INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research


**International Journal of Human Factors and Ergonomics**

ISSN online: 2045-7812 - ISSN print: 2045-7804
https://www.inderscience.com/ijhfe

# Assessing and improving supervisory behaviours at standardised meetings using self-recorded video and the operant supervisory taxonomy index

Simon Elvnäs, Ned Carter, Mikael Forsman

**Article History:**

| | |
|---|---|
| Received: | 12 July 2023 |
| Last revised: | 14 January 2024 |
| Accepted: | 17 January 2024 |
| Published online: | 24 June 2024 |

# Assessing and improving supervisory behaviours at standardised meetings using self-recorded video and the operant supervisory taxonomy index

Simon Elvnäs*

Division of Ergonomics,
Royal Institute of Technology,
Hälsovägen 11C, 141 57 Huddinge, Sweden
Email: elvnas@kth.se
*Corresponding author

Ned Carter

Swedish Association of Local Authorities and Regions,
118 82 Stockholm, Sweden
Email: nedcarter52@outlook.com

Mikael Forsman

Division of Ergonomics,
Royal Institute of Technology,
Hälsovägen 11C, 141 57 Huddinge, Sweden
Email: miforsm@kth.se

**Abstract:** The aims of this study were to use the operant supervisory taxonomy index (OSTI) with self-recorded videos: 1) to assess three crucial supervisory behaviours at recurrent workplace meetings; 2) to assess an intervention to increase supervisors' use of performance consequences; 3) to demonstrate how using OSTI with video for studying supervisor-employee interactions contributes to the field of human factors and ergonomics. Eight supervisors at a technical company video-recorded themselves at standardised daily staff meetings for ten weeks (360 videos, in total). All supervisors increased their delivery of performance consequences per meeting after participating in weekly sessions focusing on their use of this behaviour. The use of the crucial supervisory behaviours varied within and between supervisors. A combination of self-recorded videos over time and OSTI is a useful method for objectively analysing supervisor performance in detail and can contribute to improving supervisory behaviours and their impact on performance and health.

**Keywords:** OSTI; leadership; intervention; feedback; management; agile; transformational; first-line managers; organisational behaviour management; OBM.

**Reference** to this paper should be made as follows: Elvnäs, S., Carter, N. and Forsman, M. (2024) 'Assessing and improving supervisory behaviours at standardised meetings using self-recorded video and the operant supervisory taxonomy index', *Int. J. Human Factors and Ergonomics*, Vol. 11, No. 2, pp.188–207.

**Biographical notes:** Simon Elvnäs is a PhD candidate in Ergonomics at KTH – Royal Institute of Technology. In 1997, he obtained his Bachelor of Science in Occupational Therapy from Örebro University and his Master's degree in Industrial Ergonomics at Linköping University in 2009. His thesis focuses on supervisory behaviour and the gap between what leaders should do and what they actually do in their daily work.

Ned Carter is a retired psychologist, with a PhD from Uppsala University, who spent his final 15 years of employment as a work environment specialist at the Swedish Association of Local Authorities and Regions. His previous work experience also focused on work environment issues at AFA Insurance, Stockholm, Sweden, the Department of Occupational and Environmental Medicine, Uppsala University and the Swedish National Institute of Occupational Safety and Health, Solna, Sweden. He has co-authored publications on topics such as occupational injury prevention and sleep disordered breathing.

Mikael Forsman is a Professor of Ergonomics and Head of the Division of Ergonomics at CBH, KTH – Royal Institute of Technology. He works both in field and lab projects, often in collaboration with practitioners to develop smart methods to measure physical work exposure. His interests also include psychosocial exposures since there is an association between these and physical health. He is also contributing to the area of occupational medicine by teaching and supervising PhD students.

## 1 Introduction

Within the field of human factors and ergonomics (HFE), a plethora of measures and techniques for evaluating and analysing physical and cognitive activities have been developed. Measures and techniques for studying social interactions in occupational settings seem less common, especially with regards to interactions between management/supervisors and employees/subordinates. There are numerous definitions of leadership and taxonomies of leadership (Alvesson and Einola, 2019; Banks et al., 2018; Meuser et al., 2022; Bommer et al., 2018; van Knippenberg and Sitkin, 2013), and a variety of methodologies have been used to study leadership (Yukl, 1994). There is relatively little descriptive data within humans, technology and organisation (HTO)-research and literature about what leaders and supervisors do or should do to influence employees (Berglund et al., 2020; Eklund, 2000; Salvendy and Karwowski, 2021), although most definitions of leadership include some indication that the behaviour of others is to be influenced (Paul et al., 2001; Yukl, 1999; Alvesson and Spicer, 2012; Karltun et al., 2022). Given the importance attached to leadership for productivity, health and well-being, there is also a somewhat surprising lack of observational data on supervisory behaviour in real life settings (Barley and Kunda, 2001; Alvesson, 2019). Nonetheless, there seems to be innumerable books, training programs and other interventions claiming to improve leadership. Evaluations of interventions targeting leadership behaviours, especially in field settings and over time, are rare in the HFE research literature. Avolio and colleagues (Avolio and Gardner, 2005; Avolio et al., 2009) found in their meta-analysis of 100 years of leadership encompassing thousands of reports, that only 201 studies had evaluated an intervention. Of these, only 30% were

carried out in field settings and only 11 studies evaluated an intervention for more than one month. A more recent meta-analysis of leadership training with employees (Lacerenza et al., 2017) included 214 published reports and found positive effects of training. Training duration was found to be an important variable but details regarding the length of training and evaluations were not described.

Meetings seem to be a commonality for all organisations. Formal meetings are one of the major forums where leaders have opportunities to influence employees through, for example, providing information, giving instructions, monitoring work, and delivering performance consequences (i.e., feedback on specific tasks). Previous studies of leaders' self-reported distribution of their own work tasks have found that the leaders spend a substantial amount of their time in meetings, from 18 to 75 % (Arman et al., 2009; Florén and Tell, 2004; Kurke and Aldrich, 1983; Mintzberg, 1973; Tengblad, 2006; Tengelin et al., 2011; Wikström et al., 2013). Meetings provide opportunities for studying leadership behaviours (Bass and Stogdill, 1990) and for examining taxonomies of leadership behaviours.

In order to observe leadership in field settings, it is necessary to operationalise terms used in leadership theories, definitions and taxonomies. There is a profusion of taxonomies describing leadership behaviours (Yukl, 1994) and there are differences in the degree of detail and specificity in the categories included in the various taxonomies. Furthermore, the taxonomies have been developed through the use of different approaches for categorisation of leadership behaviour and this in turn affects how studies using a specific taxonomy are planned and conducted, and how data are collected and analysed.

The number of subjects included in field studies is an issue to be addressed when evaluating the validity of the results. Forsman and colleagues have demonstrated that statistically valid results can be obtained with a limited number of participants (Palmerud et al., 2012; Neumann et al., 2018). Conducting field studies in organisational settings poses several challenges, especially for researchers who wish to systematically and directly observe dynamic processes. Obtaining access to observe and record behaviour in organisational settings can be complicated, in part due to issues concerning security and integrity. The use of an external, live observer may have effects on the behaviours of interest, i.e. measurement reactivity, and the use of multiple observers to evaluate inter-observer agreement may increase the risk of such effects (Boyce et al., 2000). The logistical issues involved and costs of using one or more external observers represent additional challenges. Furthermore, live observation almost presupposes the use of a sampling procedure rather than continuous recording due to the complexity of observing and recording ongoing behaviours. Given the potential advantages that video-observations may provide, examining the feasibility of using video for leadership research in real-life work settings is warranted.

Komaki and Minnich (2002) and Lehmann-Willenbrock et al. (2015) are examples of studies using video to record supervisory behaviour in real life settings. These studies are innovative and unique in capturing temporal and situational aspects of supervisory behaviour, as well as interactions between supervisor-personnel. Komaki and Minnich (2002) video observations were used in sports to evaluate a sailing team and did not include professional supervisors and personnel in a work life setting. The study by Lehmann-Willenbrock et al. (2015) evaluated verbal interactions of leaders and team members from videotapes of 30 entire meetings from two organisations. They reported that transformational leadership style was associated with improved problem solving by

team members. The study is notable for analysing leadership at the level of actual events, rather than aggregating individual perceptions of leadership to the person level (Eberly et al., 2013; Hoffman and Lord, 2013).

The seminal research by Komaki and colleagues who developed the operant supervisory taxonomy and index (OSTI) serves as the basis for the study. Komaki et al. (1986) developed OSTI to study supervisors' behaviours, and the frequency of different interactions (with different quality levels) with their personnel. The taxonomy, with ties to organisational behaviour management (OBM), permits classification of all supervisory behaviour into one of seven mutually exclusive categories. Four of the OSTI-categories are related to the performance of employees and the remaining three are related to work but not to employee performance. In using OSTI, supervisors' behaviours are correlated to employee performance (Komaki, 1998; Johnson, 2013). A major part of OBM research in the area of supervisor-personnel interaction has dealt with feedback, a term with multiple definitions and behavioural functions (Alvero et al., 2001). Leaders' interactions with employees are one of several possible opportunities to provide feedback and/or consequences (Alvero et al., 2001). Providing feedback, or consequences, is considered to be an especially important supervisory behaviour (Alvero et al., 2001; DeNisi and Kluger, 2000; van Dierendonck et al., 2007).

The seven OSTI behavioural categories, with accompanying abbreviations, as defined by Komaki et al. (1986) are:

1 providing consequences (C)

2 monitoring performance (M)

3 providing performance antecedents, e.g., instructions, exhortations (A)

4 describing own performance (O)

5 talking about work related topics (W)

6 talking about non-work-related topics (N)

7 solitary behaviour (S)

(see Table 1 for examples). The first three (C, M, and A) are viewed as key behaviours most involved in effective supervision (Komaki et al., 1986). In this manuscript we focus on supervisors' reactions to employee reports about task performance, which Komaki described as providing consequences (C), monitoring performance (M), providing performance consequences (Komaki et al., 1986) To our knowledge, this is one of the most explicit methodologies for observing and categorising supervisory behaviour.

There are several innovative descriptive field studies using direct observation of supervisors in real-time, non-standardised situations (Avolio and Gardner, 2005; Brewer and Weber, 1994; Mattila et al., 1994; Methot et al., 1996). Meta-analyses encompassing thousands of articles on leadership concluded that very few studies have evaluated the effects of leadership interventions (Avolio and Gardner, 2005; Lacerenza et al., 2017). To our knowledge, Methot et al. (1996) is the only study to use OSTI, although in a modified form, to evaluate an intervention in a real-life setting.

This study focused on observing and classifying in greater detail what supervisors say to influence employees to accomplish desired work-related goals through their interpersonal verbal interactions. The OSTI was used to examine, via self-recorded videos, how supervisors used three important performance categories: antecedents,

monitors, and consequences in everyday work situations, i.e., in daily, ostensibly standardised, planned meetings. During the course (several months) of data collection, an intervention, as described below, was carried out to increase the frequency at which supervisors delivered performance consequences. The aim of this study was to answer the following research questions:

RQ1 How stable are supervisory behaviours at meetings with regards to both frequency and variability, individually and among individuals when a standardised meeting structure is used?

RQ2a Does an intervention focusing upon supervisors' use of performance consequences at meetings increase their use of this class of behaviours?

RQ2b Does an intervention focusing upon supervisors' use of performance consequences at meetings affect how they use performance monitors?

RQ2c Does an intervention focusing upon supervisors' use of performance consequences at meetings affect how they use performance antecedents?

RQ3 Is video self-recording and off-site coding with OSTI a useful method for collecting comprehensive data on the frequency and quality of supervisory behaviours in field settings?

## 2 Methods

### *2.1 Setting and participants*

The study was conducted in the Swedish research and development department of a multi-national technical company with more than 40,000 employees in 70 countries. Participants encompassed supervisors and staff from three different sections of this department located in two different cities, 500 kilometres apart. About one year prior to the data collection, the organisation had invested large amounts of time and money in introducing management practices which involved time-boxed iterative project cycles with continuous follow-ups. The intention was to improve uniformity of supervision and production. One of the newly introduced management practices involved daily meetings at the section level where employees reported on previous and coming performance and what hindrances they were facing. Department managers were unsure to what extent section supervisors applied the practices in a uniform matter. The practices entailed adhering to standardised formats, which at a general level dictated how supervisors' meetings with their staff were planned and conducted. Eight supervisors participated in the study. All eight had participated in the company's agile based management training program.

The study population included all supervisors ($n$ = 8) in the company's three sections, divided into groups by section 1 ($n$ = 3), 2 ($n$ = 2), and 3 ($n$ = 3). The original intention was to include nine supervisors. However, prior to the start of the intervention, one supervisor left the company and was not replaced. No data for this individual is reported here, seven men and one woman participated voluntarily in the study. They were all university-educated engineers, between the ages of 35 and 45, and with more than one-years' experience inside the company. Management experience with this and other

companies ranged from 1 to 11 years. Seven of the individuals having more than two years of experience. They all had identical work descriptions and a comparable position in the organisation hierarchy. The study was conducted with the informed written consent of the supervisors and the study design was approved by the Regional Ethic Review Board (2013/1093-31), Stockholm, Sweden.

The employees in the sections were responsible for developing high technological products with both software and hardware components. The 70 employees (five females/65 males; age range 23–62 years old) were university-educated engineers and worked both individually and in small teams.

### *2.2 Study design*

A multiple baseline design was chosen to make optimal use of the naturally occurring organisational structure, varied location of the three sections within the department and time constraints. The department at which the study was conducted granted access for a limited, ten-week period during which time there was an opportunity to describe and evaluate how supervisors performed their leadership at morning meetings and to attempt to increase supervisors' use of performance consequences at meetings. Data collection began simultaneously at all three sections. Due to time constraints, the interventions were begun at one-week intervals in random order among the three sections. The baseline periods were one, two and three weeks for the three sections. In Table 2, Sections 1, 2 and 3 respectively. The intervention periods were correspondingly nine, eight and seven weeks for the three sections.

### *2.3 Video-recordings*

The study is based on digital video-recordings of the supervisors, who filmed themselves. Filming was done for the entire daily morning meeting, with an easy-to-use compact digital video camera (Flip Mini Camcorder; Cisco, San Jose, USA) with HD-resolution. Each supervisor received instructions for the operation of the camera. They were instructed to set up the camera prior to the meeting and place it on the table next to them, at a distance of less than two metres. These instructions required less than five minutes per supervisor. The camera was placed so that only the supervisor was recorded visually, and all dialogue from the supervisor and employees was recorded on the audio track.

The contents of the cameras were transferred to external hard drives on a weekly basis. During the study, all supervisors missed one or more meetings due to conflicting obligations, illness, or vacation. No supervisor missed more than four sessions. All film material was treated as confidential material and was maintained by the first author.

The employees were given written information that the supervisor alone was being filmed as part of an educational program and that all dialogue recorded was for educational purposes only. All supervisors and employees agreed to data being collected under these conditions.

### *2.4 Coding*

Coding supervisors' behaviours was based on the OSTI developed by Komaki et al. (1986). Examples of the seven OSTI categories are provided in Table 1. Observation of supervisors involved checking for three of the seven categories:

- performance consequences (C)
- monitors (M)
- antecedents (A) in behaviour.

The recordings were coded by a trained coder. An OSTI training manual developed by Komaki was used to train coders, who observed video sequences from other settings (J. Komaki, personal communication, September 1992). The authors identified the frequencies of the three categories to be coded as they occurred in the videos used for training. Each coder was required to achieve an accuracy of 90% and an inter-observer agreement of 90%. Approximately 120 hours of training was required per coder to achieve an acceptable level of accuracy and agreement.

Video recordings was chosen over only audio recording because of the added value that observed hand movements and facial expressions was believed to contribute to in respect to the accuracy of the coding (Goltz, 1993) and to ensure that it was the supervisors' speech that was coded. Non-verbal behaviours, such as thumbs up or nods, were coded if the gesture was clearly assessed as influencing the employee. Moreover, the use of video recording made it possible to code all supervisory behaviour that occurred during each meeting. The person coding could watch all behaviour as many times as required before determining a code. Employee responses were not coded.

Approximately 90 hours of video material was obtained from more than 360 separate daily meetings. To obtain a manageable amount of data, a decision was reached to code two of the five films per week which produced a potential 160 meetings for coding. The selection of films to code was random, with the restriction that there should be at least one day between meetings. A total of 142 separate meetings were coded. No data were available for the remaining 18 meetings due to supervisor absences as described above. The number of behaviours coded ranged from 0–14 per minute with a mean of 2,5 per minute. Coding a 15-minute video required about 60 minutes for a trained observer, in large part due to repeatedly viewing passages until they were able to determine the specific category classification.

### *2.5 Intra- and inter-observer agreement*

From 142 films, five minutes from 12 randomly selected films were coded on two separate occasions by the same coder with a minimum interval of three weeks between assessing intra-observer agreement. Inter-observer agreement was assessed with a second coder evaluating the same 12 randomly selected films. Coding of a behaviour within a time frame of +/- one second and the same specific behaviour category was considered an agreement. Intra- and inter-observer agreements were calculated using the formula:

$$\frac{\#\,agreement}{\#\,of\ behaviours\ coded} \times 100 = \%$$

Overall intra-observer agreement was 94% and inter-observer agreement was 91% when tested.

**Table 1** OSTI categories and definitions, based on Komaki et al. (1986) and examples drawn from dialogues from meetings at the site

| *Category* | *Definition* | *Examples* |
|---|---|---|
| Performance consequence (C) | Indicates knowledge of performance | E: I had to add some assembly lines to fix the issue. |
| | | Su: Ok. (C) |
| | | Su: We restarted them yesterday. |
| | | Su: Mm. (C) |
| | | Su: Thomas made a change there. (C) |
| | | E: Yesterday I finished the documents and built a release. |
| | | Su: Excellent! (C) |
| | | E: Yesterday I worked on the test reports. |
| | | Su: Mm. (C) |
| Performance monitors (M) | Collects information about an employee's performance | Su: Have you sent it to the product owner for a review yet? (M) |
| | | E: Yesterday I fixed a bug in the system. |
| | | Su: Which one was that? (M) |
| | | Su: Are you done with all the tasks here? (M) |
| | | Su: What did you do yesterday? (M) |
| Performance antecedents (A) | Instructs, reminds, or conveys an expectation of performance | Su: Look into that. (A) |
| | | Su: Tell them to read it and give comments. (A) |
| | | Su: Check that with Nicholas and Kim. (A) |
| | | Su: Test it and then you can close that one. (A) |
| | | Su: We have to get someone to help us with that. (A) |

*OSTI categories not coded in this study*

| *Category* | *Definition* | *Examples of statements/interactions* |
|---|---|---|
| Own performance (O) | Refers to his or her own performance | Su: I did that yesterday. (O) |
| Work related (W) | Refers to the work but not to employee performance | Su: We have clients visiting us tomorrow. (W) |
| Non-work related (N) | Does not refer to work issues or concerns | Su: Did you see the game yesterday? (N) |
| Solitary (S) | Not oriented towards or attending to other individuals, and/or other individuals are not oriented towards or attending to the leader | |

Notes: Su = Supervisor, E = Employee, C = Consequence, M = Monitoring, A = Antecedent, O = Own performance, W = Work related, N = Non-work related and S = Solitary. Employees' statements are abridged, and the employees' names are pseudonyms.

### 2.6 Baseline

Supervisors conducted daily production meetings in accordance with standardised Scrum-procedures as determined by division management. The intention was that the supervisor of each group was to follow up on performed tasks, identify what was left to do, estimate time spent and time left and distribute necessary resources within the group. The supervisors were expected to use the following scrum-based questions to collect reports during the meetings: "What did I do in the previous working day?", "What do I plan to do today?", and "What impediments am I facing?" (Meyer, 2014). Meeting conditions were identical with regards to format (daily, start of workday; Monday to Friday, maximum 15 minutes), a common software program to log work tasks and performance and frameworks/agendas for conducting the meetings in traditional, enclosed company office space. None of the supervisors received instructions or training regarding how antecedents, monitoring, or consequences were to be given.

### 2.7 Intervention

The supervisors continued to conduct daily production meetings, using the same Scrum-procedures as during baseline. The intervention to influence supervisory behaviour was developed to meet the organisational needs as described by the departmental management, who also selected the day and time for the intervention sessions. The intervention included crucial components for leadership training as described by Lacerenza et al. (2017); sessions spaced over time, feedback, on-site training, multiple delivery methods and was adapted to existing structures, routines and processes at the organisation.

The intervention focused upon increasing supervisors' use of performance consequences at meetings. Sections were added successively on a weekly basis (Figure 1). During the first week of the intervention, the three supervisors from Section 1 participated, during the second week they were joined by the supervisors from Section 2 by using video conferencing, and during the third, also the supervisors from Section 3 participated. All supervisors participated together in these weekly sessions for the remainder of the study. Hence, three supervisors participated in all 9 sessions, two supervisors in 8 sessions and the remaining three supervisors in 7 sessions.

Each 30-minute intervention session used the scrum-based-questions for individual reports (Meyer, 2014) that the supervisors were required to use at their own daily morning meetings with employees, with the difference that they now concerned the entire previous week. The individual reports involved each supervisor describing;

1 how they had provided performance consequences at their daily meetings the preceding week
2 how they planned to provide consequences in upcoming meetings with employees
3 what obstacles there were for increasing the use of performance consequences in interactions with employees they supervised
4 responding to performance consequence related questions from their peers.

The format for each weekly 30-minute session, held Tuesday afternoons, was as follows; individual reports from each supervisor (10 minutes), two video examples for viewing

and discussion (15 minutes), and dialogue on reactions and reflections (5 minutes). The role of the first author was to initiate the individual reports, take notes on the individual reports, keep time, ensure active participation and together with the participants provide positive feedback on effort and progress. In addition, supervisors were encouraged by the first author to develop their own activities for the coming session. Following each session, the first author sent an e-mail to all participants summarising individual reports and the participants self-determined activities.

To illustrate positive or negative examples of two supervisors' attempts to deliver performance consequences, examples were chosen by the first author from video the supervisors had submitted the previous week. The examples were selected prior to and separately from coding. After the viewing, the two supervisors commented on their performance and reflected upon alternative behaviours. Through questions to the group, the first author then facilitated the supervisors' analyses and discussions focusing solely on aspects of timing, delivery and formulation of performance consequences. During the final five minutes, there was an open dialogue when the first author encouraged supervisors to express their reactions to and reflections upon the use of performance consequences.

## 3 Evaluation of the intervention

For statistical analyses non-parametric test were used, since the data were believed to be non-normally distributed. The supervisors' use of performance consequences, antecedents and monitors during baseline and intervention periods were analysed using Spearman's rank correlation to examine correlations among the three categories and Wilcoxon signed-rank test to evaluate changes in each of the three categories (MATLAB R2021b; MathWorks, Inc., Natick, MA, USA; www.mathworks.com). The mean numbers for each supervisor during the baseline and intervention periods were compared for the three categories. The analyses adopted a significance level of 5% ($p < 0.05$), and the tests were 2-sided.

## 4 Results

The three supervisory behaviours, use of performance consequences, antecedents, and monitors prior to and throughout the intervention are shown in Table 2 and Figure 1. Table 2 shows the mean number of these behaviours delivered by each of the eight supervisors, during the baseline and intervention phases.

The mean number of performance consequences delivered, the focus of the intervention, increased for all eight supervisors during the intervention phase ($p = .008$). The mean change from the baseline to the intervention period for individual supervisors ranged from 2.0 to 6.7. As compared to baseline periods, the mean number of monitors and antecedents used during the intervention phase showed non-significant decreases ($p = .844$) and ($p = .063$), respectively. However, a significant correlation was found for the degree of change in the use of performance consequences and monitors ($p = .015$). The supervisors who showed the greatest increase in the use of performance consequences also showed increases in the use of performance monitors.

**Table 2** Mean number of performance consequences, monitors and antecedents delivered by each of the eight supervisors (A-H) from the three Sections (1-3) during baseline and intervention

| *Section* | *Supervisor* | *Consequences* | | | *Monitors* | | | *Antecedents* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | *Baseline* | *Intervention* | *Change* | *Baseline* | *Intervention* | *Change* | *Baseline* | *Intervention* | *Change* |
| 1 | A | 17.0 | 19.1 | 2.1 | 6.5 | 4.8 | –1.7 | 5.0 | 5.9 | 0.9 |
| | B | 8.0 | 14.0 | 6.0 | 5.0 | 5.8 | 0.8 | 6.5 | 5.4 | –1.1 |
| | C | 0.0 | 6.1 | 6.1 | 3.0 | 7.5 | 4.5 | 3.5 | 4.1 | 0.6 |
| 2 | D | 3.5 | 8.5 | 5.0 | 3.5 | 4.5 | 1.0 | 3.8 | 0.8 | –3.0 |
| | E | 0.8 | 7.5 | 6.7 | 2.5 | 6.2 | 3.7 | 2.8 | 2.2 | –0.6 |
| 3 | F | 8.3 | 10.6 | 2.3 | 16.5 | 10.2 | –6.3 | 5.5 | 3.8 | –1.7 |
| | G | 3.0 | 5.0 | 2.0 | 3.5 | 1.4 | –2.1 | 4.2 | 1.7 | –2.5 |
| | H | 2.5 | 6.4 | 3.9 | 6.0 | 4.2 | –1.8 | 2.8 | 0.8 | –2.0 |

Notes: Data based on two meetings per week, randomly selected, for ten weeks, 142 meetings in total. The baseline period was one week for Section 1, two weeks for section 2 and three weeks for section 3. The intervention periods were correspondingly nine, eight and seven weeks for the three sections. Analysis with Wilcoxon signed ranks test showed an increase in mean number of performance consequences delivered increased during the intervention phase ($p = .008$). As compared to baseline periods, the mean number of monitors and antecedents used during the intervention phase showed a non-significant decrease for four ($p = .844$) and six ($p = .063$) supervisors, respectively.

**Figure 1** Frequency of using performance consequences, monitors and antecedents by the eight supervisors at meetings during baseline and the intervention periods (data from two randomly selected meetings per week for the ten-week period) and means for each category

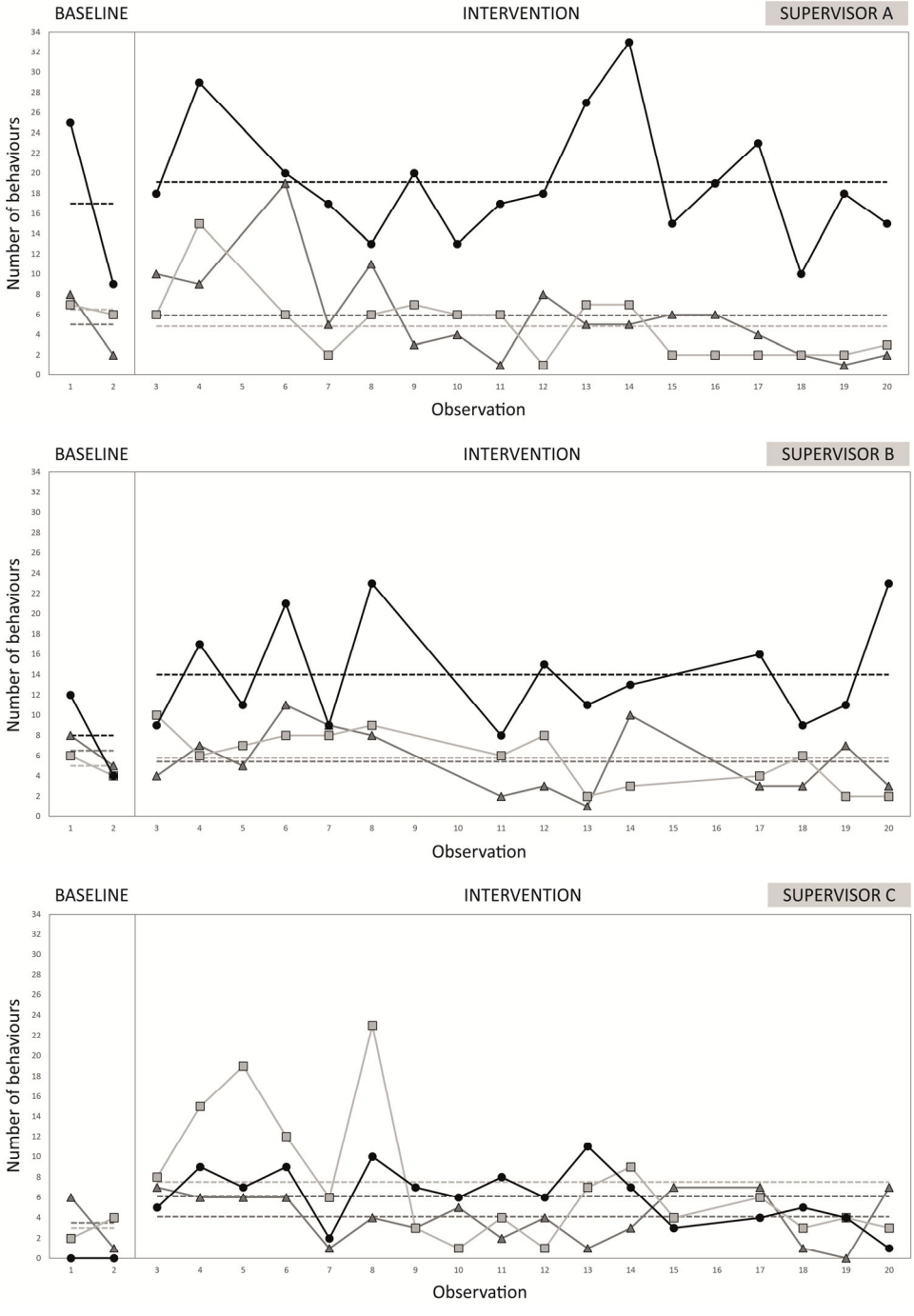

**Figure 1** Frequency of using performance consequences, monitors and antecedents by the eight supervisors at meetings during baseline and the intervention periods (data from two randomly selected meetings per week for the ten-week period) and means for each category (continued)

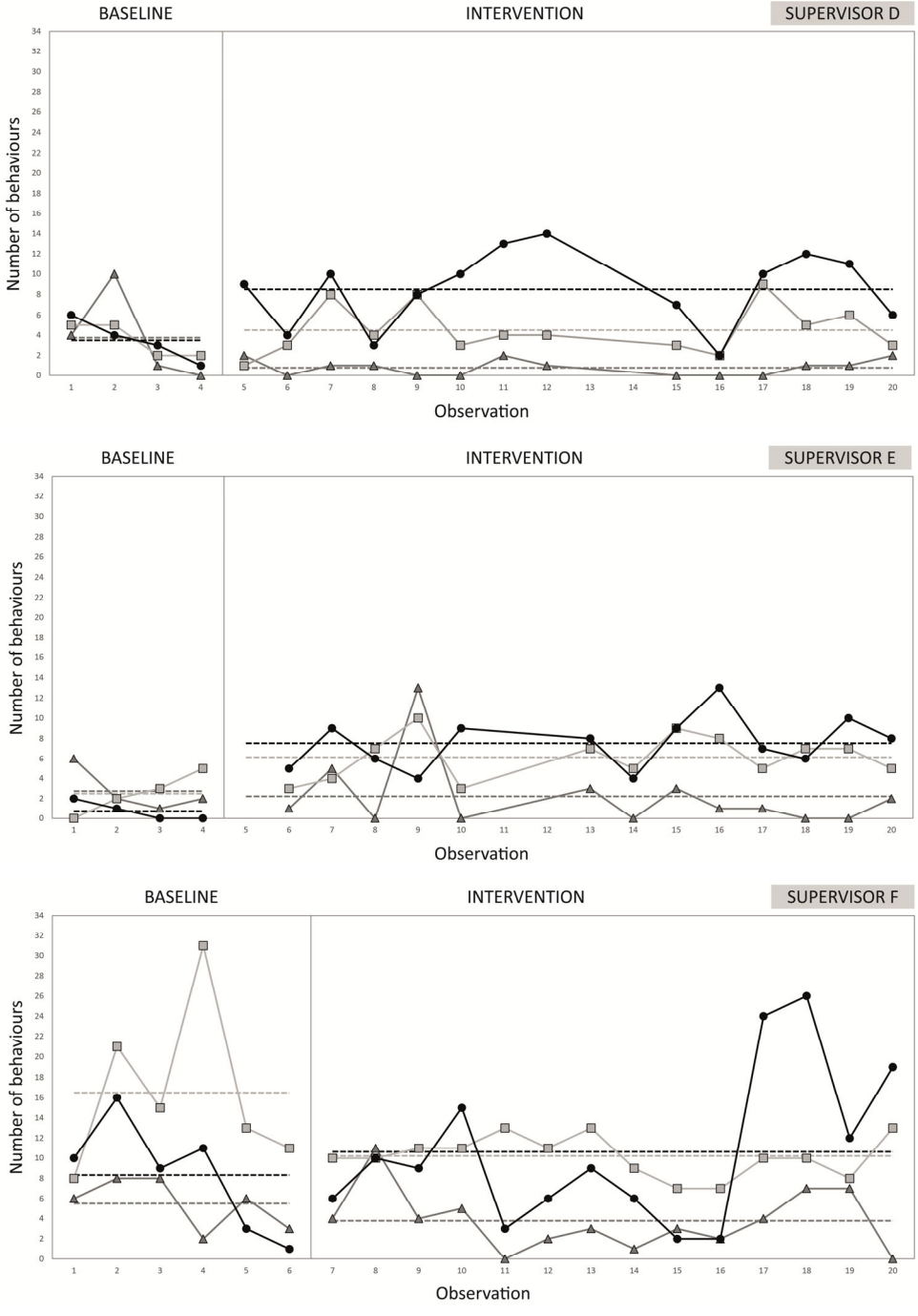

**Figure 1** Frequency of using performance consequences, monitors and antecedents by the eight supervisors at meetings during baseline and the intervention periods (data from two randomly selected meetings per week for the ten-week period) and means for each category (continued)

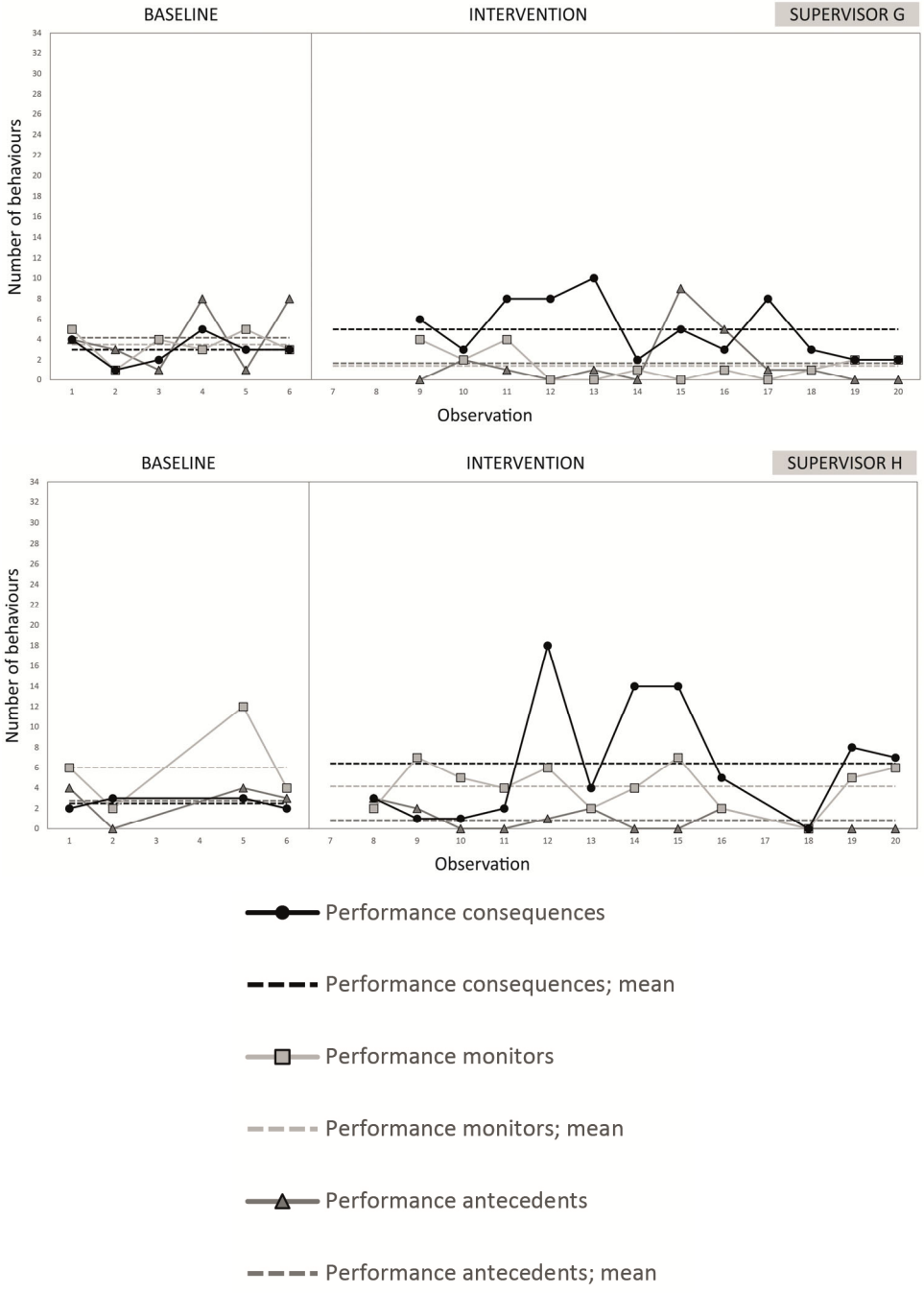

Figure 1 shows that the frequency at which the three behaviours were used by the supervisors varied both within and between supervisors during the baseline and intervention periods.

## 5 Discussion

The results showed individual variation in supervisory behaviours at meetings during both baseline and intervention periods, despite the continued, systematic use of a standardised meeting structure (RQ1). The supervisors´ use of a crucial supervisory behaviour, providing performance consequences, increased for all supervisors, and significantly at a group level, in conjunction with the intervention focusing on this behaviour (RQ2a). The results with regards to the use of performance monitors are mixed. The four supervisors showing the greatest increase in the use of performance consequences also showed increases in the use of performance monitors. The remaining four supervisors also showed increases in their use of performance consequences but there was a decrease in their use of performance monitors (RQ2b). The intervention does not appear to have influenced the use of performance antecedents by supervisors (RQ2c). The findings show that it is possible to use video self-recording and off-site coding with OSTI in field settings to evaluate the effects of supervisory training interventions for extended periods (RQ3).

Ideally, a larger number of supervisors and a lengthier base-line period would have provided more information about supervisor's use of supervisory behaviours in the standardised meeting structure being used. All supervisors at this division of the company were included. Conditions at the participating company precluded the option of using a lengthier baseline period, this since the company granted access for only ten weeks. This design did however permit evaluation of the intervention phase for seven to nine weeks, with statistically significant results.

Video recordings were used in this study. Video recordings provide potential solutions to the crucial issues of attaining accuracy and completeness of measuring behaviour (Johnston and Pennypacker, 2010). The present study extends the research by Komaki and Citera (1990), Komaki and Minnich (2002) and Fraidlin et al. (2023) who used video technology and interval sampling procedures. Video recording in the field made it possible to have a permanent record of the dynamics of all supervisory behaviour in a specific context and to code subsequently off-site. One of the major advantages with video technology is that observing, and coding do not have to be done simultaneously. This is important since the behavioural requirements of the tasks of observing and coding/scoring can be incompatible, such that either completeness or accuracy, or both may suffer (Johnston and Pennypacker, 2010). The permanent nature of video allows unlimited opportunities to replay the recordings, which for example enabled reviewing of diffuse sequences, until complete agreement was achieved.

The complete dynamics of meetings were obtained by recording for the entire meeting. Continuous observation and coding/scoring with OSTI for the entire meeting captured from 6 to 14 specific behaviours per minute as opposed to one behaviour per minute in previous studies using live sampling procedures. The results from two meetings per week for ten weeks are believed to provide realistic depiction of the leadership as performed in this context.

When sampling is used there is a risk that crucial supervisory behaviours such as using performance monitors and consequences go unobserved and unrecorded, since these occur relatively infrequently. Video recording for entire meetings over time and the use of OSTI made it possible to detect low frequency behaviours and variability. In the OSTI-coding taxonomy (Komaki et al., 1986) using performance consequences and performance monitoring, although they occur at low frequencies, are considered to be crucial for supporting employee performance and eliciting positive attitudes among employees (Komaki, 1998). To our knowledge, only one previous study has used OSTI in combination with video recordings, to evaluate a leadership intervention (Methot et al., 1996). Methot and colleagues did not use self-recording, coded only performance consequences and monitors and also used a sampling procedure.

Video self-recording by supervisors was shown to be a feasible solution to the logistical and methodological issues to be dealt with when studying supervisory behaviour in the field. An advantage with video self-recording is that no external personnel are needed at the site, which minimises the risk for observer effects upon both supervisors and employees. The costs of observer time, including travel to and from sites, have been minimised due to the use of self-recording, while the time and cost of observer training seem comparable to figures reported in other studies. Had we not used video self-recording, a minimum of eight pre-trained observers would have been required for conducting simultaneous live observations at the different work sites and to evaluate interobserver agreement.

Agile based methods (Meyer, 2014) represent one tenet of many organisational philosophies. In this tenet meeting content is standardised, and setting preconditions are essential components of the standardisation. This company had made significant investments based on this reasoning. The results from these three sections indicate that supervisors demonstrate both intra- and inter-subject variability in key supervisory behaviours in a setting where uniformity was desired and expected, given the standardised format and the agile based management training they had received before the intervention. Focusing upon increasing the use of specific supervisory behaviours can be an important complement to other organisational measures.

In this study, the supervisors who showed increases in performance monitors also showed the greatest increase in delivering performance consequences. Opportunities to provide performance consequences can be expected to increase with increased use of performance monitors. The four supervisors increasing their use of monitors during the study might have developed that strategy, even though it was not focused upon in the intervention. Considering that most previous leadership interventions last less than one month, and commonly used self-reported data (Avolio and Gardner, 2005), the present study with ten weeks of observed supervisory behaviours, contributes to valuable insights about actual supervisory behaviour over time.

While it is possible that the differences in supervisors' behaviours are specific to the three sections at this company, it seems more probable that the differences in supervisory behaviour would be even larger in organisations with fewer standardised conditions. For example, in organisations without pre-determined meeting lengths or without set meeting protocols, variability would seem more likely. This study collected data from a fixed meeting setting and is therefore not indicative of all supervisory behaviour; we do not know how these supervisors lead in other settings or more spontaneous conditions. As there is likely to be more variability outside of a standardised meeting setting, it might also be assumed that supervisors will also vary in their performance outside of a fixed

meeting setting. These findings reveal the necessity of using a sensitive measurement technology and repeated measurement to study supervisory behaviour. The topic of variability of supervisory behaviour at performance-level has been sparsely considered in previous research. As stated by Komaki and Minnich (2002), unless data are collected repeatedly, it is not even possible to reflect on how frequently supervisors' behaviours occur and how they might vary from day to day.

The results show that a realistic intervention using components proposed to be crucial for leadership training (Lacerenza et al., 2017) can influence an important supervisory behaviour, although the results were varied for the three key supervisory behaviours. While the mean for performance consequences delivered increased for all supervisors following the intervention focusing on this behaviour, the changes were neither immediate nor stable.

Furthermore, permanent video recording offers the possibility for coding by external parties or for coding the same material with more than one taxonomy. This latter possibility should be of special interest to researchers. The study by Lehmann-Willenbrock et al. (2015) evaluated supervisor-subordinate verbal interactions. The present study did not categorise employee statements in a systematic manner although this is possible given the permanent nature of the material. Another possible practical application might be to study safety climate as manifested in organisations. Clarke (2013) proposed that leadership styles have a differential effect on safety compliance and safety participation indicating that there are links between supervisory behaviours and their subsequent influence on employee behaviour.

Further research into how specific behaviours relate to how supervisors lead is needed since this could provide a better understanding of what leadership programs should be focusing upon to produce desired behaviours and outcomes, and the connection between the two. Field studies are necessary in order to observe what supervisors actually do to influence and lead employees (Barley and Kunda, 2001; Alvesson, 2019).

The availability, simplicity and low cost of digital video equipment (e.g. inexpensive or even own smartphones) should make it an attractive technology for researchers and practitioners interested in acquiring data in an endless variety of occupational or organisational settings. The use of video self-recording permitted the unobtrusive collection of extensive, permanent data on supervisors' leadership behaviours in a field setting. The comprehensive permanent record of the dynamics of supervisory behaviour could subsequently be analysed with OSTI to assess the frequency and quality of the supervisors' behaviours.

Video self-recording and OSTI can be used in daily practise and ergonomic research to study dynamic social interactions that influence both the work environment and productivity. The combination makes it possible to measure, evaluate and compare supervisory behaviour in greater detail than previously, and could be extended to study interactions between supervisors and their employees. The methodology used in this study can also be used to examine the relationship between leadership as perceived (e.g. results from questionnaires and interviews) and leadership as performed.

## Ethical considerations

All video filming was done with the informed written consent of the participants and the study design was approved by the Regional Ethic Review Board (2013/1093-31),

Stockholm, Sweden. The research reported here does not necessarily reflect the views of the Swedish Association of Local Authorities and Regions.

## Acknowledgements

Thanks to Judith Komaki for invaluable advice on the OSTI and to Leif Lund and Kalle Hedqvist for assistance in the preparation of the manuscript. Thanks to the company and all the leaders for their participation.

## References

Alvero, A.M., Bucklin, B.R. and Austin, J. (2001) 'An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985-1998)', *Journal of Organizational Behavior Management*, Vol. 21, No. 1, pp.3–29.

Alvesson, M. (2019) 'Waiting for Godot: eight major problems in the odd field of leadership studies', *Leadership*, Vol. 15, No. 1, pp.27–43.

Alvesson, M. and Einola, K. (2019) 'Warning for excessive positivity: authentic leadership and other traps in leadership studies', *The Leadership Quarterly*, Vol. 30, No. 4, pp.383–395.

Alvesson, M. and Spicer, A. (2012) 'Critical leadership studies: the case for critical performativity', *Human Relations*, Vol. 65, No. 3, pp.367–390.

Arman, R., Dellve, L., Wikström, E. and Törnström, L. (2009) 'What health care managers do: applying Mintzberg's structured observation method', *Journal of Nursing Management*, Vol. 17, No. 6, pp.718–729.

Avolio, B.J. and Gardner, W.L. (2005) 'Authentic leadership development: getting to the root of positive forms of leadership', *The Leadership Quarterly*, Vol. 16, No. 3, pp.315–338.

Avolio, B.J., Reichard, R.J., Hannah, S.T., Walumbwa, F.O. and Chan, A. (2009) 'A meta-analytic review of leadership impact research: experimental and quasi-experimental studies', *The Leadership Quarterly*, Vol. 20, pp.764–784.

Banks, G.C., Gooty, J., Ross, R.L., Williams, C.E. and Harrington, N.T. (2018) 'Construct redundancy in leader behaviors: a review and agenda for the future', *The Leadership Quarterly*, Vol. 29, No. 1, pp.236–251.

Barley, S.R. and Kunda, G. (2001) 'Bringing work back in', *Organization Science*, Vol. 12, No. 1, pp.76–95.

Bass, B.M. and Stogdill, R.M. (1990) *Bass & Stogdill's Handbook of Leadership: Theory, Research, and Managerial Applications*, Simon and Schuster, The Free Press, New York.

Berglund, M., Karltun, A., Eklund, J. and Karltun, J. (2020) *HTO – A Concept of Humans, Technology and Organisation in Interaction*, Linköping University Electronic Press, Helix Competence Centre, Linköping University, Linköping.

Bommer, W.H., Dulebohn, J.H. and Wu, D. (2018) 'Do ethical, authentic, and servant leadership explain variance above and beyond transformational leadership? A meta-analysis', *Journal of Management*, Vol. 44, No. 2, pp.501–529.

Boyce, T.E., Carter, N. and Neboschick, H. (2000) 'An evaluation of intraobserver reliability versus interobserver agreement', *European Journal of Behavior Analysis*, Vol. 1, No. 2, pp.107–114.

Brewer, M.B. and Weber, J.G. (1994) 'Self-evaluation effects of interpersonal versus intergroup social comparison', *Journal of Personality and Social Psychology*, Vol. 66, No. 2, pp.268–275.

Clarke, S. (2013) 'Safety leadership: a meta-analytic review of transformational and transactional leadership styles as antecedents of safety behaviours', *Journal of Occupational and Organizational Psychology*, Vol. 86, No. 1, pp.22–49.

Denisi, A.S. and Kluger, A.N. (2000) 'Feedback effectiveness: can 360-degree appraisals be improved?', *Academy of Management Perspectives*, Vol. 14, No. 1, pp.129–139.

Eberly, M.B., Johnson, M.D., Hernandez, M. and Avolio, B.J. (2013) 'An integrative process model of leadership: Examining loci, mechanisms, and event cycles', *American Psychologist*, Vol. 68, No. 6, pp.427–443.

Eklund, J. (2000) 'Towards a framework for quality of interactions between humans, technology and organization', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 44, pp.2-463–2-466.

Florén, H. and Tell, J. (2004) 'What do owner-managers in small firms really do? Differences in managerial behavior in small and large organizations', *Small Enterprise Research*, Vol. 12, No. 1, pp.57–70.

Fraidlin, A., Van Stratton, J.E., Mcelroy, A. and Aljadeff, E. (2023) 'Peer feedback: recommendations for behavior analysts' training and supervision', *Behavior Analysis in Practice*, Vol. 16, No. 3, pp.696–708.

Goltz, S.M. (1993) 'Dynamics of leaders' and subordinates' performance-related discussions following monitoring by leaders in group meetings', *The Leadership Quarterly*, Vol. 4, No. 2, pp.173–187.

Hoffman, E.L. and Lord, R.G. (2013) 'A taxonomy of event-level dimensions: implications for understanding leadership processes, behavior, and performance', *The Leadership Quarterly*, Vol. 24, No. 4, pp.558–571.

Johnson, D.A. (2013) 'A component analysis of the impact of evaluative and objective feedback on performance', *Journal of Organizational Behavior Management*, Vol. 33, No. 2, pp.89–103.

Johnston, J.M. and Pennypacker, H.S. (2010) *Strategies and Tactics of Behavioral Research*, Routledge, New York.

Karltun, J., Karltun, A., Havemose, K., Coelho, D.A., Kjellström, S. and Sandback, A. (2022) *Handbook for Front Line Managers*, Jönköping, Jönköping University, Jönköping University, School of Engineering, Jönköping.

Komaki, J.L. (1998) *Leadership from an Operant Perspective*, Routledge, London and New York.

Komaki, J.L. and Citera, M. (1990) 'Beyond effective supervision: identifying key interactions between superior and subordinate', *The Leadership Quarterly*, Vol. 1, No. 2, pp.91–105.

Komaki, J.L. and Minnich, M.R. (2002) 'Crosscurrents at sea: the ebb and flow of leaders in response to the shifting demands of racing sailboats', *Group & Organization Management*, Vol. 27, No. 1, pp.113–141.

Komaki, J.L., Zlotnick, S. and Jensen, M. (1986) 'Development of an operant-based taxonomy and observational index of supervisory behavior', *Journal of Applied Psychology*, Vol. 71, No. 2, pp.260–269.

Kurke, L.B. and Aldrich, H.E. (1983) ''Note—Mintzberg was right!: A replication and extension of the nature of managerial work', *Management Science*, Vol. 29, No. 8, pp.975–984.

Lacerenza, C.N., Reyes, D.L., Marlow, S.L., Joseph, D.L. and Salas, E. (2017) 'Leadership training design, delivery, and implementation: a meta-analysis', *Journal of Applied Psychology*, Vol. 102, No. 12, pp.1686–1718.

Lehmann-Willenbrock, N., Meinecke, A.L., Rowold, J. and Kauffeld, S. (2015) 'How transformational leadership works during team interactions: a behavioral process analysis', *The Leadership Quarterly*, Vol. 26, No. 6, pp.1017–1033.

Mattila, M., Hyttinen, M. and Rantanen, E. (1994) 'Effective supervisory behaviour and safety at the building site', *International Journal of Industrial Ergonomics*, Vol. 13, No. 2, pp.85–93.

Methot, L.L., Williams, W.L., Cummings, A. and Bradshaw, B. (1996) 'Measuring the effects of a manager-supervisor training program through the generalized performance of managers, supervisors, front-line staff and clients in a human service setting', *Journal of Organizational Behavior Management*, Vol. 16, No. 2, pp.3–34.

Meuser, J.D., Badura, K., Cogliser, C., Gardner, W.L., Newman, D.A., Sin, H-P., Gottfredson, R., Halliday, C.S., Scandura, T.A. and Wright, S. (2022) *Current Developments in Leader Member Exchange (LMX): An Interactive Incubator Panel Symposium*.

Meyer, B. (2014) *Agile!: The Good, the Hype and the Ugly*, Springer Science & Business Media, Switzerland.

Mintzberg, H. (1973) *The Nature of Managerial Work*, Harper & Row, New York.

Neumann, P.W., Winkel, J., Palmerud, G. and Forsman, M. (2018) 'Innovation and employee injury risk in automotive disassembly operations', *International Journal of Production Research*, Vol. 56, No. 9, pp.3188–3203.

Palmerud, G., Forsman, M., Neumann, P.W. and Winkel, J. (2012) 'Mechanical exposure implications of rationalization: a comparison of two flow strategies in a Swedish manufacturing plant', *Applied Ergonomics*, Vol. 43, No. 6, pp.1110–1121.

Paul, J., Costley, D.L., Howell, J.P., Dorfman, P.W. and Trafimow, D. (2001) 'The effects of charismatic leadership on followers' self-concept accessibility', *Journal of Applied Social Psychology*, Vol. 31, No. 9, pp.1821–1842.

Salvendy, G. and Karwowski, W. (2021) *Handbook of Human Factors and Ergonomics*, John Wiley & Sons, Hoboken, New Jersey.

Tengblad, S. (2006) 'Is there a 'new managerial work'? A comparison with Henry Mintzberg's classic study 30 years later', *Journal of Management Studies*, Vol. 43, No. 7, pp.1437–1461.

Tengelin, E., Arman, R., Wikström, E. and Dellve, L. (2011) 'Regulating time commitments in healthcare organizations: managers' boundary approaches at work and in life', *Journal of Health Organization and Management*, Vol. 25, No. 5, pp.578–599.

Van Dierendonck, D., Haynes, C., Borrill, C. and Stride, C. (2007) 'Effects of upward feedback on leadership behaviour toward subordinates', *Journal of Management Development*, Vol. 26, No. 3, pp.228–238.

Van Knippenberg, D. and Sitkin, S.B. (2013) 'A critical assessment of charismatic – transformational leadership research: back to the drawing board?', *The Academy of Management Annals*, Vol. 7, pp.1–60.

Wikström, E., Arman, R. and Dellve, L. (2013) 'Vad gör chefer med sin tid och hur kan tid och engagemang hanteras på ett mer hållbart sätt?', *Socialmedicinsk tidskrift*, Vol. 90, No. 6, pp.830–837.

Yukl, G. (1999) 'An evaluative essay on current conceptions of effective leadership', *European Journal of Work and Organizational Psychology*, Vol. 8, No. 1, pp.33–48.

Yukl, G.A. (1994) *Leadership in Organizations*, Prentice Hall, Hoboken, New Jersey.