics for Toxic Leadership Scale*

Toxic Leadership Scale	Unit 1 (N=326)				Unit 2 (N=160)			
	Statistics		Test of Norm.		Statistics		Test of Norm.	
	M	SD	SW	Sig	M	SD	SW	Sig
Self-promotion	**1,63**	*0,65*	0,85	.001	**2,11**	*0,94*	0,92	.001
Abusive leadership	**1,96**	*0,82*	0,92	.001	**2,47**	*1,15*	0,94	.001
Unpredictability	**1,93**	*0,90*	0,89	.001	**2,55**	*1,23*	0,92	.001
Narcissism	**1,91**	*0,81*	0,91	.001	**2,22**	*0,97*	0,92	.001
Authoritarian leadership	**1,75**	*0,72*	0,88	.001	**2,13**	*0,96*	0,92	.001
Toxic Leadership Index	**1,84**	*0,72*	0,91	.001	**2,30**	*0,98*	0,93	.001

The reliability of the Toxic Leadership Scale

For the calculation of internal consistency, we used the average intra-class correlation coefficient, equivalent to the Cronbach alpha coefficient, which also offers the confidence interval limits. As can be seen in Table 2, the internal consistency indices obtained are high, similar to those calculated on the Romanian adaptation sample (Popa et al., 2013) and the original American sample (Schmidt, 2008).

Table 2. *Cronbach alpha coefficients for the Toxic Leadership Scale*

Scales	No. items	Military sample (N=486)		Romanian sample (N=352)		American sample (N=216)	
		Alfa	*95%CI*	*Alfa*	*95%CI*	*Alfa*	*95%CI*
Self-promotion	5	**.80**	.77-.83	**.82**	.78-.84	**.91**	.88-.92
Abusive leadership	7	**.86**	.84-.88	**.88**	.85-.89	**.93**	.91-.94
Unpredictability	7	**.86**	.84-.88	**.89**	.87-.90	**.92**	.90-.93
Narcissism	5	**.87**	.86-.89	**.90**	.88-.91	**.88**	.85-.90
Authoritarian leadership	6	**.86**	.84-.88	**.88**	.86-.90	**.89**	.86-.91
Toxic Leadership Index	30	**.97**	.96-.97	**.95**	94-.95	**.97**	.96-.97

The factor structure of the Toxic Leadership Scale

All 30 items of the Romanian version of the TLS scale were included in an exploratory factor analysis with the *principal axis factoring* method and varimax orthogonal rotation technique, similar to the adaptation study (Popa et al., 2013). On both batches, a four-factor solution was obtained covering 57.22%, respectively 62.58% of the total variation of the responses. Although the outcome seems similar to that obtained in the adaptation study on the Romanian population (Popa et al., 2013), the composition of the factors is a little bit different and difficult to interpret at present.

The criterion validity of the Toxic Leadership Scale

To evaluate how the scale was operating in significantly different organizational contexts, the results of the managers from the first unit (*where the situation was normal*) were compared with those of the managers from the second unit (*where the situation was tense*).

If we compute the t test for independent samples the results showed the TLS means in the first unit are smaller than the averages obtained in the second unit and these differences are statistically significant in both the comparison between the managers of those units and in the overall comparison of unit scores. The value of Cohen's d ranged from 0.4 to 0.6 for the difference between the heads of units and from 0.2 to 0.6 for the difference between the deputy heads which shows a small to medium effect size.

However, taking into account that the scores of TLS on both samples present significant deviation from the normal distribution, I assumed that Bayesian inference would be more appropriate to test the difference between the two leaders. Bayesian inference is particularly useful when we do not have as much data as we would like and want to update the probability for a hypothesis as more evidence or information becomes available. During the last years, using Bayes factors has become a concrete and practical alternative to hypothesis testing than using *p* values (Wagenmakers et al., 2017). The Bayes factor quantifies the evidence in the data for the hypotheses under investigation. For instance, if $BF_{01} = 10$, the support in the data is 10 times larger for H0 than for H1. Similar rules hold for evidence in favor of H1 and if $BF_{10} = 10$, the support in the data is 10 times larger for H1 than for H0.

According to Jeffreys (1961), Bayes Factor could be interpreted as substantial (0.5-1), strong (1-2) and decisive (2<). According to guidelines presented by Kass and Raftery (1995), a Bayes factor in the range 1–3 constitutes anecdotal evidence in favor of H0 or H1, 3–20 represents positive evidence and 20–150 strong evidence.

Table 3 shows the Bayes Factors for the two types of managers compared on TLS means. In is essential to mention that Bayes Factor is display as BF_{-0} instead of BF_{10} because the alternative hypothesis specified that group 1 is less than group 2 (BF_{-0}) not only that there are significant differences between the groups (BF_{10}). Therefore, we can notice the values of Bayes Factor for the difference between the heads of unit are 2.05 for self-promotion, 8316 for abusive leadership, 1.6 for unpredictability, 25878 for narcissism, 403 for authoritarian leadership and 353 for the global TLS score. At the same time, if we compare the deputy heads the values of Bayes Factor are 32 for self-promotion, 3488 for abusive leadership, 20123 for unpredictability, 241 for narcissism, 505 for authoritarian leadership and 2009 for the global TLS score. These values prove that the alternative hypothesis which state that TLS mean in unit 1 is less than TLS mean in unit 2 have a much higher probability of being true compared to the null hypothesis. To be more precise regarding the TLS index, the support in the data is 353 times larger for H1 than for H0 for the heads of units ($BF_{-0} = 353.153$) and 2009 times larger for H1 than for H0 ($BF_{-0} = 2009.290$) for the deputy heads.

Table 3. *Bayesian Independent Samples T-Test between the leaders of the two units*

Toxic Leadership Scales	**Dif. between Heads**		**Dif. between Deputy Heads**	
	BF_{-0}	**error %**	**BF_{-0}**	**error %**
Self-promotion	2.053	~ 2.205e -5	32.781	~ 3.536e -5
Abusive leadership	8316.833	~ 7.027e -7	3488.608	~ 8.373e -7
Unpredictability	1.681	~ 6.434e -5	20123.417	~ 3.668e -7
Narcissism	25878.302	NaN	241.759	~ 2.342e -5
Authoritarian leadership	403.887	~ 7.235e -6	505.360	~ 4.139e -6
Toxic Leadership Index	353.153	~ 1.031e -5	2009.290	~ 1.768e -6

Note. For all tests, the alternative hypothesis specifies that group 1 is less than group 2

To have a better picture of this analysis, a sequential analysis was also conducted (Figure 1) and the outcome proves that evidence for Hypothesis 1 is extreme, therefore we could indeed claim that leaders of the high conflict unit obtain higher scores at the Toxic Leadership Scale than leaders of the low conflict unit.

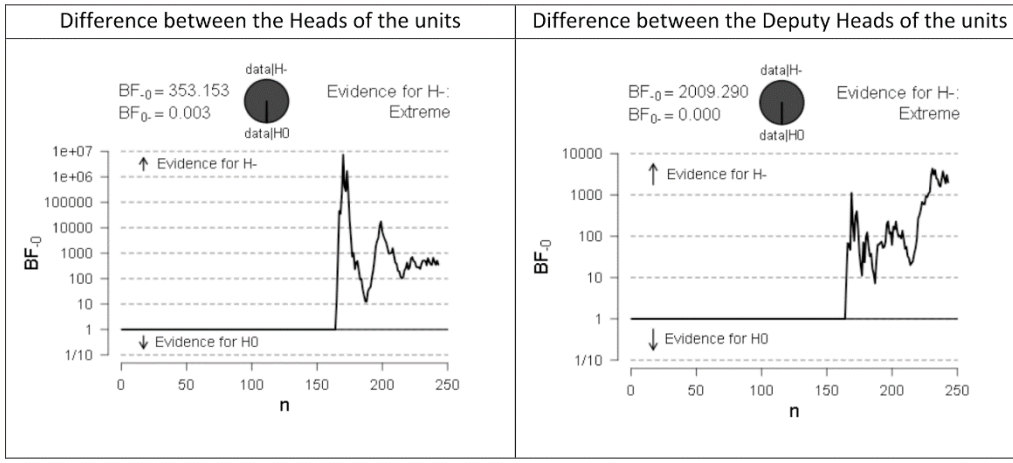

Figure 1. *Sequential analysis for testing the difference between the leaders of the two units*

The research hypothesis also assumed that the comparative analysis between managers within the same unit would produce quite similar scores. Table 4 shows the Bayes Factors in favor of H0 for the two types of units. The values of Bayes Factor for the difference between the managers within the first unit are between 4.8 and 8.1, while the Bayes Factors for the difference between the managers within the second unit are between 1.7 and 5.2. These values prove that null hypothesis which states that there are no significant differences between the mean of managers within the same unit have a higher probability of being correct compared with the alternative hypothesis. To say it differently, the support for the data is almost 7 times larger for H0 than for H1 in the first unit (BF_{01} = 6.8) and nearly 6 times larger for H0 than for H1 (BF_{01} = 5.8) in the second unit.

Table 4. *Bayesian Independent Samples T-Test between the leaders of the same unit*

Toxic Leadership Scales	**Dif. within Unit 1**		**Dif. within Unit 2**	
	BF_{01}	**error %**	**BF_{01}**	**error %**
Self-promotion	8.171	4.515e -5	5.276	9.316e -6
Abusive leadership	4.259	2.155e -5	4.211	7.671e -6
Unpredictability	7.746	4.252e -5	2.673	5.033e -6
Narcissism	8.086	4.462e -5	4.556	8.211e -6
Authoritarian leadership	4.842	2.496e -5	1.702	3.176e -6
Toxic Leadership Index	6.841	3.696e -5	5.798	1.006e -5

In Figure 2 a sequential analysis is also displayed which proves the evidence for the Null Hypothesis is moderate, therefore we could indeed state there are no significant differences shown by the Toxic Leadership Scale between the leaders of the same type of organizations.

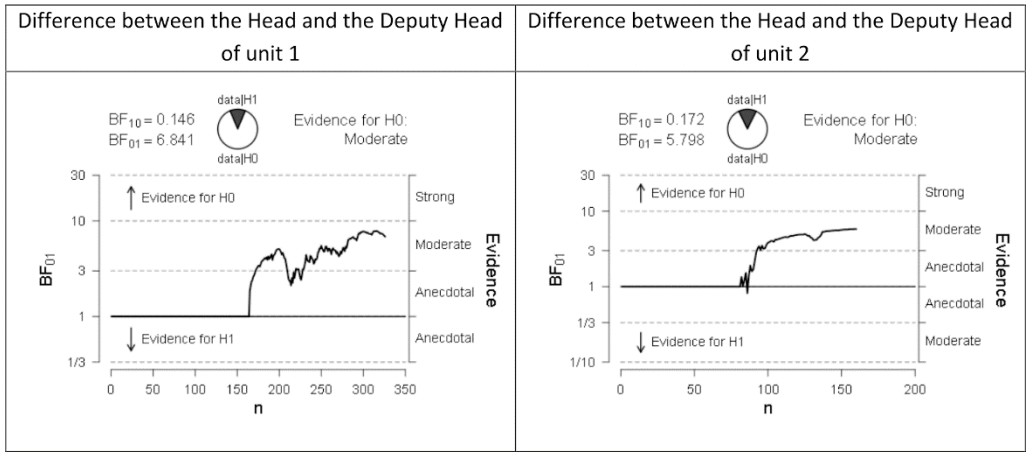

Figure 2. *Sequential analysis for testing the difference between the leaders within the same unit*

Discussion

Prior studies have noted the importance of studying negative features of leadership in general and toxic leadership in particular. However, our literature review of the literature did not reveal valid and reliable instruments for toxic leadership evaluation, especially in high stake contexts. Our study aimed at exploring the extent to which the Toxic Leadership Scale allows us to detect differences between effective and ineffective managers or the criterion validity of TLS. Two indicators were used to classify managers in good or poor: the first was related with warning signs of maladjustment regarding written complaints against them, and the second was related with perceived organizational climate and managerial efficiency. Therefore, the Toxic Leadership Scale was used to assess the managers of two military units that were significantly different in terms of institutional context (*normal versus tense*). Unlike other studies in which participants filled out TLS thinking of "*the most difficult leader they ever had*", in this study the employees have assessed their current managers and responded to the TLS in a delicate context of organizational audit.

The outcome of this study indicates that the managers in problematic organization received higher scores on TLS than the managers from the non-problematic unit which support the idea that TLS can discriminate between good and poor managers.

Main conclusions

The results of this study showed that:

a) The TLS mean scores for the military units leaders are relatively small, ranging from 1.79 to 2.14, but this is normal when it comes to assessing a boss who is in office. Although there are no norms for Romanian population regarding the TLS, we considered these results being small in rapport with the possible range of answers which is between 1 and 6, thus far lower than the theoretical average of the scale.

b) The TLS shows very good reliability (.80-.97), similar to the levels reported by other samples in Romania and America.

c) Regarding the factor structure of the TLS scale, our results indicated a factorial solution with 4 factors that is different from the solution proposed by the American author (5 factors), but similar with the one suggested by the authors who adapted the instrument to the Romanian population (4 factors).

d) Regarding the criterion validity, Bayesian t-test for independent samples provided very strong

evidence that the managers of the conflict unit achieve significantly higher scores on the Toxic Leadership Scale than the managers of the conflict-free unit that supports the criterion validity of the scale. Moreover, no significant differences have been detected between the managers of the same unit.

The current findings support the conclusion that the Toxic Leadership Scale can be used in real organizational contexts for assessing senior management even in high-stakes situations. Our study provided additional evidence of the psychometric characteristics of the Toxic Leadership Scale and has also highlighted the problem of the scale factor structure.

Limitations and future investigations

Although the relevance of TLS in assessing dysfunctional management is clearly supported by the current findings, the most controversial issue might be related with the criterions used to classify units or managers. For example, some critics could claim that the conflict between two managers is not necessarily a problem of toxic leadership. However, it is noteworthy to mention that in a military unit the management is supported by a managerial team and if those leaders are not able to work together this means that they are not able to accomplish their main job requirements. Furthermore, if their conflict increases so much that it goes beyond their unit, then this became a pressing problem for the higher level management team and could also affect the image of that organization on either a local or national level.

On the other hand it is true that many other criteria might be used to classify managers as good and poor and more research on this topic needs to be undertaken in order to analyze the criterion validity of TLS in concrete and diversified organizational contexts.

It might be also useful that future investigations should focus on testing different alternative models in various samples to clarify the factor structure of the TLS.

Finally, reflecting on the impact of this study, we could add that further work is required to establish whether the military environment in general might have a toxic leadership problem. Military organizations have some characteristics which might be fit for the development of toxic leaders: a clear hierarchy, a lot of power at the discretion of assigned commanders and the requirement for the military personnel to follow orders (*a quite perfect toxic triad*). Besides, a toxic leader can be seen as very competent and effective when performing a specific objective or task because a toxic leader can force their will upon their followers to get short-term performance. Therefore, toxic leadership potential is not easy to detect and alleviate although the consequences of toxic leadership could be evaluated much easily.

References

Ashforth, B. (1994). Petty Tyranny in Organizations, *Human Relations, 47*(7), 755–778.

Doty, J. & Fenlason, J. (2013). Narcissism and Toxic Leaders, *Military Review*, *93*(1), 55–60.

Einarsen, S., Aasland, M. S., Skogstad A. (2007). Destructive Leadership Behavior: A Definition and Conceptual Model, *The Leadership Quarterly*, *18*(3), 207–216.

Kellerman, B. (2004). *Bad Leadership: What It Is, How It Happens, Why It Matters,* Boston: Harvard Business School Press.

Gallus J. A., Walsh, B. M., van Driel, M, Gouge, M. C., & Antolic, E. (2013). Intolerable Cruelty: A Multilevel Examination of the Impact of Toxic Leadership on U. S. Military Units and Members, *Military Psychology*, *25*(6), 588–601.

Jeffreys, H. (1961). *Theory of Probability* (3rd ed.). Oxford: Oxford University.

Kass, R. E. & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795.

Lindsay, D., Watola, D. J., Lovelace, J. (2016). Why Negative Leadership Matters: Implications for Leader Development in the Military. In D. Watola, & D. Woycheshin (eds.). *Negative Leadership: International Perspective*, Canadian Defence Academy Press.

Liță, Ș. (2014a). *Factors which affect the perception of managerial effectiveness in military units*. Romanian Police National Symposium of Psychology, București [In Romanian].

Liță, Ș. (2014b). *Organizational diagnosis of interpersonal deviance: development and validation of an assessment tool.* Post-Doctoral and Early Career Scholars Workshop organized by European Group for Organizational Studies, Rotterdam, 2014.

Liță, Ș. (2016). Essential competencies of managers from Romanian Gendarmerie. In V. Mărineanu (ed.).

Mental and occupational health in military settings, Ministry of Defence Publishing House, Bucharest, 69-79 [In Romanian].

Liță, Ș., & Grigoraș, M. (2007). The role of personality traits in police leaders' behavior. *Polish Journal of Applied Psychology*, *5*(1), Wroclaw University Publishing.

Liță, Ș., & Rusan, A. (2017). *Developing an Early Warning System for detecting personnel maladaptive behavior*. National Conference of Organizational Psychology, București. [In Romanian].

Liță, Ș., Scrieciu, D., & Țintea, G. (2018). Validation of a Psychosocial Climate Assessment Questionnaire in military units. *Journal of Psychology*, Romanian Academy Publishing House *64*(2), 87–101, [In Romanian].

Liță, Ș., Scrieciu, D., Țoi, M. (2014). *Measuring organizational climate in military units*, National Conference of Organizational Psychology, Iași [In Romanian].

Lipman-Blumen, J. (2005). The Allure of Toxic Leaders: Why We Follow Destructive Bosses and Corrupt Politicians – and How We Can Survive Them, Oxford, UK: Oxford University Press.

Padilla, A., Hogan, R., Kaiser, R. B. (2007). The Toxic Triangle: Destructive Leaders, Susceptible Followers, and Conductive Environments, *The Leadership Quarterly*, *18*, 176–194.

Paulhus, D. L. & Williams, K. M. (2002). The Dark Triad of Personality: Narcissism, Machiavellianism, and Psychopathy, *Journal of Research in Personality*, 36, 556–563.

Popa, M. (2012). *Psihologie militară*. Iași: Polirom.

Popa, M., Rotărescu, V., Sulea, C., & Albulescu, P. (2013). Scala de evaluare a conducerii toxice – adaptare românească, *Psihologia Resurselor Umane*, *11*(2), 57-70.

Price, T. (2006). *Understanding Ethical Failures in Leadership*, Cambridge, UK: Cambridge University Press.

Razali, Nornadiah; Wah, Yap Bee *(2011).* Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson-Darling tests, *Journal of Statistical Modeling and Analytics. 2(1), 21–33.*

Reed, G. E. (2004). Toxic Leadership, *Military Review*, *84*(4) 67–71.

Reed, G., & Bullis, R. C. (2009). The Impact of Destructive Leadership on Senior Military Officers and Civilians, *Armed Forces & Society*, *36*(1), 5–18.

Riley, R. Hatfield, J. Paddock, A.& Fallesen, J. J. (2013) *The 2012 Center for Army Leadership Annual Survey of Army Leadership (CASAL): Main Findings*, Leavenworth, KS: Center for Army Leadership.

Rosenthal, S.A., & Pittinsky, T. L. (2006). Narcissistic Leadership, *The Leadership Quarterly, 17*(6), 617–633.

Shapiro, S. S,. & *Wilk, M. B.* (1965). *An analysis of variance test for normality (complete samples). Biometrika, 52(3–4), 591–611.*

Schmidt, A. A. (2008). *Development and validation of the Toxic Leadership Scale*, Master's thesis, University of Maryland, College Park. (http://drum.lib.umd.edu/bitstream/1903/8176/1/umi-umd-5358.pdf).

Schmidt, A. A. (2014). An Examination of Toxic Leadership, Job Outcomes, and Military Deployment. Doctoral thesis. University of Maryland.

Steele, J. P. (2011). Antecedents and Consequences of Toxic Leadership in the U.S. Army: A Two Year Review and Recommended Solutions, Fort Leavenworth: Center for Army Leadership.

Tepper, B. J. (2000). Consequences of Abusive Supervision. *The Academy of Management Journal 43*(2), 178–190.

Truhon, S. A. & McDonald, D. P. (2013). *Toxic Leadership in the U.S. Military*, Presentation, Annual Convention of the International Military Testing Association, Seoul, October 28-31.

Ulmer, Jr.W.F. (2012). Toxic Leadership: What Are We Talking About? *Army*, 47–52.

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, A. J., Love, J., Selker, R., Gronau, Q. F., ˇSm´ıra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2017). Bayesian statistical inference for psychological science. Part I: Theoretical advantages and practical ramifications, *Psychonomic Bulletin & Review*, *25*, 35–57.

Walter, F. U. (2012). Toxic Leadership: What Are We Talking About? *Army*, 47–52.

Williams, D. F. (2005). *Toxic Leadership in the U.S. Army*, Master's thesis, U.S. Army War College.