Situational leadership theory: a test from a leader-follower congruence approach

Geir Thompson and Lars Glasø
Department of Leadership and Organisational Behaviour, BI Norwegian Business School, Oslo, Norway

Abstract

Purpose – Empirical tests of Hersey and Blanchard’s situational leadership theory (SLT) have demonstrated that the assessment of follower competence and commitment, critical contextual features dictating optimal leadership style, poses several problems in testing the validity of this theory. As well, most previous studies have used leader rating as the only information source in making these assessments. The purpose of this paper is to use the degree of agreement between leader rating and follower self-rating to determine follower competence and commitment, and consequently an optimal leadership style.

Design/methodology/approach – Survey data collected from both supervisors and employees in business organizations in Norway were analyzed to test the predictions put forward in SLT.

Findings – The results show that SLT principles are supported when leader rating and follower self-rating are congruent. However, no support was obtained for Blanchard’s suggestion to apply followers’ self-ratings of competence and commitment in the case of discrepant ratings of follower development level. But data do support the contention that leader assessment would be a better basis for providing followers with appropriate direction and support.

Originality/value – So far SLT has been a very popular but as yet under-researched theory. An important contribution of the present study is by making SLT an evidence-based theory, as opposed to just being intuitively sound. As such, the authors think SLT in an updated “convergent” version should be applied in organizations and taught in leadership development programs.

Keywords Flexibility, Leadership style, Effective leadership, Situational leadership theory,
Partnering for performance

Paper type Research paper

Introduction

Hersey and Blanchard’s situational leadership theory (SLT) is often cited in academically orientated management textbooks, and is among the most widely known leadership theories popular in management training programs and school teacher-training settings. However, it remains among the least-researched views of managerial effectiveness (Thompson and Vecchio, 2009). Considering the enormous popularity of the theory over many decades, strong evidence-based testing of its prescriptive principles is disproportionately lacking. However, recent research on SLT has proposed a new research avenue where SLT predictions are more likely to hold when leader rating and follower self-rating are congruent, rather than when using leader rating alone as in previous studies (Thompson and Glasø, 2015). The present study aims at extending this leader-follower congruence approach to testing SLT.

SLT was first introduced as the “life cycle of leadership” (Hersey and Blanchard, 1969). This version presented SLT’s principles for dealing with newer employees by using more directiveness, and then gradually substituting directiveness with supportiveness as employee seniority increased. In the revised 1977 version, Hersey and Blanchard provided a thorough description of the theoretical foundation for their original model. Since then SLT has undergone a number of cosmetic and substantive changes (e.g., Blanchard, 1988; Blanchard et al., 1985, 1993) and has emerged more recently as a restated set of prescriptive principles (Blanchard, 2007, 2010). In this updated version, follower development level is a crucial moderator of the relationship between leader behavior and efficiency. The moderator
variable comprises follower competence and commitment. Competence is defined as the follower’s task-relevant knowledge, and skills gained through formal education, on-the-job training and experience. Commitment is defined as follower motivation and confidence. Motivation is follower interest in the task, and confidence is follower sense of security or self-assuredness, the extent to which the follower trusts that he or she has the ability to work independently and perform well. Leader behavior consists of two major dimensions of supportiveness and directiveness. Supportiveness refers to the leader displaying warmth and consideration, whereas directiveness refers to the leader initiating structure and monitoring results. Favorable leadership behavior is assumed to shift with changes in follower development level. For inexperienced followers, effective leadership calls for low supportive and high directive behavior. As followers become more competent, the need for directive behavior would be reduced, while the need for more supportive behavior would increase. More specifically, the moderator variable is divided into four categories. Appropriate leadership styles that match follower development comprise directing, coaching, supporting and delegating behavior. For expositional purposes, four combinations of follower development level and leader behavior are outlined: first, followers, low on competence but high on commitment, benefit from directing behavior, which entails low leader consideration combined with high leader structuring; second, followers low to moderate on competence in combination with low commitment, benefit from coaching behavior, which is high leader consideration combined with high leader structuring; third, followers who are moderate to high on competence but show variable commitment benefit from supportive behavior, characterized by high leader consideration combined with low leader structuring; and fourth, followers who are high on both competence and commitment benefit from delegating behavior, where leaders diminish both consideration and structuring behavior.

SLT has made positive contributions to our understanding of dyadic leadership by emphasizing the significance of flexible, adaptive behavior, which has become an essential principle of recent theory and research within the contingency theory tradition, for instance with respect to political skill (Brouer et al., 2013; Yukl, 2010). In addition, SLT has, since it was introduced in the 1970s, stated that it is vital to treat different followers differently, and pointed out that leaders should be aware of opportunities to develop the skills and confidence of followers. This development perspective has been adopted by several leadership theories, such as transformational leadership and servant leadership (Bass and Avolio, 1994; Konczak et al., 2000; Van Dierendonck, 2011). Finally, research has found that the two major dimensions of SLT, leader supportiveness and leader directiveness, are important predictors of employee job satisfaction and ratings of leadership effectiveness (Piccolo et al., 2012). However, despite many revisions of SLT over the years, empirical evidence of the theory’s descriptive accuracy is sparse. In our opinion the authors of the theory have made few attempts to document its validity. Furthermore, research has yielded inconclusive evidence of the theory’s empirical validity and utility in leadership training, probably due to imprecise characterization of leader-follower dynamics (Fernandez and Vecchio, 1997; Thompson and Vecchio, 2009) and ambiguity surrounding the conceptual definition of follower competence and commitment. Previous studies have employed different strategies for measuring these variables. Blank et al. (1990), for example, employed peer rating in order to avoid self-reporting bias, and to ensure the independence of leader behavior descriptions. Self-appraisals of development level have also been applied. Here it was assumed that followers are capable of commenting on their own sense of competence and commitment. The majority of studies, however, have used supervisor rating of follower development level (Vecchio, 1987; Norris and Vecchio, 1992; Fernandez and Vecchio, 1997; Vecchio et al., 2006; Thompson and Vecchio, 2009).
Although some studies differed regarding their use of rating source, they generally had one single source when assessing the moderator variable, i.e., either follower self-report of competence and commitment, or peer or leader assessment of follower competence and commitment. However, there are problems related to the validity of information gathered from rating. For example, self-rating has been shown to suffer from attribution errors, e.g., good performance is believed to result from self-rater behavior, while poor performance results from environmental factors (see Ross, 1977). Such biases are relevant for other ratings as well. For example, rating by peers and superiors may suffer from the halo effect, where positive appraisal of a person in one context might spill over into areas where leaders have little or no factual basis for their assessment (Viswesvaran et al., 2005). The halo effect can seriously influence leaders’ social perceptions, thus increasing the risk of biased diagnoses of follower development level.

To overcome such problems, objective indices of predictors have been suggested as surrogates of development level (Vecchio and Boatwright, 2002). For example, education level and years of experience in the current position have both been used as an index of follower competence (Thompson and Glasø, 2015). Objective indices are assumed to avoid attribution bias, where leaders may make assumptions about one particular attribute based on their judgment of other attributes. However, Thompson and Glasø (2015) reported only modest evidence for the notion that objective indices may be superior to subjective constructs in studying follower reaction to leader behavior. Another approach was suggested by Fernandez and Vecchio (1997), who applied follower job level rather than follower competence and commitment as a predictor of optimal leadership style. Only modest support was obtained for this approach as well.

Clearly more research is needed regarding the assessment of follower development level as described in SLT. Leader-follower congruence, as suggested by Thompson and Glasø (2015), may be a useful approach to dealing with the moderator variable. Measuring the degree of agreement between leader ratings of follower competence and commitment, and follower self-ratings, might be a core issue for assessing the moderator variable. Accordingly, the aim of the present study is fourfold: first, as suggested by Thompson and Glasø (2015), to test SLT’s predictions when leader and follower agree on follower development level. Second, a discrepant assessment of follower development level may occur when leader and follower attempt to come to some agreement. Under such circumstances, Blanchard (2010) suggested using followers’ self-rating as a basis for choosing leadership style. Accordingly, the second aim is to examine the validity of applying follower self-rating of development level for testing the predicted three-way interactions proposed by SLT. Third, a discrepant assessment of follower development level could result from inaccurate and unrealistic self-perception by followers, or from overly modest self-perception of their strengths. Accordingly, we examine whether SLT’s predictions are likely to hold when follower self-rating of competence and commitment are either higher than leader rating, or lower than leader rating. Finally, due to the problematic nature of self-ratings (Atwater et al., 2009), follower self-assessments should not be considered “true scores.” In the case of discrepant assessments, follower development level could be rated by the leader to provide an alternate perspective. Hence, our fourth aim is to examine the validity of applying leader rating of development level for testing SLT’s predicted three-way interactions. Taken together, our study aims to contribute to the situational leadership literature by first applying a congruence approach in an attempt to overcome limitations in previous research to testing the validity of SLT. In addition, we investigate how different strategies for managing discrepant assessment of follower development level may relate to effectiveness outcomes. As such, our examination provides a new avenue for testing SLT prescriptive guidelines.

**Leader-follower congruence approach**

Thompson and Glasø (2015) found that SLT’s predictions are more likely to hold when leader assessment and follower self-assessment of competence and commitment are congruent.
An omnibus test, or overall comparison of matches and mismatches, yielded a significant result for the output variable “performance” in favor of SLT principles. However, the omnibus test may mask unique pattern differences within each development level. Therefore, further comparisons were made within each of the four developmental levels. Support was obtained for SLT predictions for development levels 3 (followers moderate to high on competence but variable commitment benefit from a supportive leadership style), and 4 (followers high on both competence and commitment benefit from a delegating leadership style). Results for development levels 1 and 2 were not obtained due to inadequate sample size for conducting the necessary statistical tests. Accordingly, the present study tests SLT’s predictions for all four development levels, consistent with Fernandez and Vecchio’s (1997) suggestion that SLT’s “predictions are most likely to hold strongly when managers and their subordinates are in agreement concerning individual readiness” (p. 81). Moreover, research has demonstrated that agreement between individual self-evaluation and evaluation by others (e.g., leaders) does relate to individual outcomes such as performance (Bass and Yammarino, 1991; Atwater et al., 1998). In addition, such a test would also be in line with the concept of “partnering for performance” introduced by Blanchard (2010) as an integral component of SLT, where both leader and follower need to come to some agreement on the determination of follower development level. If agreement is reached, then it is assumed that the leader can provide the follower with an appropriate amount of direction and support. This leads to our first hypothesis:

**H1.** SLT’s predictions are more likely to hold when leader rating and follower self-rating of competence and commitment are congruent.

**Discrepant assessments**

Second, if there are gaps between follower self-assessment and leader assessment of follower development level, then differences in these assessments may pose a potential problem for the leader. Followers who do not recognize the discrepancy between their own and their leader’s perceptions may not see the need for more leader direction and/or support. In such case, followers may be less effective because leadership will not be tailored to maximize their motivation and performance. If discrepant assessment of development level is found to reflect different perspectives on the same phenomena, then Blanchard (2010) suggested that follower self-rating of competence and commitment should form the basis for selecting leadership style. Blanchard (2010) argued that a leader should not fight over development level, but rather accept follower self-rating of competence and commitment. But if follower performance does not meet agreed expectations, the assessment should be reconsidered and more direction and/or support from the leader should be given. Based on Blanchard’s suggestion, our second hypothesis is:

**H2.** SLT’s predictions are more likely to hold when follower self-rating of competence and commitment is applied.

Third, when applying follower self-rating as the basis for choosing leadership style, investigation should be undertaken to see whether SLT’s predictions are likely to hold when follower self-rating of competence and commitment exceed leader rating. Research has found over-estimators to produce self-ratings that are significantly higher than leader estimation of their development level due to the tendency to see themselves in an unrealistically positive light (Atwater et al., 2005). Consequently, they see themselves as good performers whilst others see them as poor performers (Yammarino and Atwater, 1997). This notion seems consistent with the Dunning–Kruger effect, in which poor performers seem largely unaware of just how deficient their expertise is. According to Dunning (2011), these deficits leave them with a double burden. First, their incomplete knowledge and
Misperceptions lead them to make mistakes, and second, those same deficits prevent them from recognizing their mistakes and seeing others are choosing more wisely. Inaccurate and unrealistic self-perception may indirectly prevent the follower from receiving the proper amount of direction and support. Hence, we suggest the following hypothesis:

**H3.** SLT’s predictions are not likely to hold when follower self-rating of competence and commitment are higher than leaders’ rating.

Fourth, follower self-rating of competence and commitment may be significantly lower than leader rating. Under-rating may occur because the self-rater either mis-diagnoses his or her strengths or is overly modest (Yammarino and Atwater, 1997). They might think themselves poor or average performers, but others may see them more favorably as successful and effective. They are likely doing well and trying hard to improve, but may not accurately recognize their own strengths and weaknesses (Atwater et al., 2005). This lack of self-awareness may result from overly critical self-evaluation, combined with setting high standards for themselves, continually striving to meet feedback expectations of others, and leading them to pursue easy and unchallenging improvement goals (Tekleab et al., 2008). However, their tendency to overestimate weaknesses may be compensated through hard work, resulting in greater success in task accomplishment (Atwater et al., 1998). Furthermore, positive feedback from others provides them with incentives to undertake corrective action. Hence, we examine whether SLT’s predictions are likely to hold when follower self-rating of competence and commitment fall below leader rating:

**H4.** SLT’s predictions are likely to hold when follower self-rating of competence and commitment are below leaders’ rating.

Finally, because self-assessments are viewed with suspicion due to the likelihood of inflation, assessments obtained from others in the workplace are thought to offer significant added value (Brett and Atwater, 2001; DeNisi and Kluger, 2000). Accordingly, when assessments are discrepant, leaders may provide unique information from an alternative perspective, be more aware of followers’ potential, and know how to develop their skill and confidence. Finally, leader assessment should prevail in discrepant cases because leader action ultimately governs follower development (Fernandez and Vecchio, 1997). Hence, our final hypothesis is:

**H5.** SLT’s predictions are more likely to hold when leader rating of follower competence and commitment is applied.

**Method**

**Sample and procedure**

Data were collected from 168 supervisors and 830 employees from Norwegian for-profit organizations. Business organizations were selected because previous tests of the validity of SLT have in most cases been executed in educational, health care and armed services. Data from profit-oriented firms have the potential to increase our knowledge of whether SLT principles are demonstrably valid in a business setting. Nearly 72 percent of contacted individuals responded, were predominantly male (71.4 percent), with average age of 41.9 years, and an average education of 15.2 years.

Questionnaires were distributed to the respondents while at work. Data were collected electronically and each respondent was given a unique link to the questionnaires. Confirmit was applied as a data collection system. The respondents were informed that the study was conducted solely for academic research purposes and assured of the confidentiality of their responses. Respondents were not compensated for their participation in the study.
The original questionnaires to be used in this study were developed in English. To avoid the risk of misunderstanding or misconception and ensure equivalence of item meaning, the instruments were put through a translation-back translation conversion process (Cavusgil and Das, 1997). All translated questions were field tested. Pilot testing of the instruments with a focus group of five supervisors and five followers indicated that the instruments were relevant in an industrial setting.

**Measures**

Hersey and Blanchard developed the LEAD questionnaire for assessing leadership style. The instrument was designed to provide insight into leaders' perception of their own behavior with respect to style, style range and style adaptability. However, this questionnaire has been severely criticized, as the reliability and validity of LEAD have not yet been established (Graeff, 1983; Lueder, 1985). In the present study, we apply measures of leadership that focus on the behavior of leaders rather than their perceptions and attitudes. Measures of behavior should diminish the possibility of respondent self-deception and be less subject to potential attribution influence and respondent tendencies. Hence, we applied established measures of leader supportiveness and directiveness as indices of key constructs. In order to ensure comparability, the present study relies on measures used in previous studies.

Supervisors provided data on the following instruments: a five-item performance rating scale developed by Liden and Graen (1980), (item stems: dependability, planning, know-how and judgment, overall present performance and expected future performance; anchors, 1 = Unsatisfactory, 7 = Outstanding). Responses to these five items were then summed to provide a measure of performance for each subordinate. Furthermore, each supervisor completed a modified ten-item Employee Readiness Scale (Fernandez and Vecchio, 1997) from an earlier (Blanchard, 1988) Employee Readiness Questionnaire for assessment of subordinate developmental level (sample items for follower competence: “Past job experience,” “Knowledge of the subject area;” sample items for follower commitment: “Willingness to take responsibility,” “Positive work attitude;” anchors, 1 = Low, 7 = High).

Each follower completed a package of questionnaires consisting of LBDQ XII (Stogdill and Coons, 1957), used for measuring supervisor considerateness and structuring behavior. Leader consideration was measured with a five-item scale composed of items taken from the LBDQ-XII instrument (sample item: “My supervisor’s relations with me can be described as friendly and approachable,” anchors, 1 = Never, 2 = Seldom, 3 = Occasionally, 4 = Often, 5 = Always). Leader structuring was measured with five items taken from the LBDQ-XII, using the same five-point response scale for each item, (sample items: “My supervisor schedules for me the work to be done,” and “My supervisor lets me know what is expected of me.”). Also, subordinate self-assessment of development level was measured on a modified ten-item Employee Readiness Scale (Fernandez and Vecchio, 1997) (sample items for follower competence: “Past job experience,” “Knowledge of the subject area;” sample items for follower commitment: “Willingness to take responsibility,” “Positive work attitude;” anchors, 1 = Low, 7 = High).

**Analysis and results**

The three scales that were completed by the supervisors (follower performance, follower competence and commitment) were subjected to confirmatory factor analyses (CFA) using M-Plus, in order to determine whether a three-factor solution better represented the data than a single or general factor. The CFA results for a three-factor solution were: \( \chi^2 = 789.81 \), Tucker-Lewis index (TLI) = 0.97, comparative fit index (CFI) = 0.97 and root mean square error of approximation = 0.078. In contrast, the CFA results for a one-factor solution indicated that the hypothesized three-factor model had significantly better fit over a more parsimonious model where the performance, commitment and competence items were set to
load on a single factor ($\Delta \chi^2 = 63.39, p < 0.001$). Hence, the supervisors’ responses are better described as reflecting three factors.

For the four sets of responses provided by the followers (leader directiveness, leader supportiveness, follower self-reported competence and commitment), CFA was used to contrast a four-factor solution with a single-factor solution. The results for the four-factor model were: $\chi^2 = 523.75$, TLI = 0.97, CFI = 0.97 and RMSEA = 0.067. The CFA results for a one-factor solution indicated that the hypothesized four-factor model had significantly better fit over a more parsimonious model where leader directiveness, leader supportiveness, follower self-reported competence and commitment items were set to load on a single factor ($\Delta \chi^2 = 3,754.83, p < 0.001$).

Descriptive statistics and intercorrelations among the independent and the dependent variables are displayed in Table I. The $\alpha$ coefficients which are listed on the primary diagonal of the intercorrelation matrix were in an acceptable range for all the variables of interest. Performance was significantly and positively correlated with leader consideration and direction, supervisor rating of follower competence and commitment, and follower self-rating of competence and commitment. Follower self-ratings of competence and commitment were moderately correlated with ratings of performance. Furthermore, average self-ratings of competence and commitment were higher than leader’s ratings.

**Leader-follower congruence**

$H1$ posits that SLT’s predictions are more likely to hold when leader rating and follower self-rating of competence and commitment are congruent. To test $H1$, it was necessary to make a comparison between follower self-ratings and their leader’s rating to identify whether leader rating and follower self-reports were in some agreement. Difference scores between follower self-rating and leader ratings were calculated, and subsequently individuals were classified into groups based on the magnitude of their self/leader difference. Individuals with difference scores within one standard deviation comprised in-agreement raters (Atwater and Yammarino 1997). The next step was to identify cases representing the four development levels in accordance with the terms used by Blanchard (2010). Developmental level 1 is the combination of low competence and high commitment, developmental level 2 reflects low to some competence in combination with low commitment, developmental level 3 is defined by moderate to high competence with variable commitment, and developmental level 4 is the combination of high competence and high commitment. This system of classification was applied for the present data. Based on Thompson and Vecchio (2009), development level 1 was defined as cases that were in the bottom third on competence (low) and in the top third on commitment (high). Developmental level 2 was defined as cases that were in the bottom half on competence (low to some degree
of competence) and in the bottom third on commitment (low). Developmental level 3 was defined as cases that were in the top half on competence (moderate to high) and below the top third on commitment (variable). Finally, developmental level 4 was defined as cases that were in the top third on both competence (high) and commitment (high). A case was labeled a “match” if it had the specified combination of leader supportiveness and leader directiveness within its developmental level. Other cases within the same developmental level were labeled “mismatches.” This system of classification makes a number of cases not potentially classifiable when they do not fall into the specified subsets of the two-dimensional space mapped by the detailed combinations of competence and commitment. In short, the theory only makes predictions for specific combinations of key variables. For this reason, cases were identifiable as falling outside SLT’s specified space and no prediction could be offered for such cases. In our data set, 258 cases, or 31.1 percent, were not classifiable.

Due to the lack of linear conversion of competence and commitment into development level, a polynomial regression approach was less attractive as an analytic technique. We, therefore, used Vecchio’s approach for analysis of the data (Vecchio, 1987; Norris and Vecchio, 1992; Fernandez and Vecchio, 1997; Vecchio et al., 2006; Thompson and Vecchio, 2009). Although categorization of variables as in SLT may be criticized, the theory has existed more than four decades in this format and should be tested on these premises. Analysis of the data was, therefore, done by applying omnibus testing (comparing results across conditions) and subgrouping analysis, labeled partitioned testing, for comparisons of matched and mismatched cases within each development level (i.e., cases where leader and follower attributes were in alignment with the theory’s framework were contrasted with those cases where these attributes were not in alignment). The result of the omnibus testing is presented in Table II and provides support for SLT principles. The test statistic values show that the level of mean match cases significantly exceeded the mean of mismatched cases for the output variable “performance.”

Omnibus testing showed no significant difference between male and female leaders on the output variable performance. In addition, follower rating of leader consideration and leader structuring showed that female leaders’ leadership style was more supportive (Mean = 3.97, SD = 0.74) than task orientated (Mean = 3.41, SD = 0.84). The same pattern was found for male leaders (consideration: Mean = 3.89, SD = 0.75 and task orientation: Mean = 3.35, SD = 0.80).

Furthermore, partitioned tests were conducted within each of the four developmental levels. Matches and mismatches within each developmental level were contrasted on the dependent variable “performance.” Results of the partitioned test were supportive of the theory’s predictions for all development levels, where the average value for performance was higher for matched cases than mismatched cases, and the difference was statistically significant (see Table III).

**Leader-follower incongruence**

The second hypothesis states that if incongruent, SLT’s predictions are more likely to hold when follower self-rating of competence and commitment are applied. To identify whether leader rating and follower self-reports were incongruent regarding competence and commitment, a comparison was conducted between follower self-ratings and their leader’s rating.

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
<th>n</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>6.20</td>
<td>0.46</td>
<td>83</td>
<td>11.59**</td>
</tr>
<tr>
<td>Mismatch</td>
<td>5.39</td>
<td>1.08</td>
<td>489</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *p < 0.05; **p < 0.01
The “incongruent” group was comprised of individuals with difference scores higher than one standard deviation. Analysis of the data was done by applying omnibus testing as described above. The result of the omnibus test is presented in Table IV and provides no support for Blanchard’s suggestion of applying follower self-ratings of competence and commitment as the basis for selecting leadership style, as no significant result was obtained for this notion.

Follower over-estimator
An extension of Blanchard’s suggestion of applying follower self-rating of competence and commitment was done by examining whether SLT’s predictions are likely to hold when follower self-rating of competence and commitment are higher than leader rating. The “over-estimator” group was identified and comprised of followers with difference scores higher than one standard deviation than their leader. Analysis of the data was done by applying omnibus testing as described above. The result of the omnibus test is presented in Table V.

---

### Table III.
Results of partitioned tests: test of SLT three-way interaction

<table>
<thead>
<tr>
<th>Group</th>
<th>$M$</th>
<th>SD</th>
<th>$n$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Development level 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent variable = performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>6.16</td>
<td>0.26</td>
<td>9</td>
<td>7.74**</td>
</tr>
<tr>
<td>Mismatch</td>
<td>5.39</td>
<td>1.08</td>
<td>489</td>
<td></td>
</tr>
<tr>
<td><strong>Development level 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent variable = performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>5.71</td>
<td>0.48</td>
<td>14</td>
<td>2.39*</td>
</tr>
<tr>
<td>Mismatch</td>
<td>5.39</td>
<td>1.08</td>
<td>489</td>
<td></td>
</tr>
<tr>
<td><strong>Development level 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent variable = performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>6.14</td>
<td>0.40</td>
<td>24</td>
<td>7.93**</td>
</tr>
<tr>
<td>Mismatch</td>
<td>5.39</td>
<td>1.08</td>
<td>489</td>
<td></td>
</tr>
<tr>
<td><strong>Development level 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent variable = performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>6.43</td>
<td>0.36</td>
<td>36</td>
<td>13.46**</td>
</tr>
<tr>
<td>Mismatch</td>
<td>5.39</td>
<td>1.08</td>
<td>489</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** *p < 0.05; **p < 0.01

---

### Table IV.
Results of omnibus test: test of SLT three-way interaction

<table>
<thead>
<tr>
<th>Group</th>
<th>$M$</th>
<th>SD</th>
<th>$n$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dependent variable = performance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>4.94</td>
<td>1.22</td>
<td>42</td>
<td>0.66</td>
</tr>
<tr>
<td>Mismatch</td>
<td>4.79</td>
<td>1.19</td>
<td>86</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** *p < 0.05; **p < 0.01

---

### Table V.
Results of omnibus test: test of SLT three-way interaction

<table>
<thead>
<tr>
<th>Group</th>
<th>$M$</th>
<th>SD</th>
<th>$n$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dependent variable = performance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>4.56</td>
<td>1.03</td>
<td>34</td>
<td>1.00</td>
</tr>
<tr>
<td>Mismatch</td>
<td>4.35</td>
<td>0.98</td>
<td>65</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** *p < 0.05; **p < 0.01
and, as expected, yields no significant support for SLT’s predictions being likely to hold if leaders apply ratings of competence and commitment by “over-estimators” as a basis for selecting leadership style.

**Follower under-estimator**

We examined whether Blanchard’s suggestion was likely to hold when follower self-rating of competence and commitment was lower than leader rating. The “under-estimator” group was identified and comprised of followers with lower difference scores than their leader (more than one standard deviation). Analysis of the data was done by applying omnibus testing. The result of the omnibus test is presented in Table VI and provides support for our hypothesis in that level of mean match cases significantly exceeded the mean of mismatched cases for the output variable “performance.”

Finally, **H5** states that SLT’s predictions are more likely to hold when leader rating of follower competence and commitment is applied when some agreement is unattainable. Again, the “incongruence group” was used to test the hypothesis, but now by applying leader rating of follower competence and commitment. The result of the omnibus test is presented in Table VII and provides support for the notion of using leader ratings of competence and commitment as a basis for selecting leadership style, as match significantly exceeded mismatch.

**Discussion**

The present study found support for **H1**, which posits that SLT’s predictions are more likely to hold when leader rating and follower self-rating of competence and commitment are congruent. Omnibus testing and partitioned testing showed that the level of mean matched cases significantly exceeded the mean of mismatched cases for the output variable “performance.” This result was consistent with results from Thompson and Glaso (2015), where omnibus testing yielded a significant result for the output variable “performance” in favor of SLT principles. Furthermore, the results contradict previous studies (Vecchio, 1987; Norris and Vecchio, 1992; Fernandez and Vecchio, 1997; Vecchio, Bullis, and Brazil, 2006; Thompson and Vecchio, 2009) that applied only supervisor rating of follower development level. It seems that congruence between both leader and follower ratings is central to determining follower development level, and enables leaders to provide followers with an appropriate amount of direction and support.

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
<th>n</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>6.55</td>
<td>0.21</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Mismatch</td>
<td>6.16</td>
<td>0.56</td>
<td>21</td>
<td>2.71*</td>
</tr>
<tr>
<td><strong>Notes:</strong> *p &lt; 0.05; **p &lt; 0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table VI. Results of omnibus test: test of SLT three-way interaction

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
<th>n</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>5.34</td>
<td>1.08</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>Mismatch</td>
<td>4.85</td>
<td>1.22</td>
<td>184</td>
<td>2.94**</td>
</tr>
<tr>
<td><strong>Notes:</strong> *p &lt; 0.05; **p &lt; 0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table VII. Results of omnibus test: test of SLT three-way interaction
Previous examinations of SLT have found delegating leadership to lack strong empirical grounding (Fernandez and Vecchio, 1997; Norris and Vecchio, 1992; Vecchio, 1987; Vecchio et al., 2006). It was assumed that the theory prescribed inappropriate behavior, as low supportiveness and directiveness imply lack of attention, approval and involvement by the leader. However, in light of the present findings, SLT’s suggestion that low supportiveness and low directiveness (delegating leadership style) would be favorable for skillful and motivated followers has been demonstrated. Such a finding suggests that highly motivated followers may place less value on expressions of supportiveness from their leaders. In addition, followers who desire greater independence may rely less on supervisors and express less need for supervisory intervention. This interpretation is consistent with SLT’s core principle that highly mature employees need less supervisory involvement. This finding is consistent with Yukl and Fu’s (1999) study, which found delegation is more likely to be successful when subordinates already know how to handle new job responsibilities and will take the initiative to deal with problems without waiting for direction.

Another interesting pattern is the relationship between development level and performance. Yammarino and Atwater (1997) differentiated between in-agreement good (follower self-rating and leader rating are both favorable and in agreement) and in-agreement poor (follower self-rating is low and similar to leader rating). They suggested that in-agreement good individuals would perform much better than individuals in the in-agreement poor category. Evidence for this suggestion has been found in the present study. In the in-agreement good category, (which is equivalent with development level 4), performance was 6.43 (Table III). In the in-agreement poor category (which is equivalent with development level 2), performance was 5.71 (Table III). This finding has relevance for SLT and illustrates an important point: that it is not only agreement between leader and follower that is positively related to performance, but this congruence is also affected by whether leader and follower agree that follower competence and commitment are good or poor.

H2 suggests that if there is incongruence between leader assessment and follower self-assessment of competence and commitment, then SLT’s predictions are more likely to hold when follower self-rating is applied. However, our result provides no support for Blanchard’s suggestion of applying follower’s self-ratings when some agreement is unattainable. These results seem to reflect the problematic nature of self-ratings (Atwater et al., 2009), and followers’ tendency to see themselves in an unrealistically positive light. As such, inaccurate and unrealistic self-perception form an inadequate basis for providing followers with the proper amount of direction and support. This finding is probably closely connected to H3, which posits that SLT’s predictions are not likely to hold when follower self-rating of competence and commitment are higher than leader rating. Over-estimators see no need to make improvements, and may react quite negatively and even aggressively to leader assessment of their competence and commitment, which can block the opportunity to come to some agreement. In such a situation, Blanchard’s (2010) suggestion of using follower self-assessment as a basis for selecting leadership style would prove unfavorable, as over-estimators perform far worse than expected. In other words, if leaders utilize a leadership style that an over-estimator perceives as matching their development level, leaders may not obtain the results they desire. SLT and other theories within the contingency approach are based on the principle that leader effectiveness depends on how well leader style fits the situation. Based on our findings, we suggest that leader effectiveness depends on how well leader style fits a congruent understanding of the situation. However, if some agreement is unattainable, then a favorable strategy for managing discrepant assessment of follower development level would be leader intervention, as leaders may have a perspective for assessing follower development level that outweighs follower self-rating. Support was obtained for this notion, as SLT’s predictions are more likely to hold when leader rating of follower competence and commitment is applied. Leaders should be more aware of and sensitive to under-estimators’
needs, listen to their problems and see things from their perspective. This group of followers thinks they are poor or average performers, but others see them as performing more favorably or even effectively. The result of omnibus testing provides support for this suggestion.

Finally, previous studies indicate the natural occurrence of “matches” is modest in organizations (Fernandez and Vecchio, 1997). It seems “mismatch” is widely distributed, which makes the possibility for testing SLT in field settings problematic. Hence, a statistically powerful test of the theory would require a large sample in order to capture the hypothesized range of situations. Despite the large sample in the present study, few followers were present at level 1, low on competence but high on commitment. We found 9 to represent a match with telling leadership style as shown in Table III. Furthermore, as portrayed in Table I, there was a strong positive association between competence and commitment as measured by both leaders (0.75**) or followers (0.65**). It seems that enthusiasm or high commitment occurs with increased competence, and is rare in combination with low competence. The rare occurrence of match is also consistent with findings by Zigarmi and Roberts (2017), who reported that the use of telling leadership style was very infrequent. Also at level 2 (low on competence to having some competence in combination with low commitment), few followers were represented. We found 11 followers to represent a match, as shown in Table III. The rare occurrence of this category of followers could reflect organizational recruitment requirements. As stated by many leaders at leadership training programs, they avoid hiring demotivated followers with modest competency skills. Taken together, it seems that SLT addresses situations that are rare in organizations. Due to the increasing span of control in most organizations today, we believe leaders can no longer tell their followers what, how and when to do everything. SLT could probably benefit from a revision which more accurately reflects the current leadership practices of organizations. Specifically, we propose a taxonomy consisting of three leadership styles, ranging from coaching leadership to achievement-orientated leadership, and a framework for matching each style to specific situations. Like Blanchard, we suggest using follower development level as the basis for selecting the proper leadership style, but framed differently. Accordingly, developmental level 1 should be the combination of some competence and some commitment. A coaching leadership style, characterized by some leader consideration combined with some leader structuring would presumably be an adequate leadership style to match this development level, where the purpose is to guide the follower in mastering the task and provide social-emotional support to elevate follower confidence. Over time, followers would gain competence and confidence in mastering the task (development level 2) and delegating leadership style would be an appropriate response to this situation, characterized by low leader structure and low leader consideration. Finally, a reasonable number of skilled followers would treasure a job situation where they continually experienced personal growth. Therefore, very competent and motivated followers (development level 3) should benefit from an achievement-orientated leadership style (House, 1996) where leaders set moderately difficult and challenging goals, continuously emphasize work performance improvements, and expect followers to achieve high levels of performance.

Practical implications
What do these results mean for SLT? The present study provides empirical grounds for advocating close adherence to some of SLT’s prescriptive guidelines. Blanchard (2010) has introduced “partnering for performance” as an integral component of SLT, where both leader and follower come to some agreement on follower development level. It seems that in-agreement assessments create a mutual understanding between leader and follower and elevate follower self-awareness of both strengths and weaknesses, making them more amenable to leader directiveness and supportiveness. Evidence was obtained that
in-agreement assessments are associated with performance. This “agreement approach” should guide and encourage the leader to tailor his or her leadership style to follower development level, instead of applying an intermediate leadership style with all followers.

However, no support was obtained for Blanchard’s (2010) suggestion to use follower self-ratings of competence and commitment in the case of discrepant ratings of follower development level. In contrast, data support the contention that leader assessment would be a better basis for providing follower direction and support. This finding should substantiate a revised form of “partnering for performance,” which currently does not suggest how to reach agreement. However, research has identified factors that could increase self-other rating agreement. Results indicate that individuals who reported higher levels of feedback were more likely to have self-assessments congruent with their supervisors’ ratings (Williams and Johnson, 2000). Accordingly, providing followers with progress-checks throughout the performance period may increase leader-follower agreement over time. Finally, for the under-rater group of followers, “partnering for performance” seems to be favorable, as followers need feedback and training in order to appreciate and realize their full potential. “Partnering for performance” may help leaders and followers to identify follower competence and commitment level, and subsequently enable leaders to provide followers with proper direction and support.

Strengths and limitations
A number of study strengths and limitations should be noted. First, past studies suffer from several shortcomings, for example, a major concern is the very high correlation between leader ratings of follower performance and leader rating of follower development level. Common method variance could also be substantial in previous studies. The present study tries to overcome some of these methodological issues by using two sources to assess follower development level.

Second, the present study, which has the largest number of leaders and followers to be examined in an empirical test of SLT, has applied data from business settings in Norway in order to examine the validity of SLT in a for-profit setting. This raises the question of possible limitations of the generalizability of the research from this sample of leaders and followers to other settings and other nations. Accordingly, the validity of our findings is yet to be explored in different cultural contexts like collectivistic cultures such as China and South Korea, or in a more individualistic cultural context such as USA. This would be in line with research showing that aspects of the situation impact what leader behaviors are ultimately effective (Lord et al., 2001; Shamir and Howell, 1999) such that situational leadership behaviors that are effective in one context may not necessarily translate to effective performance in a different context. However, using data from a business context could be regarded as a strength because it offers the possibility to control alternative sources of error variance, whereas different types of organizations could create problems when combining results across firms. As stated by Hair et al. (1998), results that look significant can be an artificial creation of the unique combination of a cross-firm data.

A third limitation is the cross-sectional nature of the data. The present study has relied on a cross-sectional static measurement design, where the validity of SLT has been examined by conducting a survey at a single point in time across a large number of leaders and followers. Consequently, the development aspect of leader-follower relations remains untested. An assessment of these developmental notions would require a longitudinal research design. The longitudinal approach appears essential, as leader behavior is often understood as a response to the performance of a follower over time.

Finally, congruence of follower self-rating and leader-rating of competence and commitment has been operationalized to be within one standard deviation. This interval was chosen in order to have an adequate sample size for conducting the necessary statistical
tests across the cells. It may be argued that this categorization of “some agreement” may shift the pattern of main differences in favor of SLT. A re-examination was therefore conducted, applying a half standard deviation (Atwater and Yammarino, 1997). Omnibus testing supported our analyses, as level of mean match cases significantly exceeded the mean of mismatched cases for the output variable “performance.” However, the number of match cases decreased and made partitioned test unachievable.

**Future research**

The present study has obtained evidence in support of the view that congruence in follower self-rating and leader rating is key to effective functioning of SLT. However, much needs to be learned about how the process of rating works, particularly understanding what determines follower self-rating and leader rating of development level. Self-other rating agreement research supports the contention that a number of variables influence self-other agreement, like individual characteristics such as rater’s gender, age, education level and job experience. Furthermore, contextual or situational factors may also influence follower self-rating and rating by the superior. Context factors comprise organizational situation, job pressure, political processes, organizational position, prior rating experiences, etc. (Yammarino and Atwater, 1997; Brutus et al., 1999; Ostroff et al., 2004; Vecchio and Anderson, 2009; Fleenor et al., 2010). Identifying determinants of self-other agreement in connection to SLT remains an important avenue for future research.

As mentioned above, it seems that “matching” in accordance with the precepts of SLT is rare in organizations. To our knowledge, this seems to be the case for different types of organizations and different levels within organizations. This may reflect substantial untapped potential for improvement of follower performance by deliberately altering supervisory style. Blanchard (2010) has suggested leadership training to enhance leaders’ ability to tailor their leadership style to follower competence and commitment. It is assumed that follower performance will improve via sensitization of supervisors to the theory’s principles. This suggestion for leadership training has not been tested empirically and remains among the major open questions with respect to SLT’s validity. To test the effects of a training program requires a longitudinal measurement design, where every subject in the sample must be observed on chronologically successive occasions. In the pursuit of such longitudinal data, researchers should administer a pre-test and a post-test on each subject for the effective measurement of change. Furthermore, different data collection methods could be applied, such as observation, interviews, questionnaires and ability tests.

**References**


Stogdill, R.M. and Coons, E.E. (Eds) (1957), "Leader behavior: its description and measurement", Bureau of Business Research Monograph No. 88, Ohio State University, Columbus, OH.


Yammarino, F.J. and Atwater, L.E. (1997), "Do managers see themselves as others see them? Implications of self-other rating agreement for human resources management", *Organizational Dynamics*, Vol. 25 No. 4, pp. 35-44.


Further reading


About the authors
Geir Thompson is Associate Professor of Organizational Psychology at BI Norwegian Business School. His research interests are leadership, emotions at work, interpersonal mistreatment and personality. Geir Thompson is the corresponding author and can be contacted at: geir.thompson@bi.no

Lars Glasø (PhD) is Professor of Organizational Psychology at BI Norwegian Business School. Glasø is certified as Clinical Psychologist and as Specialist in Applied Organizational Psychology (NPA). He has published extensively on issues such as leadership, emotions at work and workplace bullying.

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com

Situation leadership theory

591