

Empirical and Theoretical Validity of Hersey-Blanchard's Contingency Model

A Critical Analysis

Gavriel Meirovich and Jian Gu

Salem State University, USA

There is a certain discrepancy that can be labeled a fundamental paradox of Hersey and Blanchard's Situational Leadership Theory (SLT). On the one hand this model is intuitively appealing and popular among practitioners. On the other hand there is a consensus in the literature of a dearth of solid empirical evidence supporting the model. This study contributes to the field of contingency leadership in several ways. First, it provides comprehensive analysis of the model itself and its theoretical and empirical criticism. Second, it addresses the issue of the external validity of the model and its overlap with other theoretical constructs. Third, it utilizes an innovative approach in measurement of effects of fit between employees' readiness and leadership style. Last, it provides empirical evidence of the validity of SLT.

- Leadership style
- Readiness
- Contingency approach

Gavriel Meirovich is a Professor of Management at Salem State University, Bertolon School of Business in Salem, Massachusetts. He has earned his Ph.D. in Management from Yaroslavl University, Russia. He is presently teaching Management Theory and Practice, and Business Ethics. His research interests are in the fields of Quality, Emotions, Cross-Cultural Ethics and Contingency Leadership.

✉ Department of Management,
Bertolon School of Business, Salem
State University, 352 Lafayette
Street, Salem, MA 01930, USA

💻 gmeirovich@salemstate.edu

Jian Gu is Associate Professor of Management at Salem State University, Bertolon School of Business in Salem, Massachusetts. He has earned his Ph.D. in Management from Temple University, Philadelphia. He is presently teaching Business Policy & Strategy, and Managing in the Global Business Environment. His research interests are in the fields of Cross-Border Mergers and Cross-Cultural Management.

✉ Department of Management,
Bertolon School of Business, Salem
State University, 352 Lafayette
Street, Salem, MA 01930, USA

💻 jgu@salemstate.edu

FOR SOME TIME IT HAS been commonplace to claim that Hersey and Blanchard's Situational Leadership Theory (SLT), despite its popularity among practitioners, lacks empirical support (Goodson *et al.* 1989; Blank *et al.* 1990; Graeff 1997). This criticism was echoed by several studies addressing various aspects of leadership (see e.g. Avolio *et al.* 2003; Aldoory & Toth 2004; Derue *et al.* 2011). Objective analysis shows that as more studies on the subject appear this argument becomes less persuasive—there is certain empirical and theoretical evidence of the validity of the SLT model. The goal of this study is to undertake a comprehensive analysis of both the model itself and its theoretical and empirical criticism, and to provide empirical evidence of the model's validity. The structure of the paper is the following: First we give a concise description of the SLT theory. Second, we analyze empirical findings in regards to the model. Third, we analyze theoretical arguments against SLT. Fourth, we address the issue of external validity of the model and its overlap with other theoretical constructs. Finally we formulate and test hypotheses in regards to empirical validation of discussed theoretical concepts.

The Hersey-Blanchard Situational Model

The initial version of Hersey-Blanchard's normative model stipulated that leadership styles should match the followers' maturity (Hersey & Blanchard, 1977). In their later version maturity was replaced with the more specific term "readiness" (Hersey *et al.*, 1996, 2013). Since both terms are used in theoretical discussion in the literature, in this paper these two notions are used interchangeably. Readiness includes the willingness and abilities of followers. Four levels of readiness require using four different styles— S_1 Telling; S_2 Selling; S_3 Participating and S_4 Delegating. These styles are based on the combination of two dimensions, task orientation and people (relationship) orientation. Followers at the lowest level of readiness, R_1 , are not willing and not able to perform a task. In this case leaders have to use the S_1 Telling style, which involves high task and low people orientation. Its essence is creating a structure, i.e., providing specific and explicit directions on what tasks should be carried out (Hersey *et al.*, 2013, p. 125). Moderate level of readiness, R_2 , involves high willingness and low abilities. The right style for this level is S_2 Selling, which reflects high focus on both task and relationship. Applying this style, leaders use persuasion, explain the rationale of assignments to followers, and show consideration so as to maintain followers' motivation (*ibid.*, p. 126).

When employees reach R_3 , high level of readiness, leaders need to gain followers' participation in decision-making. S_3 Participating style involves low level of task focus and high level of people focus (*ibid.*, p. 127). Followers are knowledgeable enough so as to provide valuable input in solving problems, so they do not need much structuring, like in previous cases. The majority of the leader's effort and attention should be directed towards relationships rather than task dimensions. Followers at the highest level of readiness, R_4 , are experts

in their field; they are motivated and capable of doing the job autonomously. The behavior which fits these circumstances is the S_4 Delegating style. It includes a low focus on both task and relationship dimensions (ibid., p. 128). A leader delegates authority and responsibility to followers, who make their own decisions.

Empirical evidence and criticism

Hambleton and Gumbert (1982), using the Leader Behavior Analysis (LDA) instrument, found a positive relationship between situational leadership styles and managers' perceptions of followers' performance. They argued that in situations where the SLT model was used appropriately, subordinates' performance was significantly higher (ibid., p. 240). At the same time the authors cautioned against generalization of the results due to the problems with self-assessments of managerial style and with the representativeness of the managers sample.

Goodson *et al.* (1989) did not find a significant relationship between employees' readiness and leadership behavior in terms of initiating structure and consideration. Their findings did not support predictions of SLT with respect to the best, second best, third best and worst styles of leadership. However, careful analysis of their empirical data might produce different conclusions. The authors established the level of maturity by dividing participants' self-rated responses into quartiles, resulting in approximately equal numbers of employees in four maturity categories. Such a technique raises serious questions. Categories of maturity/readiness appear to be arbitrary. If an employee finds himself in the first quartile that does not necessarily mean that he possesses the lowest readiness, R_1 . By the same token those who are located in the highest quartile do not automatically possess the highest level, R_4 . Approximately equal numbers in four categories of maturity are suspicious. It is more plausible that due to the law of normal distribution most employees find themselves close to the central point in the intermediate groups of R_2 and R_3 levels of maturity. This proposition is supported by other studies. Cairns *et al.* (1998), in their research of senior executives within service and manufacturing businesses, reported that most followers were on R_2 and R_3 levels. The same suggestion was made by Hersey and Blanchard (2013, p. 205). Hambleton and Gumbert (1982) reported that about 90% of high-performing leaders in their study used S_2 and S_3 styles. Analysis of data presented by Goodson *et al.* points in the same direction. Their two first quartiles—ostensibly R_1 and R_2 —showed essentially the same pattern, that the Selling style produced the best outcomes. Their third and fourth quartiles showed preference for Participating style. It can be argued that based on presented data, most of the participants found themselves on R_2 and R_3 levels of readiness that fit S_2 and S_3 styles, respectively. An arbitrary method of establishing levels of readiness produces results that can significantly distort the real picture. A self-evaluation method of assessment of abilities and willingness is also far from being accurate.

Similar limitations can be found in the Blank *et al.* (1990) study that also found very limited support for the SLT theory. The only outcome variable that differed significantly between “fits” and “misfits” in relation of maturity and style was job satisfaction of followers. Maturity levels were created in this study by the same method—dividing maturity measures into quartiles. In addition, the study substitutes experience in terms of years of employment for abilities. Analysis of reported data shows that experience does not correlate with any of the outcome variables, even on zero-level. It is highly unlikely that skills and knowledge do not affect performance and other outcomes. The authors themselves admit that the instrument for measuring skills is far from being effective. On the positive side, an advantage of empirical methods used in this study is that maturity level was established through the more valid peer evaluation approach rather than through self-reports.

Vecchio, in several consecutive studies (1987; Norris & Vecchio, 1992; Thompson & Vecchio, 2009), created three categories of maturity (readiness). In order to establish levels of readiness, the recipients' population was trichotomized. This is essentially the same approach as used by Goodson *et al.* (1989) and Blank *et al.* (1990), with the one difference that there are three categories of maturity rather than four. First, it is not clear why the author consistently prefers three levels of maturity/readiness despite the fact that Hersey and Blanchard's theory differentiates between four levels of readiness and four styles. Combining R₂ and R₃ level employees in some intermediate group does not seem justified. According to the model, Selling and Participating are substantially different styles of leadership. For instance, when a leader chooses to apply participation of followers and convenes a meeting to discuss certain ideas and make group decisions, it is ill-advised for the leader to “sell” the ideas to employees. This will reduce their commitment to decisions made at this meeting. Instead, this leader should use input from employees in order to generate high-quality alternatives and make a sound choice.

Second, for the reasons explained earlier, trichotomization creates artificial categories that do not necessarily reflect the real level of followers' readiness. At the same time, a significant advantage of Vecchio's approach is that he does not employ self-evaluation: maturity is assessed by managers, and leadership style is evaluated by employees, which makes assessment of these variables more valid. The results of his studies provide partial empirical support to the model, especially at low/moderate levels of readiness. In earlier studies (1987, 1992) results were insignificant though in prescribed direction—styles that matched the model generated somewhat higher results in terms of performance, leader-member exchange quality, and satisfaction with supervision. The plausible reason for the lack of significant relationship (mentioned by Vecchio and coauthors) was the small number of “matches” between readiness and prescribed styles. This argument was echoed by Cairns *et al.* (1998), who obtained similar results. In the latest study (Thompson & Vecchio, 2009), a bigger sample size was used, and differences between matches were not only in the “right” direction but also significant for all outcome variables. As far as the highest level of readiness is concerned, empirical results (based on LBDQ instrument by

Stodgill and Coons, 1957) did not support the model's predictions—researchers did not find that prescribed Low Task/Low People style fit the highest maturity level. However, using a different instrument, researchers showed that employees with the highest level of experience responded more positively to leaders who provided greater autonomy.

Papworth *et al.* (2009) used a creative method of content analysis of transcripts of successful leader/follower interactions and found that as readiness of followers increased, supervisors appeared to become less dominant in their supervision styles. This result demonstrates that as readiness grows, task-orientated behavior of a leader becomes weaker and clearly is in line with prescriptions of the model. At the same time, the authors maintain that their findings provide only limited support to SLT since the majority of the specific speech behaviors are not in accordance with the model. For instance, according to the theory, an R_1 level supervisee should elicit more guiding, telling or directing behaviors than an R_2 level supervisee, which was not supported by empirical data. This conclusion seems doubtful since readiness levels in this study, as in some others, were established in a rather arbitrary fashion. R_1 , for example, was associated with the novice post-graduate specialist who lacks professional training; R_2 described individuals who were rapidly developing skills and expertise. An argument can be made that the former fits R_2 rather than the R_1 level while the latter approximates R_3 level of readiness. It is reasonable to assume that a new untrained employee who lacks specific knowledge but is willing to improve skills matches the description of the second maturity level and should be handled through explaining and persuading rather than control and instruction. If we upgrade levels of readiness, specific supervisors' behaviors will be in line with predictions of the model (e.g. questioning and explanation as elements of S_2 style fit R_2).

Silverthorne (2000), and later Silverthorne and Wang (2001), found that leaders who used more adaptive and flexible styles were rated as more successful than more rigid traditional managers, and that companies under their leadership demonstrated lower absenteeism and turnover and higher profitability. Although these studies did not specifically address SLT, results suggest the effectiveness of the contingency approach in leadership. It can be argued that SLT puts more emphasis on a leader's flexibility and is more aligned with these findings than other contingency leadership models such as Fiedler's contingency theory (1967). The latter advocates placement of a leader with a certain style in a situation that fits such a style.

Ralph (2004) utilized the situational leadership model as a foundation for a contextual supervision method developed by the author in the mentor-protégé context. He argued that application of this method increased the degree of matches between followers' developmental level and leaders' task and support behaviors and that when the match was achieved, followers and leaders experienced lower level of interpersonal conflicts. In a more recent study Luo and Lio (2014) tested the impact of leader-follower match based on SLT in the Chinese context. The authors found positive impact of matches between the leader's style and followers' readiness on employees' organizational citizenship behavior.

The overall conclusion is that despite certain scarcity of research devoted to this area, there is a growing empirical evidence of validity of Hersey-Blanchard's theory. This study endeavors to add to such evidence.

Theoretical criticism

The most detailed theoretical critique of the model was provided by Graeff (1983; 1997). He argued that the theory's weaknesses "include internal inconsistencies, conceptual ambiguity, incompleteness, and confusion associated with the multiple versions of the model" (1997, p. 153). He also alleged the model to be a fad that uses catchy buzzwords such as competence and commitment (ibid., p. 156). Along with valid points, such categorical criticism seems to be somewhat one-sided. First, the tone of the discussion is far from being objective. For example, sections in the article (Graeff, 1983) are titled "Problems with the normative model" and "Problems with the Lead Instrument". It would seem less judgmental if more impartial terms such as "Critical Analysis" had been used. In another instance, Graeff, in an analysis of a situation presented in Hersey-Blanchard's LEAD instrument where motivation of employees deteriorates, insists on using the construct of instrumentality (performance-outcomes link) from expectancy theory (Vroom, 1964) because its clarification "might provide the greatest leverage for restoring motivational level" (p. 288). He further argues that clarifying instrumentality fits high task/low relationship style S_1 , in contrast with the LEAD instrument which suggests using S_2 high task/high relationship style in this state of affairs. It is not quite clear why, of three links presented in the expectancy model—expectancy, valence of outcome, instrumentality—the latter is necessarily the paramount way to resolve the problem of deteriorating employee motivation. Any of these variables may be manipulated in order to restore motivation. But even if instrumentality is the most useful approach, it is by no means obvious that its clarification is commensurate with high task/low relationship style. It can be argued that it is more in line with high task/high relationship, S_2 Selling style, where significant persuasion has to be used. Therefore, S_2 style appears to fit the circumstances and is correctly recommended for this situation in the LEAD instrument. The categorical claim that "accepting the expectancy argument makes the Situational Leadership Theory much less useful diagnostically" (Graeff, 1983, p. 288) seems unfounded.

Second, accusations of faddishness are far-fetched. Unlike many fads in managerial science that came and vanished rather quickly, Hersey and Blanchard's model demonstrates considerable robustness and sustainability and remains popular after more than 30 years. As Thompson and Vecchio indicate, "the theory has stood the test of time in the marketplace of leadership training programs, as it is well-known and commonly used for training leaders" (2009, p. 838). A study by Avery and Ryan (2002) in Australia demonstrated that managers find the theory understandable and easily applicable, and perceived to be relevant to managerial roles.

Third, impartial analysis should apply the same level of precision in analyzing both sides of the discussion—the model itself and the critique of this model. Graeff (*ibid.*) uses a very high level of “resolution” while examining imperfections in Hersey and Blanchard’s model; there is no such precision when the critique is addressed. As an example, Graeff shows in detail the differences between concepts of commitment and competence in SLTII (Blanchard, 1985), versus willingness and abilities in SLT, as a proof of inconsistencies between different versions of the model. At the same time, he also incorporates a previously analyzed empirical study by Blank *et al.* (1990) that uses years of experience as a substitute for abilities in measuring followers’ readiness. As was mentioned earlier, experience is only one of the components of abilities; consequently such a substitution is problematic. The difference between competence and abilities would seem significantly smaller than between abilities and experience. It is quite easy to imagine an inexperienced employee with a high level of aptitude who performs better than a worn-out experienced employee with mediocre skills. The questionable methods of using self-reports and quartiles (Goodson *et al.*, 1989; Blank *et al.*, 1990) in establishing readiness levels and using three categories for readiness rather than four prescribed by the model (Vecchio, 1987), were among other weaknesses of empirical studies presented earlier. That is not to say that the model does not have theoretical drawbacks. Among weaknesses of the model the following are probably the most salient:

1. There is a certain inconsistency between SLT (Hersey & Blanchard, 1977; Hersey *et al.*, 1996, 2013) and SLTII (Blanchard *et al.*, 1985). Indeed there is a difference between concepts of maturity and development, willingness and commitment, ability and competence. There are considerable differences between descriptions of employees’ level of development in the 1985 version and maturity/readiness levels in the 1977 and 1996/2013 versions. Graeff (1997), for the most part correctly, pinpoints these inconsistencies. At the same time, one can argue that SLTII is the version that was developed by only one of the coauthors—K. Blanchard without P. Hersey. There is very little inconsistency between the earlier version of Hersey and Blanchard’s theory (1977) and more contemporary versions (1996, 2013). The only significant difference is that the rather pejorative term “maturity” was replaced by the more neutral “readiness”. This change seems warranted since Hersey and Blanchard emphasized the task-specific nature of follower’s level. Maturity appears to be a less specific term than readiness. So if one refers to the model of Hersey and Blanchard, the inconsistencies with the version developed by only one author may become less critical.

2. Hersey and Blanchard define levels of readiness in a dichotomized way which is rather simplistic. In particular, there are lapses in logic in descriptions of a passage between R_2 (willing-unable) and R_3 (able-unwilling) levels of readiness

where ability and willingness move in the opposite directions (Graeff, 1983). Hersey and Blanchard define an R_3 employee as able but unwilling because he or she is tired or angry with a manager or insecure in a new task (Hersey & Blanchard, 2013). These are probable situations, but they do not exhaust all the possibilities of dynamics of readiness. The logical argument can be made that when a follower possesses a reasonably high (but not the highest) level of both elements—knowledge and motivation—his readiness would also be reasonably high, or R_3 . Thus, when an initially motivated rookie employee (R_2 level) acquires a certain amount of knowledge and skills, he moves to the next level of readiness (R_3) rather than jumping to R_4 immediately. He is not a complete expert in his field, but his knowledge and experience make it plausible to elicit his input in decision-making, which will constitute the S_3 Participating style. In addition, changes in opposite directions in ability and willingness are possible but cannot be taken as a general rule.

3. The LEAD instrument developed by Hersey and Blanchard was not validated empirically and “has unknown psychometric qualities” (Vecchio, 1987). As a result most studies devoted to empirical verification of SLT use other methods that have their own limitations. The most common instrument for measuring task and relationship-oriented behavior is the LBDQ -XII survey by Stodgill and Coons (1957). Unlike LEAD, which is a customized tool and describes readiness and style in specific situations, LBDQ-XII is more general and addresses over-all leader behavior at the “macro” level. This difference in scale may generate misleading empirical results. As an example, Delegating style S_4 , characterized by low task/low relationship behaviors, can be used in specific assignments by a manager alongside overall provision of structure. Therefore it is quite possible that an R_4 employee who has considerable authority and leeway at work still may respond by assigning high values for such items from LBDQ-XII as: a manager lets group members know what is expected of them; makes attitudes clear to the group; maintains definite standards of performance, etc. In fact, most of the items in the Structure section of LBDQ-XII do not reflect delegating authority to followers. Such impreciseness might be a reason for certain inconsistency of empirical outcome generated in research. As mentioned earlier, in several studies presented above, Vecchio (1987, 1992, 2009) found no match between maturity and style on high level of maturity. So SLT, according to these outcomes seems not to work on this level. At the same time, using another instrument (2009), researchers showed higher propensity towards autonomy (which can be regarded a proxy for delegating) among employees with higher maturity.

By the same token, the low structure/high consideration quadrant in the LBDQ-XII questionnaire does not necessarily represent S_3 Participating style. The content of items is closer to the Country Club Management from Managerial Grid framework (Blake & Mouton, 1985) than to participation in decision-making and teamwork—a manager is friendly but does not explain much. If

the LEAD instrument had been improved and statistically validated, serious drawbacks in empirical research dedicated to SLT might have been avoided.

Issues with the validity of SLT

Vecchio (1987) pointed out that Hersey-Blanchard's model overlaps with several other models such as Maslow (1954), Herzberg (1966), Argyris (1957), etc. but this does not add to its validity because these models were not confirmed empirically. Thus it is reasonable to maintain that if we establish overlap of SLT with the models that were confirmed empirically, it will render theoretical support to situational theory. The later paper (Thompson & Vecchio, 2009) alludes to such a model: leader-member exchange (LMX). According to the authors, this theory can explain the relation between experience and autonomy that was presented earlier: "As job experience increases, a subordinate will gain in job expertise, and both autonomy and LMX will likely increase at the same rate. . . High LMX relationships, in turn, typically involve greater delegation and empowerment" (p. 845). However, Thompson and Vecchio imply that these connections detract from the value of Hersey-Blanchard's model because another theory provides satisfactory explanation of discussed managerial phenomena. This is not a convincing argument and it is logically inconsistent with the argument made by Vecchio in the earlier paper (1987). If a leader's behavior can be properly explained by two models and one of them is empirically confirmed, doesn't that in fact add to the value of the second?

Another model that has obvious similarities with the theory under discussion is the model of Vroom and Yetton (1973) which was later revised by Vroom and Jago (1988). This framework addresses the optimal level of followers' participation in decision-making depending on various factors; it was largely supported in empirical research (e.g., Field & House, 1990). According to this framework, if followers do not possess information, are not committed to the particular decision, and do not share organizational goals to be attained by the solution, managers have to use an autocratic style, which is just another term for Telling S_1 style. If these variables are reversed, a leader has to use more participative style. In the later version of Vroom and Jago (1988), a leader can also delegate authority for making a decision to followers when the expertise of a group is high. Knowing information clearly contributes to employees' abilities in solving problems; commitment to decisions and sharing organizational goals support their willingness. The difference between two contingency frameworks is that Vroom's model pinpoints particular decisions while Hersey-Blanchard's theory addresses more prolonged utilization of certain leadership styles.

Other theoretical frameworks that are aligned with SLT are components of organizational structure such as span of control and decentralization. One of the factors defining span of control is skills of subordinates—the higher they are, the wider the span of control (Davison, 2003; Van Fleet & Bedeian, 1977). The rationale for this is that trained employees do not require close supervision.

That logic overlaps with the SLT tenet that followers with high readiness do not need high task management style. Along similar lines, decentralization is associated with higher intrinsic motivation, demonstrating personal skills and facilitating learning and creativity (Jones, 2004; Hirst *et al.*, 2011). That also coincides with the logic of SLT. In sum, there are several solid theoretical constructs that overlap with the discussed framework and provide theoretical support to it.

Hypotheses

The curvilinear relationship postulated by SLT has a twofold meaning:

1) leadership style has to change according to readiness; and 2) the task of a manager is not to stay on the same level of readiness, but to try to enhance it. The rationale for the latter is that people on higher levels of readiness are more productive and self-fulfilling. One can predict a higher level of performance and satisfaction among employees with high readiness. Enhancing of readiness underlines the dynamic nature of the model and the need to move from lower to higher styles. Thus the following hypotheses can be formulated:

Hypothesis 1a: Performance is positively related to readiness.

Hypothesis 1b: Satisfaction is positively related to readiness.

Previous empirical studies did not provide unequivocal support for the central tenet of the model that leadership style should match level of readiness. The intent of the present research is to verify the soundness of this claim. One can postulate that the closer the actual style to the apt one, the better employee performance and satisfaction will be. This implies that deviation from a prescribed style would lead to lower employee performance and satisfaction. Furthermore, the bigger the deviation from proper style, the lower the performance will be. Consequently, we can formulate the hypotheses:

Hypothesis 2a: Performance is negatively related to the deviation from the apt leadership style given the readiness level.

Hypothesis 2b: Satisfaction is negatively related to the deviation from the apt leadership style given the readiness level.

Thompson and Vecchio established that respondents with high experience had a more positive response towards autonomy, which can serve as a proxy for delegating. Since it can be expected that most respondents find themselves in intermediate levels of readiness R_2 and R_3 , participants with higher levels of readiness should express even stronger propensity towards participative style S_3 . Based on these arguments, we hypothesize that:

Hypothesis 3a: Most respondents find themselves on intermediate levels of readiness— R_2 and R_3 .

Hypothesis 3b: Participative practices are positively related to readiness.

Methods

Sample

The subjects for this study were 155 graduate business school students at a state university in the Northeast. All the participants were working, 61% full time, 39% part-time; 79% were working in service organizations while 21% were working in manufacturing companies; 55% were females. Mean age of participants was 28 years, average work experience was 7 years; 42% worked as professionals, 18% were managers, and others were in clerical/administrative and service positions. The choice of a college student sample is apparently one of convenience, but the MBA student population reflected a variety of types of organizations and leadership styles.

Measures

Readiness

Components of readiness—willingness and ability—were measured by seven and six items, respectively. Both scales represent a compilation of items selectively adopted from previous studies (Blank *et al.*, 1990 (e.g. “I do extra on the job”); Wagner, 1975 (e.g. “If the work were only more interesting I would be motivated to perform better”)) and developed for this research (e.g. “I am enthusiastic about the job”). Response categories for the foregoing variables ranged between 1 = strongly agree and 4 = strongly disagree. Readiness was calculated as a sum of means of willingness and abilities. Consequently, its scale range is between 2 and 8. Levels of readiness were established by dividing this scale into four equal portions. The measure for readiness demonstrated sufficient reliability with Cronbach alpha at 0.707.

Leader’s style

The LEAD instrument developed by Hersey and Blanchard has unknown validity. LBDQ-XII, developed by Stodgill & Coons (1957), is the most used and is a well-known tool for measuring leaders’ behaviors. Therefore, despite certain reservations explained earlier in the paper, it was decided to use this instrument in the present study as well. Two dimensions of leadership behaviors are measured through scales of initiating structure and consideration of LBDQ. Cronbach alpha for these scales were 0.926 for structure and 0.885 for consideration.

Satisfaction

Satisfaction was measured by 18 items of Satisfaction with Supervision scale from Job Descriptive Index (JDI) instruments (Smith *et al.*, 1969). This instrument was also used in previous studies (Blank *et al.*, 1990; Vecchio 1987). Cronbach alpha for the scale was 0.930.

Performance

Four items were created to measure perceived performance (e.g. "I am not successful at this work"). Response categories ranged from 1 = strongly agree to 5 = strongly disagree. Cronbach alpha for the scale was 0.710.

Participation

Three items were used to measure participation (e.g. "Manager encourages involvement in decisions"). Response categories ranged from 1 = strongly agree to 6 = strongly disagree. Cronbach alpha for the scale was 0.902.

Distance from apt leadership style

Our proposition predicts that the more the observed leadership deviates from the optimum leadership style, the worse the performance will be. Therefore, the key measurement to test the central proposition is the distance of actual style from optimal or apt style. To measure such distance, first an optimal leadership style was designated for each level of readiness (R). For lowest readiness R=1, there is no case since there are no participants that found themselves on this level. Since the sample included graduate students, which implies a certain level of maturity, this data appears genuine. For R=2, the best leadership style is "high task, high people". The curve of the model tends to approximate centers of cells of prescriptive styles rather than their corners (Hersey *et al.*, 2013). This implies that the model involves moderately low and high values rather than extremely low and high values. Both task T and people P orientation range from 1 to 6. Consequently, optimal styles would contain middle of the range for both low and high values of T and P, i.e. 2 for low (1-3) and 5 for high (4-6) values. Therefore, the optimum leadership at R=2 is designated as (T=5, P=5); for R=3 level, the optimum leadership "low task orientation, high people orientation" is designated as (T=2, P=5); for R=4 level, the optimum leadership is "low task and low people orientation", and thus designated as (T=2, P=2). For each observed case in the sample, a "deviation from best leadership" is calculated as distance=square root [(observed T - apt T at this level)² + (observed P - apt P at this level)²]:

$$\text{For } R_2, \text{ distance} = \sqrt{(T-5)^2 + (P-5)^2}$$

$$\text{For } R_3, \text{ distance} = \sqrt{(T-2)^2 + (P-5)^2}$$

$$\text{For } R_4, \text{ distance} = \sqrt{(T-2)^2 + (P-2)^2}$$

Where R is readiness level, T is task orientation, and P is people orientation.

Results

Table 1 shows the means, standard deviations, and correlations among the study variables.

Table 1 Means, standard deviations, correlation coefficients among major variables

		Readiness	Tasks	Relationship	Performance	Satisfaction	Participation
Mean		5.928	4.393	4.225	4.063	3.711	3.634
Std deviation		.657	1.011	.962	.658	.765	1.035
Readiness	Pearson correlation	1	.104	.246**	.404**	.212**	.290**
	Sig.		.197	.002	.000	.008	.000
Tasks	Pearson correlation	.104	1	.610**	.379**	.622**	.370**
	Sig.	.197		.000	.000	.000	.000
Relationship	Pearson correlation	.246**	.610**	1	.537**	.846**	.736**
	Sig.	.002	.000		.000	.000	.000
Performance	Pearson correlation	.404**	.379**	.537**	1	.598**	.543**
	Sig.	.000	.000	.000		.000	.000
Satisfaction	Pearson correlation	.212**	.622**	.846**	.598**	1	.682**
	Sig.	.008	.000	.000	.000		.000
Participation	Pearson correlation	.290**	.370**	.736**	.543**	.682**	1
	Sig.	.000	.000	.000	.000	.000	

** Correlation is significant at the 0.01 level (2-tailed).

H_{1a} and H_{1b} propose that performance and satisfaction are positively related to readiness. These hypotheses were tested by performing two regression models, regressing performance and satisfaction on readiness. Results are reported in Table 2.

Table 2 Regression results: performance and satisfaction on readiness

Model	Dependent variable	Adjusted R square	F value	Independent variables	Standardized coefficients	t	Sig.
1	Performance	.157	29.789	(Constant)		3.775	.000
				Readiness	.404**	5.458	.000
2	Satisfaction	.039	7.183	(Constant)		4.102	.000
				Readiness	.212**	2.680	.008

** coefficient is significant at .01 level (two sided)

* coefficient is significant at .05 level (two sided)

As Table 2 shows, readiness significantly and positively contributes to both performance and satisfaction. Therefore, lending support to H1a and H1b.

H2a and H2b suggest that performance and satisfaction are negatively related to the deviation from the proper leadership style given the readiness level. For each level of readiness we defined both the proper style and a distance from the proper leadership style for each observation as described before. This study used regression tests of performance and satisfaction on three explanatory variables—readiness, distance from the desired style, and the interaction between readiness and distance. The interaction item is included because the authors suspect that the readiness level may modify the relation between the dependent variables (performance and satisfaction) and the deviation from the desired leadership. The results of tests are shown in Table 3.

Table 3 Regression results: participation and satisfaction on readiness, distance, and readiness * distance

Model	Dependent variable	Adjusted R square	F value	Independent variables	Standardized coefficients	t	Sig.
1	Performance	.170	11.530	(Constant)		3.472	.001
				Readiness	.126	.748	.456
				Distance	-1.108*	-2.079	.039
				Distance * readiness	1.247*	2.028	.044
2	Satisfaction	.143	9.561	(Constant)		5.928	.000
				Readiness	-.445**	-2.598	.010
				Distance	-2.453**	-4.530	.000
				Distance * readiness	2.827**	4.524	.000

** coefficient is significant at .01 level (two sided).

* coefficient is significant at .05 level (two sided).

The adjusted R square is 17.0% with F value at 11.530 for the performance regression. For the satisfaction regression, the adjusted R square and F value are 14.3% and 9.561, respectively. Overall, both regression models are significant. The coefficient for distance is significantly negative, as expected. In other words, the more deviation from the desired leadership at each particular readiness level, the lower the level of performance and satisfaction. The significant and negative effect of distance from prescribed leadership on the dependent variables lends support to both H2a and H2b.

There are interesting results for the coefficients of the interaction items. The interaction distance * readiness is positive and significant. It indicates that readiness can modify the hypothesized relation between appropriate leadership and performance (and satisfaction): although deviation from the optimum best leadership given a readiness level adversely affects performance, the readiness

itself can weaken this relation. More explicitly, at high levels of readiness, the effect of deviation from best leadership style is mitigated by readiness and is less harmful to performance and satisfaction. The plausible explanation is that employees at this level have stronger intrinsic motivation and abilities that compensate partially for errors in management.

Hypothesis 3 proposes that readiness is positively related to participation. Table 4 shows that the Pearson correlation coefficient between the two variables is .290 and significant on .01 level. Table 5 also shows that in the simple linear regression model, readiness is positively related to participation, and this relation is significant at the .01 level.

Table 4 Correlation between readiness and participation

		Participation
Readiness	Pearson correlation	.290**
	Sig. (2-tailed)	.000
	N	155

** Correlation is significant at the 0.01 level (2-tailed).

Table 5 Regression results: participation on readiness

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.
		B	Std error	Beta	B	Std error
1	(Constant)	.932	.726		1.284	.201
	readiness	.456**	.122**	.290**	3.743	.000

Adjusted R square is .078, F value is 14.010.

This result suggests that the more ready the employees are to perform, the more likely it is that they are encouraged to participate in decision making. Therefore, hypothesis 3 is confirmed.

Discussion

There is a certain discrepancy that can be labeled a fundamental paradox of Hersey-Blanchard’s theory. On the one hand, this model is intuitively appealing, and almost self-evident. Indeed is it reasonable to delegate authority and discretion (S_4 style) to a person who is incapable of carrying out an assignment independently? Does it make sense to solicit input from an employee (S_3 style) who does not possess knowledge of an issue? Is it appropriate to give specific instructions (S_1 style) to an expert who might know particular details better than a manager? On the other hand, there appears to be a consensus in the literature of a dearth of solid empirical evidence supporting the model. This study

addresses this paradox and contributes to the field of contingency leadership by utilizing an innovative approach in measuring the effects of fit between employees' readiness and leadership style and providing empirical validation of SLT.

The hypothesized relationships were confirmed statistically. We found that there is a significant relationship between readiness and performance and satisfaction. Outcome variables were positively related to both abilities and willingness. This implies that readiness matters for managerial practices. SLT does not include additional situational variables besides readiness, and constitutes a simpler framework than other contingency models. Consequently it is more easily understood and applied by practitioners (Avery & Ryan, 2002). However, such simplicity is a double-edged sword. Disregarding other situational factors raises questions of the model's ecological validity and has drawn criticism of oversimplification (Goodson *et al.*, 1989; Vecchio *et al.*, 2006). Our findings reinforce the position that abilities and motivation are significant factors that should be taken into account in creating an apt leadership style. They also emphasize the dynamic nature of the model and the benefits of enhancing both components of readiness and gradually moving from lower to higher styles.

The central empirical result of this study is that deviation from an apt style impairs performance and satisfaction. The closer a leader's behavior is to the prescriptions of the model, the better both outcome variables will be. For measuring effect of "fit" of a leader's style to employees' readiness this research employed a method based on measuring distance between proper style and actual style. Our method is an alternative and less subjective way of defining "second best", "third best" and the "worst" style described in the LEAD instrument (Hersey and Blanchard, 1977).

Our approach seems to have advantages over techniques used previously in the literature. For instance, method of comparison between matches and mismatches (Cairns *et al.*, 1998; Vecchio, 1987; Thompson and Vecchio, 2009) contains two potential weaknesses. First, in most previous studies, the number of matches was rather small, which resulted in insignificant outcomes, although the differences were mostly in a predicted direction. The low frequency of "matches" was mentioned as a difficulty in conducting tests of SLT (Cairns *et al.*, 1998; Thompson and Vecchio, 2009). Second, the term "mismatch" pools all the "improper" styles in one category regardless of the distance from the proper style. For instance if the best style is S_1 Telling, all other styles would be labeled as "mismatch". However, S_2 Selling is much closer to S_1 than S_4 Delegating. Since the magnitude of an error is significantly smaller when employing S_2 rather than S_4 , performance is supposed to be higher in the former case. These differences are not revealed by simple comparison between matches and mismatches; hence, some information is lost. Another method—hierarchical regression analysis using three way multiplicative interaction product task * consideration * maturity/readiness (Goodson *et al.*, 1989; Vecchio, 1987)—also misses the effect of distance. In addition such an approach has another possible drawback. In this formula maturity/readiness is supposed to moderate the effect of style on performance, and task * consideration is supposed to reflect leader's style. The combination high task * low consideration will produce the

same numbers as low task*high consideration. However these identical numbers will stand for very different styles: S_1 Telling and S_3 Participating, respectively. Therefore the meaning of this triple multiplication is not clear.

As we expected, most respondents found themselves at intermediate levels of readiness R_2 and R_3 . This supports theoretical considerations and empirical results of previous research (Cairnes *et al.*, 1998; Hersey *et al.*, 2013). Due to the fact that the majority of respondents tend to be closer to the middle point in terms of readiness, we can express further methodological reservation with establishing readiness/maturity levels through quartiles (Blank *et al.*, 1990; Goodson *et al.*, 1989) or tricycles (Vecchio, 1987), the method that was criticized earlier in this paper. Such techniques create arbitrary categories that do not accurately reflect real levels of followers' skills and willingness. As a result, the whole notion of fit and misfit between style and readiness may become distorted.

Our research found a significant positive relationship between readiness and participative practices. This study utilized a separate scale for measuring participative practices. The reason for developing a new scale was a concern that due to the more general character of the LBDQ –XII instrument (Stodgill & Coons, 1957), it might not capture specific leaders' behaviors prescribed by the model. For instance, low task/high consideration behaviors in this survey do not necessarily overlap with participating practices and reflect S_3 Participating style. As we expected, since the majority of respondents found themselves in intermediate levels of readiness, those with higher readiness expressed stronger propensity towards participation. This represents a fit between R_3 readiness and S_3 style, and constitutes an empirical validation of Hersey-Blanchard's model. Our result is in line with previous findings by Thompson and Vecchio (2009) that followers with higher experience showed more positive response towards autonomy. Both studies demonstrate that: 1) followers with high readiness have a preference for high level styles; and 2) using more specific scales than LBDQ-XII produces consistent results in favor of SLT. One possible explanation, which should be tested in future research, is that the LBDQ-XII instrument is too general. An employee can have substantial discretion, provide inputs to decisions, and be engaged and participative, but still need a certain degree of structure from a manager.

Limitations and future research

A few limitations should be taken into account in interpreting our results. First, self-assessment of readiness is far from being optimal. An improved approach would be one where readiness is evaluated by managers while leadership style is assessed by subordinates. Since the respondents in this study were MBA students working in different organizations, such an approach was not possible. A more sophisticated design of the study that would involve such cross-evaluation is warranted and is clearly a venue for future research.

Second, the use of cross-sectional methodology prevents definitive conclusions regarding causal relationships. As in most studies in the field, all variables are measured at the same time. However, if a manager uses a style which matches or mismatches readiness, favorable or unfavorable outcomes would not appear immediately. A longitudinal study would be beneficial in order to confirm the cause-effect relationships. The challenging task is to establish the apt size of the lag between measurements of independent variables (readiness and style) and outcome variables (performance and satisfaction) in order to capture effects of matches and mismatches.

Another challenging and fruitful venue for future research would be the development and validation of an instrument for measurement of leaders' behaviors that is more specific than the LBDQ-XII questionnaire and more acceptable to researchers than the LEAD survey. In our opinion, the lack of such an instrument is an impediment to empirical confirmation of Hersey-Blanchard's model, a task long overdue.

References

- Aldoory, L., & Toth, E. (2004). Leadership and Gender in Public Relations: Perceived Effectiveness of Transformational and Transactional Leadership Styles. *Journal of Public Relations Research*, 16 (2), 157-183.
- Argyris, C. (1957). *Personality and organization*. New York, NY: Harper and Row.
- Avery, G., & Ryan, J. (2002). Applying situational leadership in Australia. *Journal of Management Development*, 21(4), 242-262.
- Avolio B., Sosik J., Jung D., & Berson Y. (2003). Leadership models, methods, and applications: Small steps and giant leaps. In W. Borman, R. Klimoski, D. Ilgen, & B. Weiner (Eds.), *Handbook of psychology* (Vol. 12, 277-307). New York, NY: John Wiley & Sons.
- Blake, R., & Mouton, J.S. (1985). *The Managerial Grid III*. Houston, TX: Gulf.
- Blanchard, K., Zigarmi, P., & Zigarmi, D. (1985). *Leadership and the One Minute Manager*. New York, NY: Morrow.
- Blank, W., Weitzel, R., & Green, S. (1990). A test of Situational Leadership Theory. *Personnel Psychology*, 43(3), 579-597.
- Cairns, T., Hollenback, J., Preziosi, R., & Snow, W. (1998). Technical note: A study of Hersey and Blanchard's Situational Leadership Theory. *Leadership & Organization Development Journal*, 19(2), 113-116.
- Davison, B. (2003). Management span of control: How wide is too wide? *Journal of Business Strategy*, 24(4), 22-29.
- Derue, D.S., Nahrgang, D.H., Wellman, N., & Humphrey, S.E. (2011). Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. *Personnel Psychology*, 2011, 64, 7-52.
- Fiedler, F. (1967). *A theory of leadership effectiveness*. New York, NY: McGraw-Hill.
- Field, R., & House, R. (1990). A test of Vroom-Yetton model using manager and subordinate reports. *Journal of Applied Psychology*, 75(3), 362-366.
- Goodson, J., McGee, G., & Cashman, J. (1989). Situational Leadership Theory: A test of leader prescriptions. *Group and Organizational Studies* 14(4), 446-461.
- Graeff, C. (1983). The Situational Leadership Theory: A Critical View. *Academy of Management Review*, 8(2), 285-291.

- Graeff, C. (1997). Evolution of Situational Leadership Theory: A critical review. *Leadership Quarterly*, 8(2), 153-170.
- Hambleton, R., & Gumpert, R. (1982). The validity of Hersey and Blanchard's theory of leader effectiveness. *Group and Organization Studies*, 7(2), 225-242.
- Hersey, H., & Blanchard, K. H. (1977). *Management of organizational behavior: Utilizing human resources*, 3rd ed., Englewood Cliffs, NJ: Prentice Hall.
- Hersey, H., Blanchard, K., & Johnson, D. (1996). *Management of organizational behavior: Utilizing human resources*, 7th ed., Upper Saddle River, NJ: Prentice Hall.
- Hersey, H., Blanchard, K., & Johnson, D. (2013). *Management of organizational behavior: Leading human resources*, 10th ed., Upper Saddle River, NJ: Prentice Hall.
- Herzberg, F. (1966). *Work and the nature of man*. New York, NY: World Publishing.
- Hirst, G., Van Knippenberg, D., Chin-Hui, C., & Sacramento, C. (2011). How does bureaucracy impact individual creativity? A cross-level investigation of team contextual influences on goal orientation-creativity relationships. *Academy of Management Journal*, 54(3), 624-641.
- Jones, G. (2004). *Organizational theory, design and change*. 4th ed. Upper Saddle River, NJ: Pearson/Prentice Hall.
- Luo, H., & Lio, S. (2014). Effect of situational leadership and employee readiness match on organizational citizenship behavior in China. *Social Behavior and Personality*, 42(10), 1725-1732.
- Maslow, A. (1954). *Motivation and personality*. New York, NY: Harper and Row.
- Norris, W., & Vecchio, R. (1992). Situational Leadership Theory: A replication. *Group & Organization Management*, 17(3), 331-345.
- Papworth, M., Milne, D., & Boak, G. (2009). An exploratory content analysis of situational leadership. *Journal of Management Development*, 28(7), 593-606.
- Silverthorne, C. (2000). Situational Leadership Theory in Taiwan: A different culture perspective. *Leadership and Organization Development Journal*, 21(2), 68-74.
- Silverthorne, C., & Wang, T. (2001). Situational leadership style as a predictor of success and productivity among Taiwanese business organizations. *The Journal of Psychology*, 135(4), 399-412.
- Smith, P., Kendall, L., & Hulin, C. (1969). *The measurement of satisfaction in work and retirement*. Chicago, IL: Rand-McNally.
- Stogdill, R., & Coons, A. (1957). *Leader behavior: Its description and measurement (Research Monograph No. 88)*. Columbus, OH: Ohio State University, Bureau of Business Research.
- Thompson, G., & Vecchio, R. (2009). Situational Leadership Theory: A test of three versions. *The Leadership Quarterly*, 20(5), 837-848.
- Van Fleet, D., & Bedeian, A. (1977). A history of span of control. *Academy of Management Review*, 10(2), 356-372.
- Vecchio, R. (1987). Situational Leadership Theory: An examination of a prescriptive theory. *Journal of Applied Psychology*, 72(3), 444-451.
- Vecchio, R., Bullis, R., & Brazil, D. (2006). The utility of Situational Leadership Theory: A replication in a military setting. *Small Group Research*, 37(5), 407-424.
- Vroom, V. (1964). *Work and motivation*. New York, NY: John Wiley & Sons.
- Vroom, V., & Jago, A. (1988). *The new leadership: Managing participation in organizations*. Englewood Cliffs, NJ: Prentice Hall.
- Vroom, V., & Yetton, P. (1973). *Leadership and Decision-Making*. Pittsburgh, OH: University of Pittsburgh Press.
- Wagner, F., & Morse, J. (1975). A measure of individual sense of competence. *Psychological Reports*, 36, 451-459.

