Golestan University of Medical Sciences

Journal of Research Development in Nursing and Midwifery
(J Res Dev Nurs Midw)
Online ISSN: 2588-3038

Self-observer agreement in leadership practices of nursing supervisors: A comparison of supervisor and team member ratings

Najmeh Mohammadi [1], Mohammad Amin Karimi [1], Akram Rabbani Javadi [2], Fatemeh Heshmati Nabavi [3]*

1. Student Research Committee, Faculty of Nursing and Midwifery, Mashhad University of Medical Sciences, Mashhad, Iran
2. Staff Education Supervisor, Dr. Sheikh Hospital, Mashhad University of Medical Sciences, Mashhad, Iran
3. Nursing and Midwifery Care Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

* Correspondence: Fatemeh Heshmati Nabavi. Nursing and Midwifery Care Research Center, Mashhad University of Medical Sciences, Mashhad, Iran. Tel: +989155002213; Email: heshmatinf@mums.ac.ir

Abstract

Background: Leadership practices among nursing supervisors are essential to effective teamwork in hospital settings. However, supervisors' self-ratings may differ from team members' perceptions. Evidence of such discrepancies in nursing contexts remains limited, and understanding them may help improve leadership evaluation and guide targeted professional development. Therefore, this study aimed to compare nursing supervisors' self-rated leadership practices with team members' observer ratings and to assess the agreement between these two assessments.

Methods: This cross-sectional analytical study was conducted from February to March 2025 among nursing supervisors in 43 public and private hospitals of Khorasan Razavi Province, Iran. All eligible supervisors (n = 60) were included through census sampling, and 3-4 team members per supervisor (n = 201) were selected by simple random sampling. Data were collected using the Leadership Practices Inventory (LPI) self and observer versions. Mean team-member ratings were calculated for each supervisor. Paired-samples t-tests and Bland–Altman analysis were used to compare ratings and assess agreement. Data were analyzed using STATA version 19.5. The significance level was set at $p < 0.05$.

Results: Most participants in both groups were female. Nursing supervisors assigned higher scores to their leadership practices than did their team members across all leadership dimensions and the total LPI score (All $p < 0.001$). The mean (SD) total score was 82.22 (12.25) for self-ratings versus 68.66 (14.57) for observer ratings. Bland–Altman analysis indicated a systematic positive bias (Mean difference = 13.56, CI 95%: 8.56 to 18.56; LoA 95%: -24.39 to 51.51). Repeated-measures ANOVA revealed significant within-group differences across dimensions for both self-ratings ($p < 0.001$) and observer ratings ($p < 0.001$), with "Enabling Others to Act" scoring highest and "Inspiring a Shared Vision" scoring lowest in both groups.

Conclusion: Nursing supervisors consistently rated their leadership practices higher than team members. This suggests that self-ratings alone may not adequately characterize leadership practices. Therefore, these findings suggest that healthcare organizations implement self-observer assessments to bridge the perception gap between supervisors and team members, thereby facilitating more targeted and effective leadership development interventions.

Article Type: Research Article

Article History

Received: 11 July 2026
Received in revised form: 26 August 2026
Accepted: 10 September 2026
Available online:
DOI: 10.29252/jgbfnm.X.X.X

Keywords

Leadership
Nurse Administrators
Professional Competence
Interprofessional Relations
Nursing

OPEN ACCESS
© The author(s)

Highlights

What is current knowledge?

- Nursing leadership practices are predominantly evaluated through single-source self-assessment tools, such as self-version of the Leadership Practices Inventory (LPI).
- Leadership literature indicates that self-evaluations frequently diverge from how leadership behaviors are experienced and perceived by team members.

What is new here?

- Supervisors rated their leadership practices significantly higher than team members across all dimensions, with the largest divergence in Challenging the Process.
- Supervisors and their team members exhibited a concordant behavioral hierarchy, rating Enabling Others to Act highest and Inspiring a Shared Vision lowest.
- Bland-Altman analysis revealed wide limits of agreement, indicating that self- and observer ratings are not interchangeable.
- Multisource assessment provides a complementary, calibrated framework for targeted nursing leadership development.

Introduction

In healthcare organizations, effective nursing leadership plays a pivotal role in cultivating a healthy practice environment and improving the overall quality of patient care (1,2). Evidence indicates that specific leadership behaviors and relational styles significantly enhance nurses' job satisfaction, foster organizational commitment, and promote staff retention (1-3). Recent implementation efforts in Iran further illustrate the organizational relevance of nursing leadership. In a pilot action-research study conducted in six Iranian university-affiliated hospitals, transformational leadership, structural empowerment, and systematic monitoring of nursing indicators were incorporated into broader efforts to improve the quality of nursing care (4). Furthermore, supportive leadership practices directly bolster nurse resilience and cultivate positive work behaviors across clinical teams (3,5). To support nurse managers in their leadership roles, healthcare institutions should provide structured leadership development programs that incorporate experiential learning and reflection to strengthen leadership competencies (6,7).

Nursing supervisors, as middle-level nursing managers, are responsible for coordinating professional activities, supervising clinical practice, facilitating communication between managerial and frontline levels, and supporting quality improvement initiatives (8,9). Effective fulfillment of these responsibilities requires the active demonstration of

core leadership practices. According to Kouzes and Posner, exemplary leadership comprises five observable practices: Modeling the Way, inspiring a Shared Vision, Challenging the Process, Enabling Others to Act, and Encouraging the Heart (10). Evaluating these five core domains provides a comprehensive diagnostic framework to capture leadership strengths and pinpoint precise areas requiring professional growth (11,12).

However, assessing and advancing leadership practices in clinical settings presents a critical operational challenge. Traditionally, leadership evaluations have relied predominantly on single-source self-assessments. Yet, evidence demonstrates that self-ratings alone are inherently vulnerable to leniency bias, frequently concealing critical leadership "blind spots" (13-15). While supervisors evaluate themselves through the lens of their internal intentions, subordinates assess leadership primarily on the basis of observed daily behaviors and interactions (14,15). Discrepancies between self-evaluations and observer perspectives may reveal important gaps in self-awareness, which can impair workplace communication, erode trust, and hinder organizational effectiveness (13,14). Therefore, cross-evaluating supervisors' practices alongside their subordinates' perspectives is important for identifying these blind spots and provide nuanced insights into which specific leadership practices demand targeted developmental interventions (13,16).

In the Iranian hospital system, nursing supervisors work closely with their designated team members (Locally referred to as "Rabet"). These team members are nurses appointed within clinical units to act as focal points between supervisors and ward staff in implementing educational activities and unit-based educational initiatives (17). Under the guidance of nursing supervisors, they facilitate communication, support the implementation and follow-up of educational programs, and contribute to monitoring educational processes at the ward level (17,18). Given their continuous interaction with supervisors in carrying out these responsibilities, team members are well-positioned to observe supervisors' routine leadership practices and provide informed assessments of their leadership practices.

Despite the recognized necessity of self-observer assessments, there is a notable scarcity of research evaluating the level of agreement and discrepancies between nursing supervisors and their direct team members across the five LPI dimensions. Evaluating the agreement between these two perspectives is not merely comparative; it is essential to determine the concordance of perceived practices, illuminate hidden behavioral blind spots, and guide evidence-based, tailored leadership development programs in nursing administration. Therefore, this study aimed to compare nursing supervisors' self-rated leadership practices with team members' observer ratings and to assess the agreement between these two assessments.

Methods

Study design and setting

This cross-sectional analytical study was conducted from February to March 2025 across 43 public and private hospitals of Khorasan Razavi (Including the metropolitan area of Mashhad and its affiliated counties), a large province in northeastern Iran. All participating institutions operate under the regulatory and administrative supervision of Mashhad University of Medical Sciences (MUMS).

Participants and sampling

This study specifically focused on Staff Development and Health Education Supervisors whose primary organizational responsibilities include conducting training needs assessments, planning and executing continuous professional development for nursing personnel, monitoring patient health education, and providing clinical mentorship. In addition, these supervisors serve as active members of the central Staff Development and Health Education committees at MUMS, participating in periodic collaborative review meetings and peer-review audits across affiliated hospitals. A census sampling approach was used for the supervisors. All 60 eligible supervisors across the 43 participating hospitals were invited and enrolled (Ranging from 1 to 2 supervisors per hospital). Depending on institutional bed capacity, hospital size, and the total number of staff, some supervisors held a combined dual role covering both Staff Development and Health Education, whereas in larger tertiary hospitals, these responsibilities were assigned to distinct individual supervisors. To evaluate each supervisor's leadership practices through a multisource (Self-observer) assessment framework, team members directly collaborating with these supervisors at the ward level were recruited. It is important to note that they were exclusively full-time clinical bedside staff nurses without any formal or informal administrative, managerial, or supervisory authority within their respective clinical units (i.e., they were neither head nurses nor charge nurses/ward supervisors). In the organizational hierarchy of Iranian hospitals, supervisors do not primarily manage direct, routine bedside clinical tasks with all staff nurses; rather, their core leadership practices-such as training needs assessments, instructional planning, staff empowerment, competence evaluation, and clinical coaching-are operationalized through direct, continuous collaboration with these team members. Because team members actively participate in regular educational meetings, project implementations, and feedback sessions with the supervisors while maintaining their clinical bedside roles, they are well positioned to observe and assess leadership practices, potentially reducing observational and rater biases. Previous multisource feedback research has demonstrated that aggregating ratings from multiple raters improves the reliability of observer assessments by minimizing individual rater bias (19-21), with small rater groups being methodologically robust and widely utilized in healthcare leadership research (22). Initially, four team members were selected for each supervisor from their respective departmental staff pools using simple random sampling without replacement. For this purpose, eligible staff in each department were listed and assigned unique identification numbers, and four numbers per supervisor were drawn using a computer-generated random sequence (Generated via Microsoft Excel's random number generator). This yielded an initial enrollment of 60 supervisors and 240 team members. After excluding 39 team members due to incomplete questionnaires, the final statistical analysis included all 60 supervisors and 201 team members. In the final dataset, 39 supervisors had ratings from three observers and 21 supervisors had ratings from four observers. Accordingly, each supervisor was evaluated by three or four observers. The participant flowchart is summarized in Figure 1.

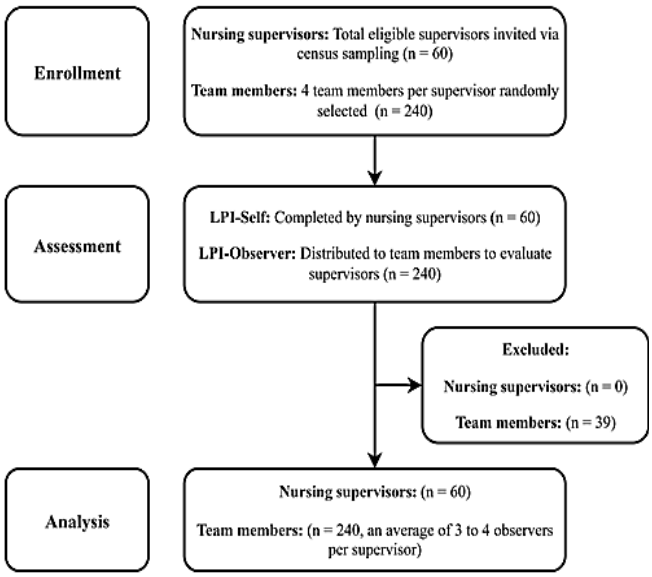

Figure 1. Participant flowchart

Inclusion and exclusion criteria

The inclusion criteria for nursing supervisors were having an official appointment as a Staff Development and Health Education supervisor, holding at least a bachelor's degree in nursing, and having a minimum of six months of active experience in the supervisory role. For team members, the inclusion criteria were being a clinical nurse actively collaborating with the designated educational supervisor, holding at least a bachelor's degree in nursing, and having at least six months of direct collaborative working experience with that supervisor.

Participants who withdrew prior to completion and questionnaires with missing responses exceeding 10% of items (Or missing critical demographic/Primary scale items) were excluded.

Measurements

The data collection instruments comprised a demographic questionnaire and two parallel versions of the Leadership Practices Inventory (LPI): the LPI-Self and the LPI-Observer.

Demographic characteristics

The demographic questionnaire collected background characteristics from both supervisors and team members, including age, gender, educational level, total years of clinical nursing experience, and years of experience in their current position. Additionally, interest in the nursing profession was assessed using a single-item question: "To what extent are you interested in the nursing profession?" rated on a 10-point numeric scale ranging from 1 ("Very low / Not interested at all") to 10 ("Very high / Extremely interested"), where higher scores indicated greater professional interest.

Leadership Practices Inventory (LPI)

Leadership practices were assessed using the 30-item Leadership Practices Inventory (LPI) developed by Kouzes and Posner (10), which evaluates five core exemplary leadership practices (6 items per dimension): Modeling the Way, Inspiring a Shared Vision, Challenging the Process, Enabling Others to Act, and Encouraging the Heart. Both versions evaluate identical behavioral dimensions and employ the same 5-point Likert scale ranging from 1 ("Never") to 5 ("Always"), but differ in their grammatical perspective and target evaluator:

LPI-Self: Completed by nursing supervisors for self-evaluation. All 30 items are phrased in the first-person singular (e.g., "I set a personal example of what I expect from others" or "I actively listen to diverse points of view"). Supervisors rate how frequently they perceive themselves engaging in each leadership behavior. Raw scores for each subscale range from 6 to 30, and the overall raw score ranges from 30 to 150. Higher scores reflect a higher self-perceived frequency of leadership practices.

LPI-Observer: Completed by designated team members to evaluate the observable leadership behaviors of their assigned supervisor. The 30 items directly correspond to the self-version but are formulated in the third-person singular (e.g., "He/She sets a personal example of what he/she expects from others" or "He/She actively listens to diverse points of view"). Observers used the same 5-point Likert scale (1 to 5) and scoring procedures. The possible raw score range is identical to the self-version (6 to 30 per subscale; 30 to 150 overall), with higher scores reflecting more frequently observed leadership behaviors. For each supervisor, the scores provided by the corresponding 3-4 team members were averaged to obtain a composite observer rating.

To facilitate direct comparison across dimensions and with the total score, raw scores for both versions were linearly transformed to a standardized 0-100 scale prior to statistical analysis.

Translation, validity, and reliability

The instruments were culturally adapted and translated into Persian using a standard forward-backward translation procedure. Two independent bilingual translators translated the original English versions into Persian. Following review and synthesis by the research team, an independent bilingual expert, blinded to the original text, back-translated the Persian version into English. Content validity was quantitatively evaluated by an expert panel of 10 faculty members specializing in nursing administration, psychometrics and instrument development; the mean item content validity ratio (CVR) was 0.867, and the Scale-level Content Validity Index using the averaging method (S-CVI/Ave) was 0.933, confirming excellent content validity. Internal consistency was evaluated independently for each version. In the present study, Cronbach's alpha for the LPI-Self was 0.87 for the total scale (Subscale range: 0.83 to 0.91). For the LPI-Observer, Cronbach's alpha was 0.83 for the overall scale (Subscale range: 0.75 to 0.91), demonstrating satisfactory reliability for both instruments (Supplementary Tables 1 and 2).

Data collection

Data were collected using a dual-mode approach to ensure confidentiality, response integrity, and accurate dyadic matching:

Supervisors (In person): Following an in-person briefing on the study objectives and ethical rights during a university-organized nursing managers' seminar, participating supervisors provided written informed consent and completed hard-copy, pre-coded questionnaires independently under researcher supervision.

Team members (Electronic survey): Observers completed the LPI-Observer via Porsline, an encrypted online survey platform. Each team member received an individualized, single-use web link sent via secure SMS to their personal phone. The survey introductory page provided detailed information about the study objectives and participants' ethical rights and assured respondents of anonymity prior to questionnaire access.

Dyadic matching was achieved through embedded alphanumeric codes within each link without recording identifiable personal data. Completed questionnaires were reviewed for completeness and subsequently entered into the statistical database for analysis. To eliminate evaluation apprehension and bias, supervisors were not given access to the online platform, individual responses, or participation status.

Data analysis

Data were analyzed using Stata version 19.5 (StataCorp LLC, College Station, TX 77845, USA). Descriptive statistics, including means, standard deviations, frequencies, and percentages, were used to summarize participant characteristics and study variables. Because the primary unit of analysis was the individual supervisor, ratings from the 3-4 team members per supervisor were averaged into a single composite observer score. This aggregation reduces individual rater variance, provides a consensus-based evaluation of each supervisor's leadership practices, and allows direct paired comparisons with the supervisor's self-rating.

The normality of the difference scores between self-assessments and observer-assessments was evaluated using the Shapiro-Wilk test. Because the differences followed an approximately normal distribution, paired-samples t-tests were used to determine whether statistically significant mean differences existed between self- and observer-ratings across the five dimensions and the total leadership score. Repeated-measures analysis of variance (ANOVA) was also used to examine differences across the five leadership practices dimensions, followed by Bonferroni-adjusted pairwise comparisons. The assumption of sphericity was assessed using Mauchly's test. Since the assumption of sphericity was met for both models ($p > 0.05$), sphericity-assumed results are reported. To evaluate the degree of agreement and identify potential systematic bias between supervisors' self-ratings and observer ratings, Bland-Altman analysis was conducted, reporting the mean difference (Bias) and 95% limits of agreement (LoA). Distinguishing between difference testing (Paired t-test) and agreement evaluation (Bland-Altman analysis) ensured a comprehensive assessment of rater divergence and concordance. A two-sided p-value < 0.05 was considered statistically significant.

Results

A total of 261 participants were included in the study, comprising 60 nursing supervisors and 201 team members. The nursing supervisors had a mean (SD) age of 40.03 (6.20) years and a mean (SD) total nursing work experience of 16.28 (6.11) years. Team members had a mean (SD) age of 34.34 (6.25) years and a mean (SD) work experience of 10.07 (5.87) years. Further demographic and professional characteristics of the participants are presented in Table 1.

Paired-samples t-tests were performed to compare supervisors' self-ratings and observer ratings on the LPI. Higher mean scores were observed for self-ratings than for observer ratings across all five leadership dimensions and for the total LPI score. Differences between the two types of ratings were statistically significant for all comparisons (All $p < .001$). For the total LPI score, the mean (SD) self-rating was 82.22 (12.25) compared with 68.66 (14.57) for observer ratings. Furthermore, repeated-measures ANOVA revealed significant within-group differences across leadership dimensions for both supervisors ($p < 0.001$) and team members ($p < 0.001$). Post-hoc pairwise comparisons were performed using Bonferroni adjustments. Detailed comparisons are presented in Table 2.

Bland-Altman analysis was performed to assess agreement between supervisors' self-ratings and the mean ratings provided by their team members on the LPI. For the total LPI score, the mean difference (Self-Observer) was 13.56 points (95% CI: 8.56-18.56), with 95% limits of agreement ranging from -24.39 to 51.51 points. Across the five LPI dimensions, mean differences ranged from 11.92 to 15.82 points. The

positive mean difference indicates that supervisors systematically rated their leadership practices higher than their team members. No substantial proportional bias was observed. Detailed estimates of the limits of agreement and their corresponding 95% confidence intervals are presented in Table 3. Additionally, the Bland-Altman plot for the total LPI score is presented in Figure 2.

Table 1. Demographic and professional characteristics of nursing supervisors and team members

Characteristic		Nursing Supervisors (n = 60)	Team Members (n = 201)
Age (Years)	Mean (SD)	40.03 (6.20)	34.34 (6.25)
Total Nursing Work Experience (Years)	Mean (SD)	16.28 (6.11)	10.07 (5.87)
Experience in Current Role (Years)	Mean (SD)	4.47 (3.83)	3.05 (1.61)
Interest in Nursing Profession*	Mean (SD)	8.55 (1.40)	7.04 (2.27)
Sex			
Female	n (%)	50 (83.3)	155 (77.1)
Male		10 (16.7)	46 (22.9)
Educational Level			
Bachelor's Degree	n (%)	33 (55.0)	183 (91.0)
Master's Degree		27 (45.0)	18 (9.0)

Note. Values are presented as mean (Standard deviation) or number (Percentage), as appropriate. *Interest in nursing profession was measured on a 10-point scale, with higher scores indicating greater professional interest

Table 2. Self- and observer-rated leadership practices across LPI dimensions: Inter-rater comparisons and within-group dimension differences

LPI Dimensions	Supervisors' Self-Ratings (n = 60)	Aggregated Team Members' Observer-Ratings (n = 60)	Paired-Samples t-Tests	
	Mean (SD)	Mean (SD)	t	P-value
D1	79.93 (15.43)	67.60 (13.10)	4.66	P < 0.001
D2	78.47 (16.26)	66.56 (14.96)	4.22	P < 0.001
D3	83.40 (13.19)	67.58 (14.42)	6.21	P < 0.001
D4	85.07 (10.94)	71.82 (15.34)	5.22	P < 0.001
D5	84.24 (13.36)	69.44 (16.71)	5.11	P < 0.001
Total score	82.22 (12.25)	68.66 (14.57)	5.43	P < 0.001
ANOVA (Repeated Measures)	-	-	-	-
F (df)	F (4,236) = 8.80	F (4,236) = 14.02	-	-
P-value	P < 0.001	P < 0.001	-	-
Pairwise Differences (Post-Hoc)	-	-	-	-
D1 vs D2	P = 1.000	P = 1.000	-	-
D1 vs D3	P = 0.054	P = 1.000	-	-
D1 vs D4	P = 0.003	P < 0.001	-	-
D1 vs D5	P = 0.012	P = 0.203	-	-
D2 vs D3	P = 0.005	P = 1.000	-	-
D2 vs D4	P < 0.001	P < 0.001	-	-
D2 vs D5	P < 0.001	P = 0.003	-	-
D3 vs D4	P = 1.000	P < 0.001	-	-
D3 vs D5	P = 1.000	P = 0.188	-	-
D4 vs D5	P = 1.000	P = 0.027	-	-

Note. The 60 aggregated observer-ratings were derived from ratings provided by 201 team members, with 3-4 team-member ratings averaged for each supervisor; therefore, all paired analyses were conducted on 60 supervisor-level pairs. Paired-samples t-tests compared supervisors' self-ratings with aggregated observer ratings. Within-group differences across the five LPI dimensions were evaluated using repeated-measures ANOVA separately for self-ratings and aggregated observer ratings, followed by Bonferroni-adjusted pairwise comparisons. Agreement between self-ratings and aggregated observer ratings was assessed using Bland-Altman analysis. LPI, Leadership Practices Inventory; D1 = Modeling the Way; D2 = Inspiring a Shared Vision; D3 = Challenging the Process; D4 = Enabling Others to Act; D5 = Encouraging the Heart.

Table 3. Bland–Altman analysis of agreement between supervisors' self-ratings and mean team members' ratings

LPI Scale	Mean Difference (Bias) (95% CI)	95% LoA	Lower LoA, 95% CI	Upper LoA, 95% CI
Modeling the Way	12.33 (7.04 to 17.61)	-27.80 to 52.45	-38.34 to -20.06	44.71 to 63.00
Inspiring a Shared Vision	11.92 (6.27 to 17.56)	-30.93 to 54.76	-42.19 to -22.67	46.50 to 66.02
Challenging the Process	15.82 (10.72 to 20.92)	-22.86 to 54.50	-33.03 to -15.40	47.05 to 64.67
Enabling Others to Act	13.25 (8.17 to 18.33)	-25.30 to 51.79	-35.43 to -17.87	44.36 to 61.93
Encouraging the Heart	14.80 (9.00 to 20.59)	-29.15 to 58.74	-40.70 to -20.68	50.27 to 70.29
Total Score	13.56 (8.56 to 18.56)	-24.39 to 51.51	-34.36 to -17.07	44.19 to 61.48

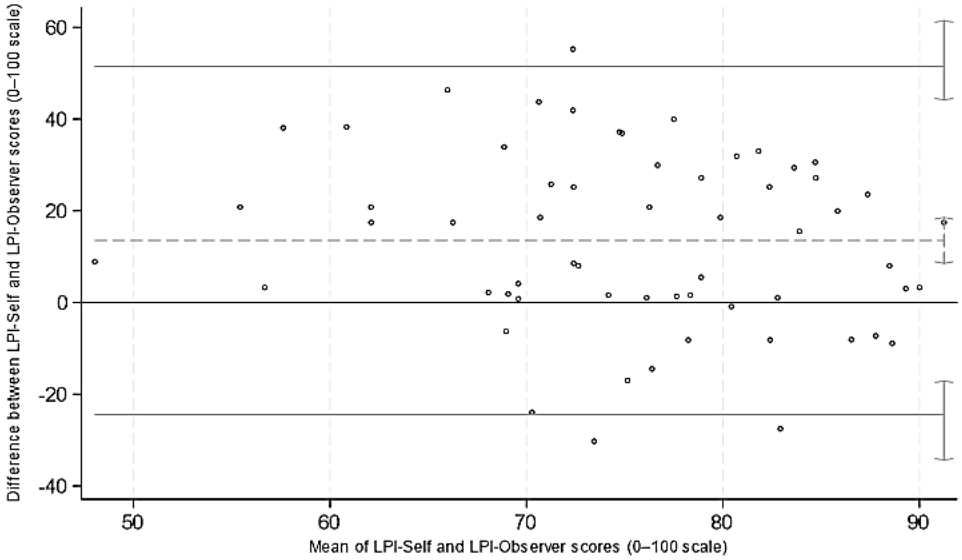

Figure 2. Bland-Altman plot of agreement between supervisors' self-ratings and mean team members' ratings for the total LPI score. The solid horizontal line represents the mean difference (Bias), and the upper and lower horizontal lines represent the 95% limits of agreement. Each point corresponds to one supervisor and represents the difference between self- and team-member ratings plotted against their mean score.

Discussion

The present study compared nursing supervisors' self-assessments of leadership practices with ratings provided by supervisors' team members using the LPI. The findings revealed a consistent pattern across all leadership practices dimensions and the overall score. On average, nursing supervisors rated their leadership practices significantly higher than did their team members, with mean differences ranging from approximately 12 to 16 points on the 0-100 scale. Bland-Altman analysis demonstrated a systematic average positive bias in self-ratings; however, the relatively wide limits of agreement indicated substantial individual-level variation in the magnitude of disagreement between supervisors and observers.

The overall leadership practice scores observed in this study were generally comparable to those reported in previous studies using the LPI among nursing leaders. Studies conducted among nurse managers have consistently reported relatively high self-rated leadership scores across the five leadership practices, suggesting that nursing leaders often perceive themselves as frequently demonstrating exemplary leadership behaviors (23,24). However, most previous nursing studies have relied primarily on self-assessment data and have therefore provided limited insight into how leadership behaviors are perceived by others working closely with nurse leaders. The present study contributes to this literature by incorporating observer ratings from supervisors' team members and demonstrating that leadership practices may be viewed less favorably by observers than by leaders themselves.

Our finding that supervisors systematically assigned higher self-ratings align closely with multisource nursing leadership literature that employs both self- and observer-based assessments. For example, a previous study examining transformational leadership practices among nurse managers found that managers evaluated their own leadership behaviors significantly more positively than did the staff nurses working with them (25). In a similar vein, a large-scale evaluation utilizing the LPI demonstrated that nurse leaders consistently scored their leadership practices higher across all five dimensions than did their followers (26). Although direct numerical comparisons across studies are limited by differences in metric scaling across the literature (Raw summative scores vs. normalized 0-100 scales), the direction and pattern of this perceptual discrepancy remain consistent across healthcare settings. Furthermore, broader research utilizing 360-degree leadership feedback approaches has shown that leadership ratings often vary systematically according to the source of evaluation, frequently identifying developmental needs that remain unrecognized through self-assessment alone (14,16). Together, these consistent findings underscore that observer perspectives provide indispensable, complementary information for evaluating leadership practices and addressing managerial blind spots in nursing.

Across both groups, a consistent ranking of leadership practices emerged, underscoring a shared recognition of strengths in leadership practices and operational challenges despite significant differences in absolute scores. Both supervisors and team members ranked 'Enabling Others to Act' and 'Encouraging the Heart' as the most prominent practices. Notably, post-hoc pairwise comparisons revealed that team members evaluated 'Enabling Others to Act' as significantly distinct from all other dimensions, highlighting that supportive, relational, and empowerment-focused behaviors are particularly visible in daily clinical practice. Conversely, 'Inspiring a Shared Vision' consistently received the lowest scores in both groups. This suggests mutual awareness of a persistent operational challenge: in fast-paced clinical environments driven by immediate administrative tasks and crisis management, supervisors have limited capacity and strategic bandwidth to articulate long-term visions. In contrast, the widest self-observer perceptual gap was observed in 'Challenging the Process'. While supervisors perceived themselves as proactive initiators of innovation and change, team members perceived substantially less improvement in bedside practice. This divergence likely reflects organizational bureaucracy and institutional barriers in healthcare systems, where supervisory problem-solving efforts often fail to translate into observable workflow modifications. These findings emphasize that leadership development should not only cultivate visionary communication but also establish actionable mechanisms to bridge the gap between supervisory change initiatives and tangible unit-level practice.

The observed discrepancies align with broader self-other agreement literature. While systematic mean differences demonstrate supervisors rating their practices more favorably than team members (13,14)-likely influenced by self-enhancement tendencies, leniency bias, or divergent access to performance information-these findings should not be construed as definitive evidence of "overestimation" or rater inaccuracy, as observer ratings do not represent an objective gold standard. From a practical and measurement perspective, the wide limits of agreement without substantial proportional bias indicate that self- and observer ratings are not interchangeable and that the positive mean bias is relatively uniform. Crucially, the Bland-Altman analysis highlights substantial heterogeneity across individual dyads, ranging from high concordance to pronounced perceptual gaps. Such divergence-spanning all five LPI dimensions-reflects distinct experiential standpoints, observational opportunities, and relational-contextual dynamics (e.g., interaction frequency and quality, shared tenure, open communication, and informal feedback channels) rather than a single, uniform deficiency. Consequently, institutionalizing multi-source feedback mechanisms can reconcile these dyadic perceptual differences, facilitate constructive dialogue, and support calibrated leadership development in nursing.

To our knowledge, this study is among the few studies in nursing management to directly compare self- and observer-rated leadership practices using both versions of the LPI. In addition, the inclusion of multiple observers for each supervisor was intended to provide a broader view of leadership practices across unit contexts. Nevertheless, several limitations should be acknowledged. The study was conducted in a single healthcare system, which may limit generalizability to other organizational contexts. Furthermore, although observer ratings were averaged across multiple team members, other sources of leadership feedback, such as staff nurses, peers, or senior administrators, were not included. Additionally, while aggregating ratings across multiple observers per supervisor was intended to capture a broader perspective of leadership practices, inter-rater reliability among observers (e.g., intraclass correlation coefficients) was not statistically evaluated, representing a methodological limitation.

Conclusion

This study demonstrates a consistent pattern of perceptual differences between supervisors and their team members across all five dimensions of the LPI. On average, supervisors rated their leadership practices higher than did team members, although substantial variability existed across individual dyads, as evidenced by the wide limits of agreement. These differences should be understood as reflecting distinct vantage points and observational opportunities rather than an objective overestimation by supervisors or underestimation by observers; neither rating represents an absolute measure of true leadership efficacy. These findings highlight the importance of incorporating observer perspectives alongside self-assessments when evaluating leadership practices in nursing management. Multisource assessment approaches may offer a more comprehensive understanding of leadership strengths and developmental needs, potentially foster constructive dialogue and prompt reflection among nursing leaders; however, their effects on self-awareness require further investigation and cannot be inferred causally from the present study. Future longitudinal and interventional designs should explore factors associated with self–observer agreement and examine whether feedback interventions might enhance leadership effectiveness and team collaboration among nursing leaders.

Acknowledgement

The authors would like to express their sincere gratitude to all the supervisors and their team members at the participating hospitals for their invaluable cooperation and support during the data collection process.

Funding Sources

This study was financially supported by Mashhad University of Medical Sciences and was based on the first author's Master's thesis.

Ethical Statement

The study was approved by the Ethics Committee of Mashhad University of Medical Sciences (IR.MUMS.NURSE.REC.1404.028). Written informed consent was obtained from all participants before data collection. Confidentiality and anonymity of participants were ensured, and participants were informed of their right to withdraw from the study at any time without consequences.

Conflicts of Interest

The authors declare no conflicts of interest.

Author Contributions

N.M.: Investigation, Data collection, and Data curation, Including the matching of self-reported and Observer data. M. A. K.: Methodology, Formal data analysis, and Revising the original draft of the manuscript. A. R. J.: Methodology and Data collection. F. H. N.: Conceptualization, Methodology design, and Overall study supervision. All authors contributed to the critical review and editing of the manuscript and approved the final submitted version.

Data Availability Statement

The datasets generated and analyzed during the current study are not publicly available due to participant privacy concerns but are available from the corresponding author upon reasonable request.

Use of Artificial Intelligence

During the preparation of this manuscript, the authors utilized ChatGPT (OpenAI) solely to assist with English language editing, grammatical correction, and improvements in readability. The authors carefully reviewed, verified, and revised all generated text, and they assume full responsibility for the integrity, scientific validity, and final content of this publication. Artificial intelligence was not employed for study design, data collection, statistical analysis, or drawing scientific conclusions.

References

1. Hult M, Terkamo-Moisio A, Kaakinen P, Karki S, Nurmeksela A, Palonen M, et al. Relationships between nursing leadership and organizational, staff and patient outcomes: A systematic review of reviews. Nurs Open. 2023;10(9):5920-36. [View at Publisher] [DOI] [PMID] [Google Scholar]
2. Cummings GG, Tate K, Lee S, Wong CA, Paananen T, Micaroni SP, et al. Leadership styles and outcome patterns for the nursing workforce and work environment: A systematic review. Int J Nurs Stud. 2018;85:19-60. [View at Publisher] [DOI] [PMID] [Google Scholar]
3. Alsadaan N, Salameh B, Reshia FAAE, Alruwaili RF, Alruwaili M, Awad Ali SA, et al. Impact of nurse leaders behaviors on nursing staff performance: a systematic review of literature. Inquiry. 2023:60:469580231178528. [View at Publisher] [DOI] [PMID] [Google Scholar]
4. Ebadi A, Ashrafizadeh H, Khanali Mojen L, Mirzaie S, Kazemzadeh Z, Ariamloo P, et al. Feasibility study of using the magnet hospital model in improving quality of nursing care in Iran: An action research pilot study. J Res Dev Nurs Midw. 2025;22(3):2-7. [View at Publisher] [DOI] [Google Scholar]
5. Wei H, Roberts P, Strickler J, Corbett RW. Nurse leaders' strategies to foster nurse resilience. J Nurs Manag. 2019;27(4):681-7. [View at Publisher] [DOI] [PMID] [Google Scholar]
6. Abou Hashish E, Farghaly S. Exploring how nurse managers' knowledge of succession planning affects their leadership and organisational resilience. Nurs Manag (Harrow). 2021;28(6):21-8. [View at Publisher] [DOI] [PMID] [Google Scholar]
7. Esquisábel-Soteras B, Olano-Lizarraga M, Pardavila-Belio MI, Vázquez-Calatayud M. Developing effective transition programmes for first-line nurse managers: A scoping review of evidence, barriers, and best practices. Nurse Educ Pract. 2025:88:104537. [View at Publisher] [DOI] [PMID] [Google Scholar]
8. Filomeno L, Forte D, Di Simone E, Di Muzio M, Tartaglini D, Lommi M, et al. Systematic review and psychometric properties analysis of first-, middle-, and top-level nurse manager's core competencies instruments. J Nurs Manag. 2024:2024:2655382. [View at Publisher] [DOI] [PMID] [Google Scholar]
9. Heidari A, Kazemi SB, Kabir MJ, Khatirnamani Z, Lotfi M, Rafiei N, et al. The relationship of communication skills with leadership style and conflict management strategies of head nurses: A cross-sectional study in Northern Iran. J Res Dev Nurs Midw. 2024;21(1):38-42. [View at Publisher] [DOI] [Google Scholar]
10. Posner BZ, Kouzes JM. Development and validation of the leadership practices inventory. Educational and psychological measurement. 1988;48(2):483-96. [View at Publisher] [DOI] [Google Scholar]
11. Posner BZ. Investigating the reliability and validity of the Leadership Practices Inventory®. Adm. Sci. 2016;6(4):17. [View at Publisher] [DOI] [Google Scholar]
12. Kouzes JM, Posner BZ. The Leadership Challenge: How to Make Extraordinary Things Happen in Organizations. 7th ed. Hoboken (NJ): Wiley; 2023. [View at Publisher]

13. Lee A, Carpenter NC. Seeing eye to eye: A meta-analysis of self-other agreement of leadership. The leadership quarterly. 2018;29(2):253-75. [View at Publisher] [DOI] [Google Scholar]
14. Nielsen K, Tafvelin S, von Thiele Schwarz U, Hasson H. In the eye of the beholder: How self-other agreements influence leadership training outcomes as perceived by leaders and their followers. J Bus Psychol. 2022;37(1):73-90. [View at Publisher] [DOI] [Google Scholar]
15. Pronin E, Hazel L. Humans' bias blind spot and its societal significance. Curr Dir Psychol Sci. 2023;32(5):402-9. [View at Publisher] [DOI] [Google Scholar]
16. Emam SM, Fakhry SF, Abdrabou HM. Leaders development program by 360 degree feedback: reflection on head nurses' leadership practices. BMC Nurs. 2024;23(1):772. [View at Publisher] [DOI] [PMID] [Google Scholar]
17. Shoghi M, Sajadi M, Oskuie F, Dehnad A, Borimnejad L. Strategies for bridging the theory-practice gap from the perspective of nursing experts. Heliyon. 2019;5(9):e02503. [View at Publisher] [DOI] [PMID] [Google Scholar]
18. Nabavi FH, Vanaki Z, Mohammadi E, Yazdani S. Balancing power: A grounded theory study on partnership of academic service institutes. J Adv Med Educ Prof. 2017;5(3):124-33. [View at Publisher] [PMID] [Google Scholar]
19. Fleenor JW. What Can We Learn from Research on Multisource Feedback in Organizations? In: Rollett W, Bijlsma H, Röhl S, editors. Student Feedback on Teaching in Schools: Using Student Perceptions for the Development of Teaching and Teachers. Cham: Springer International Publishing; 2021. p. 221-36. [View at Publisher] [DOI] [Google Scholar]
20. Markham SE, Smith JW, Markham IS, Braekkan KF. A new approach to analyzing the Achilles' heel of multisource feedback programs: Can we really trust ratings of leaders at the group level of analysis? Leadersh Q. 2014;25(6):1120-42. [View at Publisher] [DOI] [Google Scholar]
21. Lockyer J, Sargeant J. Multisource feedback: an overview of its use and application as a formative assessment. Can Med Educ J. 2022;13(4):30-5. [View at Publisher] [DOI] [PMID] [Google Scholar]
22. Jackson D, Michaelides G, Dewberry C, Jones A, Toms S, Schwencke B, et al. Uncertainty about Rater Variance and Small Dimension Effects Impact Reliability in Supervisor Ratings. Hum Perform. 2022;35(3-4):278-301. [View at Publisher] [DOI] [Google Scholar]
23. Alharbi SM, Almansour AM. Leadership behaviours and practices assessment among nurse managers in riyadh second health cluster. BMC Nurs. 2025;24(1):969. [View at Publisher] [DOI] [PMID] [Google Scholar]
24. Silva VLdS, Camelo SHH, Soares MI, Resck ZMR, Chaves LDP, Santos FCd, et al. Leadership practices in hospital nursing: a self of manager nurses. Rev Esc Enferm USP. 2017;51:e03206. [View at Publisher] [DOI] [PMID] [Google Scholar]
25. Uslu Sahan F, Terzioglu F. Transformational leadership practices of nurse managers: the effects on the organizational commitment and job satisfaction of staff nurses. Leadersh Health Serv (Bradf Engl). 2022;35(4):494-505. [View at Publisher] [DOI] [PMID] [Google Scholar]
26. Nanyonga RC, Bosire EN, Heller DJ, Bradley E, Reynolds NR. Predictors of nursing leadership in Uganda: a cross-sectional study. Health Policy Plan. 2020;35(Supplement_1):i51-i64. [View at Publisher] [DOI] [PMID] [Google Scholar]

Cite this article as:

Mohammadi N, Karimi MA, Rabbani Javadi A, Heshmati Nabavi F. Self-observer agreement in leadership practices of nursing supervisors: A comparison of supervisor and team member ratings. *J Res Dev Nurs Midw*. 2026;X(X):X. http://dx.doi.org/10.29252/jgbfnm.X.X.X

Supplementary Table 1. Leadership Practices Inventory Self Version (LPI-Self)

#	Dimension	Statement
1	Model the Way	I set a personal example of what I expect of others.
2	Inspire a Shared Vision	I talk about future trends that will influence how our work gets done.
3	Challenge the Process	I seek out challenging opportunities that test my own skills and abilities.
4	Enable Others to Act	I develop cooperative relationships among the people I work with.
5	Encourage the Heart	I praise people for a job well done.
6	Model the Way	I make certain that people adhere to the principles and standards that have been agreed upon.
7	Inspire a Shared Vision	I describe a compelling image of what our future could be like.
8	Challenge the Process	I challenge people to try out new and innovative ways to do their work.
9	Enable Others to Act	I actively listen to diverse points of view.
10	Encourage the Heart	I make it a point to let people know about my confidence in their abilities.
11	Model the Way	I follow through on the promises and commitments that I make.
12	Inspire a Shared Vision	I appeal to others to share an exciting dream of the future.
13	Challenge the Process	I actively search for innovative ways to improve what we do.
14	Enable Others to Act	I treat others with dignity and respect.
15	Encourage the Heart	I make sure that people are creatively recognized for their contributions to the success of our projects.
16	Model the Way	I ask for feedback on how my actions affect other people's performance.
17	Inspire a Shared Vision	I show others how their long-term interests can be realized by enlisting in a common vision.
18	Challenge the Process	I ask "What can we learn?" when things don't go as expected.
19	Enable Others to Act	I involve people in the decisions that directly impact their job performance.
20	Encourage the Heart	I publicly recognize people who exemplify commitment to shared values.
21	Model the Way	I build consensus around a common set of values for running our organization.
22	Inspire a Shared Vision	I paint the "big picture" of what we aspire to accomplish.
23	Challenge the Process	I identify measurable milestones that keep projects moving forward.
24	Enable Others to Act	I give people a great deal of freedom and choice in deciding how to do their work.
25	Encourage the Heart	I tell stories of encouragement about the good work of others.
26	Model the Way	I am clear about my philosophy of leadership.
27	Inspire a Shared Vision	I speak with genuine conviction about the higher meaning and purpose of our work.
28	Challenge the Process	I take initiative in anticipating and responding to change.
29	Enable Others to Act	I ensure that people grow in their jobs by learning new skills and developing themselves.
30	Encourage the Heart	I get personally involved in recognizing people and celebrating accomplishments.

Supplementary Table 2. Leadership Practices Inventory Observer Version (LPI-Observer)

#	Dimension	Statement
1	Model the Way	Sets a personal example of what he/she expects of others
2	Inspire a Shared Vision	Talks about future trends that will influence how our work gets done
3	Challenge the Process	Seeks out challenging opportunities that test his/her own skills and abilities
4	Enable Others to Act	Develops cooperative relationships among the people he/she works with
5	Encourage the Heart	Praises people for a job well done
6	Model the Way	Makes certain that people adhere to the principles and standards that have been agreed upon
7	Inspire a Shared Vision	Describes a compelling image of what our future could be like
8	Challenge the Process	Challenges people to try out new and innovative ways to do their work
9	Enable Others to Act	Actively listens to diverse points of view
10	Encourage the Heart	Makes it a point to let people know about his/her confidence in their abilities
11	Model the Way	Follows through on promises and commitments he/she makes
12	Inspire a Shared Vision	Appeals to others to share dream of the future
13	Challenge the Process	Actively searches for innovative ways to improve what we do
14	Enable Others to Act	Treats people with dignity and respect
15	Encourage the Heart	Makes sure that people are creatively recognized for their contributions to the success of our projects
16	Model the Way	Asks for feedback on how his/her actions affect other people's performance
17	Inspire a Shared Vision	Shows others how their long-term interests can be realized by enlisting in a common vision
18	Challenge the Process	Asks "What can we learn?" when things don't go as expected
19	Enable Others to Act	Involves people in the decisions that directly impact their job performance
20	Encourage the Heart	Publicly recognizes people who exemplify commitment to shared values
21	Model the Way	Builds consensus around a common set of values for running our organization
22	Inspire a Shared Vision	Paints the "big picture" of what we aspire to accomplish
23	Challenge the Process	Identifies measurable milestones that keep projects moving forward
24	Enable Others to Act	Gives people a great deal of freedom and choice in deciding how to do their work
25	Encourage the Heart	Tells stories of encouragement about the good work of others
26	Model the Way	Is clear about his/her philosophy of leadership
27	Inspire a Shared Vision	Speaks with genuine conviction about the higher meaning and purpose of our work
28	Challenge the Process	Takes initiative in anticipating and responding to change
29	Enable Others to Act	Ensures that people grow in their jobs by learning new skills and developing themselves
30	Encourage the Heart	Gets personally involved in recognizing people and celebrating accomplishments