The Dimensionality and Measurement of Destructive Instructor-Leadership

Paul T Balwant, PhD, MSc, BSc Management
Department of Management Studies, The University of the West Indies, St. Augustine, Trinidad
+1 (868)792 1600
paul.balwant@sta.uwi.edu

Kamal Birdi, PhD and BSc Psychology, MSc Occupational Psychology
Institute of Work Psychology, Management School, The University of Sheffield, Sheffield, UK
+44 (0)114 222 3288
k.birdi@sheffield.ac.uk

Ute Stephan, PhD Psychology
Aston Business School, Aston University, Birmingham, UK
+44(0)121 204 3183
u.stephan@aston.ac.uk

Corresponding author:

Paul T Balwant
Email: paul.balwant@sta.uwi.edu
Mailing address: Department of Management Studies, The University of the West Indies, St. Augustine, Trinidad.

The Dimensionality and Measurement of Destructive Instructor-Leadership

Abstract

The conceptualization of destructive leadership has received increasing attention in recent times. Accordingly, researchers have developed a theoretical model of destructive leadership that highlights two manifestations as follows: (1) leading followers towards goals that contradict the organization's interests and (2) the use of harmful methods in leading followers. The two manifestations of destructive leadership point to the concept being multidimensional. However, researchers rarely investigate the dimensionality of destructive leadership when measuring the concept in general and in instructor-student relationships. Moreover, the most prominent measure of destructive leadership fails to capture its two manifestations adequately. To address the apparent mismatch between the theory and measurement of destructive instructor-leadership, we enhance an existing measure of destructive leadership. Using a sample of 174 students from the UK, the findings indicated that the two manifestations of destructive instructor-leadership can be measured by 13 items, and was composed of three dimensions including, irresponsibility, victimization, and callous communication. These findings, along with limitations and suggestions for future research are discussed.

Keywords

Introduction

The most central theme of leadership is that of 'influence'. Yukl (2006) examined numerous leadership definitions in the literature and explained that leadership is a process of intentional influence over others to direct them towards a goal. This influence process has been investigated primarily from a positive perspective (Bligh, Kohles, Pearce, Justin, & Stovall, 2007). Specifically, researchers have focused on leaders using supportive behaviours towards their followers, e.g., socialized charisma, inspiration, support, authenticity, etc. However, researchers are becoming increasingly interested in the dark side of leadership, particularly destructive leadership (Schyns & Hansbrough, 2010). Destructive leadership has been investigated primarily in a corporate context (Schyns & Schilling, 2013), but such behaviours may also exist in instructor-student relationships, i.e., destructive instructor-leadership (Balwant, 2017). In the present paper, our research question is, 'What are the dimensions of destructive instructor-leadership and how can the multidimensional concept of destructive instructor-leadership be measured?'

Literature review

To date, research on destructive forms of leadership has focused primarily on its antecedents and outcomes (Mackey, Frieder, Brees, & Martinko, 2017; Schyns & Schilling, 2013). In contrast, few researchers have examined the dimensionality of destructive forms of leadership. For instance, Mitchell and Ambrose (2007) showed that abusive supervision may consist of two dimensions – active personal abuse (e.g., 'tells me my thoughts and feelings are stupid') and passive acts of abuse (e.g., 'doesn't give me credit for work requiring a lot of effort'). Mitchell and Ambrose (2007) then used only the active dimension for the purpose of

their study. Moreover, the passive dimension includes certain behaviours that are arguably

active, e.g., blaming the follower or invading the follower's privacy.

Goodyear, Crego, and Johnston (1992) also offered a dimensionality for unethical leader

behaviours by investigating psychologists leading students during their research. The authors

found that unethical leader behaviours consist of incompetence, inadequate supervision,

supervision abandonment, intrusion of values, abusive supervision, exploitive supervision, dual

relationships, encouragement to fraud, and authorship issues. Some of these dimensions infer a

multidimensional structure for destructive forms of leadership, e.g., intrusion of values, abusive

supervision, exploitive supervision, and encouragement to fraud. However, other dimensions

seem to indicate the absence of leadership, e.g., incompetence, inadequate supervision, and

supervision abandonment. Additionally, the purpose of Goodyear et al.'s (1992) study was to

investigate ethical issues in the instructor-student relationship rather than destructive leadership.

The findings from Mitchell and Ambrose (2007) and Goodyear et al. (1992) suggest that

destructive leadership may be a multidimensional construct. Yet, extant research ignores the

dimensionality issue, and instead assumes that destructive leadership is a single construct

(typically using labels such as abusive supervision, petty tyranny, undermining, etc.[1]) (Ashforth,

1994; Duffy, Ganster, & Pagon, 2002; Schyns & Schilling, 2013; Tepper, 2000). Therefore, a

research gap exists between the way in which destructive leadership is represented in extant

research and its potential multidimensional nature. In order to address this research gap on

---

[1] In the destructive leadership literature, a myriad of terms are used to describe destructive leadership. In addition to the examples provided here, other common terms in the literature include tyrannical leadership, unsupportive managerial behaviours, despotic leadership, toxic leadership, strategic bullying, and more (Krasikova, Green, & LeBreton, 2013). Krasikova et al. (2013) explain that the multiple terminologies used to describe destructive leadership are problematic for scientific communication and progress in empirical research. For this reason, we only use 'abusive supervision' when referring to studies that used that term. In the upcoming paragraphs, we highlight Krasikova et al.'s (2013) unified definition of destructive leadership that ties together the multiple terms under one construct.

destructive leadership dimensionality, the purpose of the present study is twofold: (1) to investigate the dimensionality of destructive leadership; and (2) to develop and validate a measure that is aligned with the potential dimensionality of destructive leadership. To do so, we investigate destructive leadership in a unique context – instructor-student relationships.

The dynamics in instructor-student relationships are sometimes examined via a leadership lens. Leadership in this instructor-student context is often referred to as instructor-leadership. Instructor-leadership can be defined as 'a process whereby instructors exert intentional influence over students to guide, structure, and facilitate classroom activities and relationships in a class' (Balwant, 2017, p. 2). A key point from this definition is that intent is necessary to characterise instructors' behaviours as leadership. For instance, an instructor who comes to a class and reads lecture slides ad verbatim, without any intentions to guide and influence students, is showing a lack of leadership (for a more extensive discussion of instructors as leaders, see Balwant, 2016). In the instructor-leadership literature, most research has emphasized supportive teaching behaviours in higher education (Baba & Ace, 1989; Balwant, 2016; Dawson, Messe, & Phillips, 1972). However, the classroom context is also conducive to destructive forms of leadership (for a review of the educational approach to destructive leadership, see Balwant, 2017).

To understand the potential dimensionality of destructive instructor-leadership, we must start with a theoretical model that explains the meaning of destructive leadership. Krasikova, Green, and LeBreton (2013) reviewed and integrated the extant state of destructive leadership research, and propose that destructive leadership is

> 'volitional behaviour by a leader that can harm or intends to harm a leader's organization and/or followers by (a) encouraging followers to pursue goals that contravene the legitimate interests of the organization and/or (b) employing a leadership style that involves the use of harmful methods of influence with followers, regardless of justifications for such behaviour' (p. 1310).

This definition highlights two manifestations of destructive leadership, which may offer insight into the potential dimensionality of destructive instructor-leadership. The first manifestation describes leading followers towards goals that contradict the organization's interests, e.g., encouragement to fraud. The second manifestation describes the use of harmful behaviours in leading followers towards goals, e.g., verbal abuse.

Following the two manifestations proposed by Krasikova et al. (2013), Balwant (2017) offered a multidimensional concept of destructive instructor-leadership. Balwant (2017) highlighted the differences and similarities between destructive instructor-leadership and other concepts examined by educational researchers, e.g., teacher misbehaviours (Kearney, Plax, Hays, & Ivey, 1991) and disconfirmation (Ellis, 2000). In so doing, Balwant (2017) conducted an inductive study of destructive instructor-leadership which shows that the concept may consist of callous communication, chaotic carelessness, and irresponsibility. Callous communication is the use of harmful verbal and nonverbal behaviours, and is directly related to the second manifestation of destructive leadership. Chaotic carelessness refers to leader behaviours that create disorder while purposely neglecting followers, and seems to capture both manifestations of destructive leadership. Irresponsibility means that a leader neglects duties and responsibilities in leading followers towards goals, and appears to be closely linked to the first manifestation of destructive leadership.

In light of the aforesaid recent developments in the destructive instructor-leadership literature, we remain critical of using the most prominent measure of destructive leadership – Tepper's (2000) Abusive Supervision Scale – to fully capture destructive instructor-leadership (Schyns & Schilling, 2013). Tepper's (2000) scale appears to focus entirely on the second manifestation of destructive leadership as proposed by Krasikova et al. (2013). Even for this

second manifestation of destructive leadership, Tepper's (2000) scale seems to ignore harmful

methods that are nonverbal as proposed by Balwant (2017). Given these shortcomings of

Tepper's (2000) scale, our aim is to build upon the findings of Balwant (2017) to augment

Tepper's (2000) scale so that the latter is better aligned with the conceptualization of destructive

leadership as proposed by Krasikova et al. (2013). Our aim here directly attends to a call by

Balwant (2017) to 'adapt and refine Tepper's (2000) instrument for the instructor-student

relationship' (p. 23).

To develop the measure of destructive instructor-leadership, we follow Hinkin's (1998)

conceptual framework and guide for the development of scales that is based upon established

psychometric principles in survey research. Hinkin's (1998) scale development process

comprises of six steps. In the present paper, we follow the first five of Hinkin's (1998) six-step

process[2]. First, <u>item generation</u> entails the construction of items to measure the construct. Here,

we use the Abusive Supervision Scale (Tepper, 2000) items as a starting point, and then add

items based on interview findings from Balwant (2017). In other words, we combine the well-

established items from Tepper's (2000) Abusive Supervision Scale with newly created items

derived from an inductive approach used by Balwant (2017). In writing the newly created items,

we kept the statements simple and short, and phrased items in a manner that is similar to

Tepper's items in order to maintain a consistent perspective (Hinkin, 1998) (see Table 1). Again,

for consistency, we used the same five-point scale used in Tepper's Abusive Supervision Scale.

Each point describes the frequency with which an instructor uses the behaviour stated by the

item. With respect to content adequacy of the newly created items (Hinkin, 1998), Balwant

---

[2] The sixth step of Hinkin's guide is replication. Replication involves repeating the study including all statistical analyses for an independent sample in order to improve confidence about reliability and validity findings (Hinkin, 1998). As such, replication within this paper would considerably add to the paper's complexity and length. For these reasons, we omitted the sixth step in this paper, but later offer suggestions for a replication study.

(2017) provided each student with the definition of destructive leadership, and asked students to describe incidents that were consistent with the provided definition. Therefore, the newly created items should reasonably measure destructive instructor-leadership.

[Insert Table 1 near here]

The second step is questionnaire administration (Hinkin, 1998). In distributing the questionnaires to samples of students, the new measure, i.e., Tepper's (2000) already established items along with our newly created items, is administered along with other established measures. For instance, we include a measure of supportive leadership in order to show that destructive leadership is not simply the opposite of supportive leadership. Here, we include the most popular supportive leadership theory – transformational leadership (Judge & Piccolo, 2004). Transformational leadership means that a leader shifts followers 'beyond immediate self-interests through idealized influence (charisma), inspiration, intellectual stimulation, or individualized consideration' (Bass, 1999, p. 11). We choose transformational leadership because meta-analytic findings show that transformational and destructive leadership are related to similar outcomes, but in opposite directions, e.g., satisfaction, attitude towards leader, and performance (Derue, Nahrgang, Wellman, & Humphrey, 2011; Judge & Piccolo, 2004; Schyns & Schilling, 2013; Wang, Oh, Courtright, & Colbert, 2011). Therefore, both constructs should be negatively related, but the relationship should not be high enough to suggest redundancy between the two concepts.

In addition to transformational leadership, we measure follower outcomes in order to examine criterion-related validity. Specifically, we measure students' satisfaction, extra effort, and individual performance because these are all established outcomes in the destructive leadership literature (Schyns & Schilling, 2013). Satisfaction refers to an overall level of

satiation or contentment. Destructive leaders are deceptive and victimize their followers, and thus followers are likely to be dissatisfied with such leadership. Followers may also respond to destructive leadership by reducing their efforts and performance when they try to distance themselves from the source of distress, e.g., withdrawing mentally and/or physically from the instructor's classes.

Third, Hinkin (1998) suggested that for <u>initial item reduction</u>, exploratory factor analysis (EFA) should be used to refine the scale. EFA is useful for determining the underlying structure of an item set. Here, we expect that three factors would emerge from the data as proposed by Balwant (2017), i.e., callous communication, chaotic carelessness, and irresponsibility. The EFA is described in more detail later on. After assessing the dimensionality of the measure, internal reliability can then be assessed using Cronbach's alpha. Hinkin (1998) suggested a coefficient alpha of at least 0.70 for exploratory measures.

For the fourth and fifth steps, Hinkin described the statistical techniques to be used. The fourth step is <u>confirmatory factor analysis (CFA)</u>. This technique is used to confirm the prior analyses and provide further evidence of construct validity. The fifth step is <u>convergent, discriminant, and criterion validity</u>. For these two steps, Hinkin (1998) explained conventions in using relevant statistical techniques, e.g., chi-square, goodness of fit indices, correlations between constructs, and regressions. In the present paper, we investigate all three forms of validity. We elaborate on the fourth and fifth steps in more detail in our upcoming presentation of the findings.

Methods

Research approach

We used a quantitative approach in this study because our aim was to develop a measure of destructive instructor-leadership. Specifically, we conducted a psychometric study of destructive instructor-leadership, and thus used various statistical techniques as described in our literature review and in accordance with Hinkin's (1998) scale development guide.

Sampling design and size

In quantitative research, nonprobability sampling is often used, but is not ideal for generalization purposes (Morling, 2014). Nonetheless we used a non-probability approach, i.e., we used a convenient sample of undergraduate students from the UK. We used this non-probabilistic approach because (1) the population of all undergraduate students in the UK was extremely large, thus making randomization cumbersome; (2) we had limited resources with respect to time and researchers to assist with data collection; and (3) students may be unwilling to participate in survey research, likely because of survey fatigue (Adams & Umbach, 2012). Taken together, these challenges made randomization unfeasible and would have likely resulted in a small sample size within our study's timeframe. To determine an appropriate sample size for a psychometric study, including its various statistical tests, we followed the recommendations by Hair et al. (2009) and Hinkin (1998), and set a goal of collecting at least 150 completed questionnaires.

Participants

We used the same sample examined in Balwant, Birdi, Stephan, and Topakas (2018) , but different measures and analyses were used in this study[3]. After accounting for missing data[4], outliers, and invalid responses, the sample size consisted of 174 undergraduate students studying at universities located in England (n = 161, 92.5%), Scotland (n = 7, 4.0%), Wales (n = 5, 2.9%), and Northern Ireland (n = 1, 0.6%). The students were from various faculties including Social Sciences (n = 43, 24.7%), Natural Sciences (n = 32, 18.4%), Arts and Humanities (n = 30, 17.2%), Medicine, Dentistry, and Health (n = 17, 9.8%), Engineering (n = 15, 8.6 %), Law (n = 9, 5.2%), Film (n = 7, 4.0%), and other faculties (n = 19, 10.9%). The sample included 49 males ranged from 18 to 40 years ($M = 21.63$, $SD = 4.72$) and 122 females ranged from 17 to 29 years ($M = 20.45$, $SD = 3.01$). Participants had to be undergraduate students in order to be included in the study. There were no other inclusion/exclusion criteria.

Materials

Preceding the questionnaire, brief instructions were given to participants asking them to rate one specific lecturer who taught them in the previous semester. In so doing, participants were required to (1) choose a lecturer who taught at least half of the classes for the module and (2) recall the grade received for said module. Participants were also provided with a definition of a module, which may have been useful for international students. A module was defined as a

---

[3] We intentionally designed two separate papers prior to collecting data. Each paper had unique research questions and separate theoretical implications that were too expansive for a single paper to address. Although there is a minor overlap between the two papers (i.e., transformational instructor-leadership was the focus of the other paper), the present study and the other paper include substantially unique variables, different findings, and different theoretical and practical implications.

[4] Four cases had extremely high levels of missing data, and were deleted from the analysis. For the remaining destructive instructor-leadership variables, missing data was very low (< 2%). Therefore, we used the EM approach to estimate missing data for these variables.

series of lectures on a subject that typically lasts a semester[5]. Participants were then asked to provide the lecturer's name, and this name was subsequently used to frame the upcoming questions (which was labelled '<Name>' in the present paper).

Destructive instructor-leadership. We measured destructive instructor-leadership using Tepper's Abusive Supervision Scale along with the supplementary items provided by Balwant (2017). Both measures were moderately correlated ($r = .50$) indicating that while they may both tap into the same concept, the supplementary items from Balwant (2017) still measured unique behaviours that may capture the first manifestation of destructive instructor-leadership. For the destructive instructor-leadership items, students were given instructions to indicate how often an instructor intentionally used the leader behaviours.

The Abusive Supervision Scale consisted of 15 items that were represented on a 5-point continuum (1 = I cannot remember <Name> ever using this behaviour with me; 2 = <Name> very seldom used this behaviour with me; 3 = <Name> occasionally used this behaviour with me; 4 = <Name> used this behaviour moderately often with me; 5 = <Name> used this behaviour very often with me) with higher scores indicating more abusive supervision. All items are shown in Table 1 ($\alpha = 0.83$).

The 15 supplementary items based on the findings from Balwant (2017) were represented on the same 5-point continuum used in Tepper's Abusive Supervision Scale, and therefore higher scores were indicative of higher destructive instructor-leadership. All items are listed in Table 1 ($\alpha = 0.83$).

Cronbach's α for the 30 items, including both Tepper's scale and the supplementary items from Balwant (2017), was .96.

---

[5] The term 'module' is generally used in the UK whereas 'course' is typically used in the US, Canada, and the Caribbean.

Transformational instructor-leadership. We measured transformational instructor-leadership using 17 items from the Multi-factor Leadership Questionnaire (MLQ). The MLQ items were represented on a 5-point continuum (0 = not at all; 1 = once in a while; 2 = sometimes; 3 = fairly often; 4 = frequently, if not always) with higher scores indicating higher transformational instructor-leadership. These items were adapted to the higher education module context using Pounder's word modifications (Pounder, 2008). In a separate analysis, we tested different factor structures of the MLQ as outlined by Antonakis et al. (2003) and found the best support for three transformational instructor-leadership dimensions including (1) charisma (11 items, e.g., '<Name> enthusiastically talked about what to do to make the module a success') (α = 0.92); (2) intellectual stimulation (3 items, e.g., '<Name> made me see a problem from different angles') (α = 0.75); and (3) individualized consideration (3 items, e.g., '<Name> treated me as a person – an individual entity, not just one among many students') (α = 0.86). Cronbach's α for the composite scale was .94.

Effectiveness, extra effort, and satisfaction. We used nine items from the MLQ to measure (1) perceived effectiveness of the instructor (4 items) (α = 0.90); (2) extent to which the instructor is able to motivate students to give extra effort (3 items) (α = 0.90); and (3) satisfaction with the instructor (2 items) (α = 0.78). All nine items were represented on a 5-point continuum (0 = not at all; 1 = once in a while; 2 = sometimes; 3 = fairly often; 4 = frequently, if not always) with higher scores indicating more positive outcomes.

Academic performance. We measured academic performance using students' self-reported final grade for a completed module. Students provided either their actual grade percentage or the range in which their grade percentage lie. We converted the commonly used

UK grade percentages to a 7-point scale (1 = no grade, 2 = 1-39, 3 = 40-44, 4 = 45-49, 5 = 50-59, 6 = 60-69, 7 = 70-100).

Demographic. We measured age in years and coded biological gender as '0' for male and '1' for female.

Procedures

We distributed the questionnaire during the second semester of the 2014/2015 academic year. Participants were asked to rate instructors from the first semester of said academic year. This approach of examining a completed module ensured that (1) students were sufficiently familiar with their instructor and (2) a grade could be provided for the module. We then distributed the questionnaire to five students in a small pilot study. The intention of the pilot study was to check understanding of item wordings, and the pilot study participants did not identify any issues. After the pilot study, we distributed the questionnaire in two ways.

First, we sent an email to all undergraduate students at a university located in England. In the email message, participants were given a brief description of the study, a link to an information sheet, and a link to the online questionnaire. Each participant could have opted to receive a free personality evaluation along with entry into a £40 prize voucher draw. The sample from this survey consisted of 102 students.

Second, the questionnaire was distributed to students at other UK universities via Qualtrics panel service. Qualtrics was used to source 100 undergraduate students from the UK. Because each of the participants from the Qualtrics panel was likely rewarded with an external incentive, we used two attention filters in the questionnaire to improve the quality of the data, i.e., verify that respondents were reading the questions carefully and following instructions. An example of an attention filter was, 'Please select "Strongly agree" for this statement'. From the

Qualtrics panel, 205 students completed the questionnaire, with 100 students being filtered out via the attention filters. Therefore, the Qualtrics panel was used to source a total of 105 students. Note that for the Qualtrics panel, participants were not (1) offered the option to receive a free personality evaluation and (2) entered into the prize draw.

Ethical considerations

We provided potential participants with an information sheet. The information sheet contained brief details regarding why the study was being done and what completing the questionnaire involved. Furthermore, the information sheet covered ethical issues such as (1) participation being voluntary, with the option of withdrawing at any time; (2) protocol to follow if anything goes wrong, e.g., contacting the university registrar or consulting a counselling service or personal doctor; (3) confidentiality of responses; (4) how the findings of the study would be used; and (5) details of the committee granting ethical approval of the study. The University Research Ethics Committee (UREC) at The University of *Name Withheld for Anonymity* granted ethical approval for this study.

Results

We used two computer programs to conduct statistical analyses in the present paper. First, we used SPSS Statistics for Windows, version 22 (IBM Corp., 2013) to test normality, component structures, convergent validity, and criterion validity. Second, we used the Lavaan package (Rosseel, 2012) for R (R Core Team, 2013) to conduct confirmatory factor analyses and test both convergent and discriminant validity.

Normality

For the 30 destructive instructor-leadership variables, all were severely non-normal, i.e., all of the kurtosis *z*-scores greatly exceeded the critical value of ±2.58 (ranging from 12.69 to

275.77), and 16 of the skewness z-scores greatly exceeded the critical value of ± 2.58 (ranging

from 12.31 to 53.63)[6]. With such large deviations from normality, χ^2 derived from maximum

likelihood estimation gets excessively large (Blunch, 2008; Byrne, 2001). Also, such non-

normality could lead to analyses failing to converge, improper solutions, and underestimated fit

indices (Blunch, 2008; Byrne, 2001). Because the deviations from normality were so severe, data

transformations recommended by Tabachnick and Fidell (2007) and Hair et al. (2009) did not

lead to any notable improvements in normality. When deviations from normality are so severe

that no transformation improves the distributions, Tabachnick and Fidell (2007) suggested

dichotomizing the variable. Therefore, we dichotomized all of the destructive instructor-

leadership variables as 0 for 'leader did not use the behaviour' and 1 for 'leader used the

behaviour'. This change in coding means that the destructive instructor-leadership variables now

measured use of destructive instructor-leadership rather than frequency of leadership behaviour.

Component structure

Following Hinkin's guidelines, we conducted a principal component analysis (PCA) on

the 30 destructive instructor-leadership items for initial item reduction. For the PCA, we used

oblique rotation (Promax) in order to allow the components to correlate as expected. Various

tests were used to determine the number of components to extract (i.e., Kaiser's criterion = 9,

Velicer's Revised Minimum Average Partial (MAP) test = 2, and Horn's parallel analysis = 3).

Given that Kaiser's criterion led to 5 components with weak loadings, we tested two-, three-, and

four-component solutions. Both the three-component and four-component structures produced

---

[6] The skewness of the destructive instructor-leadership variables indicated that destructive leadership is rare in higher education module teaching. The rarity of destructive leadership in the higher education module context coincides with its rarity in corporate contexts (Hubert & Veldhoven, 2001). Therefore, while our skewed data presented certain challenges when using statistical tests, we were pleased to discover that destructive instructor-leadership is not a common phenomenon.

the clearest structures with stronger components than the two-component solution. However, the final four-component solution contained one component that comprised of only two feedback items (inaccurate and poor feedback). In a four-component structure, the feedback component would need to be enhanced, perhaps by adding more feedback oriented questions to the survey. Moreover, in the four-component structure, the remaining three components contained fewer stronger loadings than the three-component solution. For these reasons, along with the parallel analysis results, we specified a three-component structure.

For the three-component structure, we conducted several re-specifications and deleted 17 items in an iterative process due to poor representation by the component structure[7]. Component loadings were expected to be greater than .425 based on a sample size of 174 (Hair et al., 2009). Additionally, communalities should be greater than .5 (Hair et al., 2009). In the first iteration, items 9, 13, 19, and 22 had communalities that were less than .3, and thus these items were deleted first because of their very poor representation by the component structure. After deleting the four stated items, the findings of a second iteration indicated that items 2, 6, 17, 20, 21, and 28 had communalities that were less than .4, and thus were deleted. With all communalities now greater than .4, we addressed items with low communalities along with cross-loading issues. A cross-loading meant that a given item loaded at .32 or higher on two or more components (Tabachnick & Fidell, 2007). For the third iteration, very problematic cross-loadings for items 14, 24, 27, and 30 were identified and deleted. For the fourth iteration, the communality for item 18 dropped below .4, and therefore this item was deleted. For the fifth iteration, the communality

---

[7] Even though many items were deleted during the item reduction stage, our intent was to reduce the initial item set to a smaller set of items. Hinkin (1998) explains that the item reduction creates a more parsimonious representation of the initial item set. Following Hinkin's guide, we only retained items that clearly loaded on a single factor. Moreover, deletion of many items during scale development is not uncommon in psychometrics, e.g., the Five-Factor inventory (John & Srivastava, 1999).

for item 7 was less than .5, and the item was dissimilar to other items on its component, i.e., item 7 described not giving credit for work whereas the other items on that component described communication behaviours. For these reasons, we deleted item 7. For the sixth iteration, the communality for item 25 dropped below .5, and this item was the weakest loading item on its five-item component. For these reasons, we deleted item 25. For the seventh iteration, all communalities were acceptable, and there was a clear component solution with no cross-loadings. Therefore, after six iterations of item reduction, the final component solution was represented by 13 items and explained 70.10% of the variance (see Table 2).

The components were named as follows:

Component 1. Irresponsibility: The items that loaded on this component related to unscrupulous and dishonourable leadership actions for which followers are deceived.

Component 2. Victimization: The items on this component related to a leader harassing and picking on a follower.

Component 3. Callous communication: The items that loaded on this component described harmful communication actions by the leader.

The subscale for each of the three components was reliable given that each Cronbach's alpha value exceeded .7 as shown in Table 2 (Hinkin, 1998).

[Insert Table 2 near here]

These findings partially supported the three dimensions proposed by Balwant (2017). Specifically, both callous communication and irresponsibility were fairly well represented. However, chaotic carelessness was not an emergent dimension. Still, two items that were initially expected to load on a chaotic carelessness factor, loaded on two other factors. For instance, misleading behaviour loaded on the factor that seemed to represent irresponsible behaviour.

Also, giving the silent treatment loaded on a factor that appeared to represent callous communication. These loadings indicated that chaotic carelessness may overlap with other dimensions instead of being a distinct dimension as proposed by Balwant (2017). Also, the victimization dimension is not proposed by Balwant (2017), but is an emergent factor from Tepper's (2000) scale. Finally, in a separate analysis, Tepper's scale by itself did not contain an irresponsibility dimension – the dimension that is related to the first manifestation of destructive leadership.

Confirmatory factor analysis

Following Hinkin's (1998) guide, for the next step we conducted a CFA to reproduce the derived 13-item three-factor solution. Given that all of the observed variables were categorical, we used diagonally weighted least squares (DWLS) estimation. DWLS provides more accurate estimates and fit indices for models when variables have a small number of categories (Mîndrilă, 2010). Details of the final CFA are given in Table 2. We followed the guidelines by Hair et al. (2009) to assess model fit, and thus report one incremental fit index, i.e., Comparative Fit Index (CFI), and one absolute fit index, i.e., Root Mean Square Error of Approximation (RMSEA). All of the robust fit indices indicated good model fit (χ^2/df = 1.02, CFI = 0.97, RMSEA = 0.01). χ^2 was 63.03 (df = 62) and p > .05.

We tested a series of competing models in order to determine whether the data could be represented by a better fitting model than the proposed measurement model. The baseline model used for all comparisons was the three-factor model from our initial CFA. First, a higher-order factor was tested, and the fit results were the same as the three-factor model, thus indicating that a single higher-order factor of destructive instructor-leadership could represent the data. This result was expected because the number of parameters to be estimated in both the baseline and

higher-order models was the same, i.e., these two models were empirically equivalent (Lee &

Hershberger, 1990). <u>Second</u>, we tested a one-factor model, and this model fitted significantly

worse than the baseline three-factor model ($\Delta\chi^2$ [3] = 24.37, $p < .001$). <u>Third</u>, we tested a two-

factor model combining the irresponsibility and victimization factors. For this model, the fit

results were significantly worse than the baseline three-factor model ($\Delta\chi^2$ [2] = 9.96, $p < .05$).

<u>Fourth</u>, we tested a two-factor model combining the victimization and callous communication

factors. Again, the fit results were significantly worse than the baseline model ($\Delta\chi^2$ [2] = 18.07, p

$< .001$). <u>Fifth</u>, we tested a two-factor model combining the irresponsibility and callous

communication factors into a single factor, and this model fitted significantly worse than the

baseline model ($\Delta\chi^2$ [2] = 15.03, $p < .001$). Overall, the model comparisons indicated that

destructive instructor-leadership can be represented by either three first-order constructs or a

single second-order construct.

Convergent validity

To assess convergent validity, we examined the correlation matrix for the expected

correlations between the destructive instructor-leadership dimensions and other instructor-

leadership dimensions (see Table 3), factor loadings, average variance extracted (AVE), and

construct reliability (Hair et al., 2009). <u>First</u>, each of the destructive instructor-leadership

dimensions shared at least one significant and negative association with a transformational

instructor-leadership dimension. More importantly, the combined destructive instructor-

leadership construct was significantly and negatively associated with the combined

transformational instructor-leadership construct as expected ($r = -.27$, $p < .01$). In addition, the

destructive instructor-leadership dimensions and overall construct also (1) shared expected

negative associations with active forms of instructor-leadership (i.e., contingent reward and

active management by exception), and (2) expected positive associations with other forms of

negative instructor-leadership (i.e., passive management by exception and laissez-faire). These

correlations all indicate good convergent validity. Second, all factor loadings exceeded the

minimum threshold of .5, and most loadings met or exceeded the ideal threshold of .7 (Hair et

al., 2009). Third, AVE was greater than 50% for each dimension (Hair et al., 2009). Fourth, all

construct reliability values exceeded 0.70 (see Table 2 for reported factor loadings, AVE, and

construct reliabilities). Overall, convergent validity for the three-factor solution was good.

[Insert Table 3 near here]

Discriminant validity

The discriminant validity for the destructive instructor-leadership dimensions were all

fairly good. For instance, the highest interconstruct correlation was 0.62, and thus did not exceed

the cut-off value of 0.85 proposed by Kline (2011). Furthermore, all three interconstruct

correlations were positive, which was an indication of good nomological validity, i.e., this

finding suggested that the three dimensions tapped into destructive instructor-leadership.

To further investigate discriminant validity, we compared the average variance extracted

(AVE) for each of the three destructive instructor-leadership constructs with that of its squared

interconstruct correlation (SIC) with the three transformational instructor-leadership constructs

(Hair et al., 2009). This comparison reveals whether the destructive instructor-leadership

constructs are truly distinct from the transformational instructor-leadership constructs (i.e., not

simply opposite leadership behaviours). The findings suggested excellent discriminant validity

because the AVE values for irresponsibility (.74), victimization (.52), and callous

communication (.53) all greatly exceeded the SIC values for each of the transformational

instructor-leadership dimensions (ranging from .00 to .15). Therefore, there was good evidence

that the destructive instructor-leadership constructs capture phenomena that the transformational

instructor-leadership constructs do not.

Criterion validity

To establish criterion-related validity, multiple regression models were estimated for

effectiveness, satisfaction, extra effort, and academic performance. For each model, the

independent variables entered included the three destructive instructor-leadership dimensions[8] as

well as the control variables of age and gender (see Table 4). For all of the outcomes, including

effectiveness, satisfaction, extra effort, and academic performance, neither age nor gender was a

significant predictor. Callous communication was a significant and strong predictor of

effectiveness (β = -0.32, p < .001), satisfaction (β = -0.36, p < .001), and extra effort (β = -0.36, p

< .001) in the expected direction. Irresponsibility was also a significant predictor of extra effort

(β = -0.17, p < .05) in the expected direction. Victimization was not significant in any of these

models. Moreover, none of the destructive instructor-leadership dimensions were significant

predictors in the academic performance model. Overall, these findings partially correspond with

previous research which showed that destructive leadership was negatively related to individual

performance (i.e., effort and performance) and satisfaction (Schyns & Schilling, 2013).

[Insert Table 4 near here]

Incremental validity

We used hierarchical regression to further test for incremental validity of the destructive

instructor-leadership constructs beyond other negative forms of instructor-leadership, i.e.,

laissez-faire and passive management by exception (see Table 5). As explained by Krasikova et

al. (2013), destructive leadership and ineffective leadership are conceptually different because

---

[8] Each destructive instructor-leadership dimension was measured by its respective EFA regression weight.

the former describes volitional behaviour, whereas the latter deals with a leader's inability to mobilize followers towards goals. Therefore, empirically, destructive instructor-leadership should contribute additional variance above laissez-faire and passive management by exception.

[Insert Table 5 near here]

For the hierarchical regression model, the variate's assumptions were met and multicollinearity was not problematic. Age and gender were entered in the first step, followed by laissez-faire and passive management by exception in step 2, and then the three destructive instructor-leadership dimensions in step 3. The incremental validity findings showed that (1) one of the destructive instructor-leadership dimensions explained additional variance above laissez-faire and passive management by exception in predicting effectiveness (50% more variance), and satisfaction (60% more variance), and (2) two of the destructive instructor-leadership dimensions explained additional variance above extra effort (68.42% more variance). Specifically, callous communication and irresponsibility remained significant in their respective models from Table 5, even when laissez-faire was a significant predictor. These findings support the notion that destructive instructor-leadership goes beyond these other forms of negative instructor-leadership, i.e., unsupportive instructor-leadership or non-leadership. See Table 6 for a summary of the findings from each of Hinkin's steps.

[Insert Table 6 near here]

Discussion

The aim of this paper is to develop a measure of destructive instructor-leadership that is aligned with the conceptualization proposed by Krasikova et al. (2013). In following Hinkin's (1998) framework, we (1) adapt items from Tepper's (2000) scale and generate items from Balwant (2017), (2) administer the questionnaire, (3) reduce the number of items via exploratory

factor analysis, (4) conduct a confirmatory factor analysis, and (5) conduct validity tests. The findings suggest that destructive instructor-leadership can be measured by thirteen items that capture three dimensions including, irresponsibility, victimization, and callous communication. On the one hand, irresponsibility is related to the first manifestation of destructive leadership that describes leading followers towards contradictory goals. On the other hand, victimization and callous communication are directly linked to the second manifestation of destructive leadership that describes the use of harmful behaviours in leading followers.

The findings also indicate that certain items from Tepper's (2000) scale are poorly represented in a factor structure. Specifically, six of the fifteen items from Tepper's scale were not well represented by the three-factor structure in this paper. Mitchell and Ambrose (2007) illustrated similar issues with Tepper's scale because five items were poorly represented by their two-factor structure. Perhaps some of these deleted items may indicate that certain destructive leadership behaviours from the corporate context are not applicable to the higher education module context. For e.g., given the short-term distant nature of higher education modules, it is unlikely that instructors would be able to recall past instances of students' mistakes and failures (item 6 in Table 1). Alternatively, some of Tepper's items may need to be reworded or supplemented by similar behavioural items in developing other destructive instructor-leadership dimensions. For e.g., not giving credit for work requiring a lot of effort (item 7 in Table 1) is likely to be applicable to higher education teaching, and may need to be further refined. In addition to this shortcoming, arguably the biggest flaw with Tepper's scale is that it largely ignores the first manifestation of destructive leadership as proposed by Krasikova et al. (2013), i.e., leading followers towards contradictory goals.

For the new measure, there was fairly good support for criterion and incremental validity of the destructive instructor-leadership dimensions. The callous communication dimension appears to be the strongest predictor of the examined outcomes, i.e., perceived effectiveness, satisfaction, and extra effort. Also, the new measure contributes additional variance beyond measures of transformational instructor-leadership and ineffective instructor-leadership. Overall, this paper contributes a measure of destructive instructor-leadership that is aligned with the conceptualization of destructive leadership as proposed by Krasikova et al. (2013), and that is built using nine items from Tepper's 15-item measure and four items from Balwant's (2017) interview study.

In developing the destructive instructor-leadership measure, this paper also contributes to the dimensionality of destructive leadership. The dimensionality of destructive leadership has largely been ignored in leadership research. As explained earlier, previous studies often used a single construct representing destructive leadership, which was found to be negatively related to follower outcomes (Schyns & Schilling, 2013). In this paper, we show the importance of examining the individual dimensions of destructive leadership. Specifically, when destructive leadership is examined at the multidimensional level, each dimension shares varying associations with follower outcomes (recall the criterion validity tests as summarized in Table 6). In light of our findings, certain destructive leadership dimensions may be more harmful than others depending on the follower outcome. To address our opening research question, we show that (1) destructive instructor-leadership is composed of three dimensions including irresponsibility, victimization, and callous communication and (2) these three dimensions can be captured by a 13-item measure.

Limitations and directions for future research

A potential limitation in the present paper is the item wording of the destructive instructor-leadership measure. Here, the items are worded so that each student is asked to rate his/her instructor's actions towards him/her. This phrasing is used in the original version of Tepper's scale, and ensures consistency between other leadership measures, e.g., MLQ. However, Balwant (2017) explains that students sometimes perceive their instructor as being a destructive leader based on the instructor's actions towards the class in general, i.e., a bystander effect. Therefore, the personal leadership measure used in this paper does not capture students' perceptions of destructive instructor-leadership that are based upon their observations of the instructor's actions towards other students. This limitation may partially explain why the destructive instructor-leadership variables are severely positively skewed.

To incorporate the bystander effect, and thus potentially increase variation in responses, future research can include an opening statement such as, 'Report how frequently you observe your instructor/supervisor using the following actions towards you and/or other students/employees'. Each item would then state the leader's actions towards the follower and/or other followers, e.g., ridiculed me and/or other students/employees. Alternatively, future research can even compare findings between (1) followers who experience destructive instructor-leadership on a personal one-on-one level and (2) followers who experience their instructor using destructive instructor-leadership behaviour towards others (i.e., a bystander's perspective). Such a comparison can reveal whether destructive instructor-leadership effects vary according to the type of interaction between instructor and student.

A second limitation in the present paper is that we rely on cross-sectional data. Cross-sectional data cannot be used to determine direction of causality. It is possible that the reverse

relationships are true. For instance, it is plausible to expect that a leader may react to poor performing followers by using callous communication. To test causality, future research can adopt randomized experimental designs. An ethical issue in using randomized trials is the exposure of individuals to destructive leadership behaviours. One way to minimize this potential harm is to screen participants and exclude those that suffer from depression or social anxiety. However, this selection bias can affect the validity of the results. An alternative approach is to use a computer-simulation. Participants can be immersed into a computer-simulated environment that simulates the experience of being exposed to a destructive leader. The use of a simulated leader minimizes the ethical issue of having 'real-life' leaders use harmful behaviours towards participants. Simulated environments can be further enhanced via the use of virtual reality technology – a technology that is becoming increasingly popular and feasible for research purposes (Kiss, 2014).

A third limitation of the present study is that we did not conduct a replication study. As stated earlier, a replication study is needed to enhance the generalizability of the newly developed measure in this study (Hinkin, 1998). In a replication study, we suggest that convergent validity be further examined to determine how well the newly developed measure correlates with other measures that tap into the same concept, e.g., Duffy et al.'s (2002) Social Undermining Scale.

A final note is that we investigate leadership from an instructor-student perspective. This perspective is not a limitation given the purpose of our study. However, we caution against using these findings in a corporate context. Even though an instructor-student relationship is similar to a supervisor-employee relationship, the former is characterized by unique nuances that are absent in the corporate context. For instance, students pay for their tuition, and are generally not

accountable to their instructors in the same way that employees are accountable to their

supervisors or a CEO. Also, the length of instructor-student relationships are typically shorter

than supervisor-employee relationships (for a more detailed discussion on context differences,

see Balwant, 2017). For these reasons, we suggest that our study be replicated not only in a

module/course context but also in a corporate context.

Practical implications and conclusion

The present study offers an important practical implication for identifying destructive

instructor-leadership in higher education. While our study shows that destructive instructor-

leadership may be rare, it still exists. Unfortunately, destructive instructor-leader behaviours can

go unnoticed by instructors themselves and is hardly mentioned in higher education (Boice,

1996). Moreover, quantitative questions in module/course evaluations are generally targeted

towards measuring supportive teaching practices rather than destructive instructor-leadership.

With this in mind, we have a few suggestions for identifying destructive instructor-leadership.

First, open-ended questions in module evaluations can be content analysed and cross-referenced

to our newly developed measure in order to determine if an instructor is using destructive leader

behaviours. Second, if an instructor receives poor module evaluations, our measure can be

included in future module evaluations to determine if the instructor is simply teaching poorly or

engaging in behaviours that are more aversive, i.e., destructive instructor-leadership. Third, if an

instructor receives poor module evaluations, trained observers can use our measure to observe

and note destructive instructor-leader behaviours in an unbiased manner (Boice, 1996). Once one

or more of these three methods detect destructive instructor-leadership, then a training

programme geared towards the changing of these leader behaviours may be necessary.

A detailed explanation of a training programme for changing destructive instructor-leadership is beyond the scope of the present paper. However, we have a few recommendations for how our measure can be used to design and implement a training programme. First, our measure can be used to prepare instructional objectives (Mager, 1997). For instance, a training session on callous communication may have an objective such as 'refrain from using words or actions that ridicule, put down, and show disinterest towards students in all settings, e.g., in the presence of other students, when students do not know the answer to a question, private meetings, etc.'. Second, our measure can be useful for the training methods chosen (Werner & DeSimone, 2011). For instance, if behavioural modelling is used, trainees can be shown visual examples (e.g., videos and/or skits) of destructive instructor-leader behaviours that are based on items from our measure. Finally, when evaluating the training programme, our measure can be useful for ongoing evaluation of the training programme as well as to determine whether trainees transferred their learning to the classroom. Here, our measure can be used along with module evaluations to conduct pre-training and post-training comparisons.

In summary, we followed Hinkin's (1998) guide for scale development in order to enhance Tepper's (2000) measure of abusive supervision so that it is better aligned with the conceptualization of destructive instructor-leadership. The new 13-item measure captures three dimensions of destructive instructor-leadership including irresponsibility, victimization, and callous communication. We found fairly good support for the new measure of destructive instructor-leadership and its dimensions. Nonetheless, we recommend that the new measure be administered to (1) an independent student sample and (2) employee samples in corporate contexts because doing so would improve the generalizability of said measure.

References

Adams, M. J. D., & Umbach, P. D. (2012). Nonresponse and online student evaluations of

teaching: Understanding the influence of salience, fatigue, and academic environments.

Research in Higher Education, *53*(5), 576–591. https://doi.org/10.1007/s11162-011-

9240-5

Antonakis, J., Avolio, B. J., & Sivasubramaniam, N. (2003). Context and leadership: an

examination of the nine-factor full-range leadership theory using the Multifactor

Leadership Questionnaire. *The Leadership Quarterly*, *14*(3), 261–295.

https://doi.org/10.1016/S1048-9843(03)00030-4

Ashforth, B. E. (1994). Petty tyranny in organizations. *Human Relations*, *47*, 755–778.

https://doi.org/10.1177/001872679404700701

Baba, V. V., & Ace, M. E. (1989). Serendipity in leadership: Initiating structure and

consideration in the classroom. *Human Relations*, *42*(6), 509–525.

https://doi.org/10.1177/001872678904200603

Balwant, P. T. (2016). Transformational instructor-leadership in higher education teaching: A

meta-analytic review and research agenda. *Journal of Leadership Studies*, *9*(4), 20–42.

https://doi.org/10.1002/jls.21423

Balwant, P. T. (2017). The dark side of teaching: destructive instructor leadership and its

association with students' affect, behaviour, and cognition. *International Journal of

Leadership in Education*, *20*(5), 577–604.

https://doi.org/10.1080/13603124.2015.1112432

Balwant, P. T., Birdi, K., Stephan, U., & Topakas, A. (2018). Transformational instructor-

leadership and academic performance: a moderated mediation model of student

engagement and structural distance. *Journal of Further and Higher Education*, 1–17.

https://doi.org/10.1080/0309877X.2017.1420149

Bass, B. M. (1999). Two decades of research and development in transformational leadership.

European Journal of Work and Organizational Psychology, *8*(1), 9–32.

https://doi.org/10.1080/135943299398410

Bligh, M. C., Kohles, J. C., Pearce, C. L., Justin, J. E. (Gene), & Stovall, J. F. (2007). When the

Romance is Over: Follower Perspectives of Aversive Leadership. *Applied Psychology:*

An International Review, *56*(4), 528–557. https://doi.org/10.1111/j.1464-

0597.2007.00303.x

Blunch, N. (2008). *Introduction to Structural Equation Modelling Using SPSS and Amos*. Los

Angeles: SAGE Publications Ltd.

Boice, B. (1996). Classroom incivilities. *Research in Higher Education*, *37*(4), 453–486.

https://doi.org/10.1007/BF01730110

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications,*

and programming. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Dawson, J. E., Messe, L. A., & Phillips, J. L. (1972). Effect of instructor-leader behavior on

student performance. *Journal of Applied Psychology*, *56*(5), 369–376.

https://doi.org/10.1037/h0033436

Derue, D. S., Nahrgang, J. D., Wellman, N., & Humphrey, S. E. (2011). Trait and behavioral

theories of leadership: An integration and meta-analytic test of their relative validity.

Personnel Psychology, *64*(1), 7–52. https://doi.org/10.1111/j.1744-6570.2010.01201.x

Duffy, M. K., Ganster, D. C., & Pagon, M. (2002). Social Undermining in the Workplace. *The*

Academy of Management Journal, *45*(2), 331–351.

Ellis, K. (2000). Perceived teacher confirmation: The development and validation of an instrument and two studies of the relationship to cognitive and affective learning. *Human Communication Research*, *26*(2), 264–291.

Goodyear, R. K., Crego, C. A., & Johnston, M. W. (1992). Ethical issues in the supervision of student research: A study of critical incidents. *Professional Psychology: Research & Practice*, *23*(3), 203–210.

Hair, J. F. J., Black, W. C., Babin, B. J., & Anderson, R. (2009). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*(1), 104–121. https://doi.org/10.1177/109442819800100106

Hubert, A. B., & Veldhoven, M. van. (2001). Risk sectors for undesirable behaviour and mobbing. *European Journal of Work and Organizational Psychology*, *10*(4), 415–424. https://doi.org/10.1080/13594320143000799

IBM Corp. (2013). *SPSS Statistics for Windows, Version 22.0*. Armonk, NY: IBM Corp.

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research, 2nd ed.* (pp. 102–138). New York, NY, US: Guilford Press.

Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, *89*(5), 755–768. https://doi.org/10.1037/0021-9010.89.5.755

Kearney, P., Plax, T. G., Hays, E. R., & Ivey, M. J. (1991). College teacher misbehaviors: What

students don't like about what teachers say and do. *Communication Quarterly*, *39*(4),

309–324. https://doi.org/10.1080/01463379109369808

Kiss, J. (2014, March 25). Oculus: Facebook buys virtual reality gaming firm for $2bn. *The*

Guardian. Retrieved from

http://www.theguardian.com/technology/2014/mar/25/facebook-buys-virtual-reality-

gaming-firm-oculus

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York:

Guilford Press.

Krasikova, D. V., Green, S. G., & LeBreton, J. M. (2013). Destructive leadership: A theoretical

review, integration, and future research agenda. *Journal of Management*, *39*(5), 1308–

1338. https://doi.org/10.1177/0149206312471388

Lee, S., & Hershberger, S. (1990). A simple rule for generating equivalent models in covariance

structure modeling. *Multivariate Behavioral Research*, *25*(3), 313–334.

https://doi.org/10.1207/s15327906mbr2503_4

Mackey, J. D., Frieder, R. E., Brees, J. R., & Martinko, M. J. (2017). Abusive supervision: A

meta-analysis and empirical review. *Journal of Management*, *43*(6), 1940–1965.

https://doi.org/10.1177/0149206315573997

Mager, R. F. (1997). *Preparing instructional objectives: A critical tool in the development of*

effective instruction. Atlanta, GA: Center for Effective Performance.

Mîndrilă, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS)

estimation procedures: A comparison of estimation bias with ordinal and multivariate

non-normal data. *International Journal of Digital Society (IJDS)*, *1*(1).

Mitchell, M. S., & Ambrose, M. L. (2007). Abusive supervision and workplace deviance and the moderating effects of negative reciprocity beliefs. *Journal of Applied Psychology*, *92*, 1159–1168. https://doi.org/10.1037/0021-9010.92.4.1159

Morling, B. (2014). *Research methods in psychology: Evaluating a world of information* (2 edition). New York: W. W. Norton & Company.

Pounder, J. S. (2008). Full-range classroom leadership: Implications for the cross-organizational and cross-cultural applicability of the transformational-transactional paradigm. *Leadership*, *4*(2), 115–135. https://doi.org/10.1177/1742715008089634

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.

Schyns, B., & Hansbrough, T. (2010). *When leadership goes wrong: Destructive leadership, mistakes, and ethical failures*. Information Age Publishing.

Schyns, B., & Schilling, J. (2013). How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. *The Leadership Quarterly*, *24*(1), 138–158. https://doi.org/10.1016/j.leaqua.2012.09.001

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed). Boston: Pearson/Allyn & Bacon.

Tepper, B. J. (2000). Consequences of abusive supervision. *The Academy of Management Journal*, *43*(2), 178–190. https://doi.org/10.2307/1556375

Wang, G., Oh, I.-S., Courtright, S. H., & Colbert, A. E. (2011). Transformational leadership and performance across criteria and levels: A meta-analytic review of 25 years of research.

Group & Organization Management, *36*(2), 223–270.

https://doi.org/10.1177/1059601111401017

Werner, J. M., & DeSimone, R. L. (2011). *Human resource development* (6 edition). Mason,

OH : Andover: Cengage Learning.

Yukl, G. (2006). *Leadership in Organizations* (6th ed.). New Jersey: Prentice Hall.