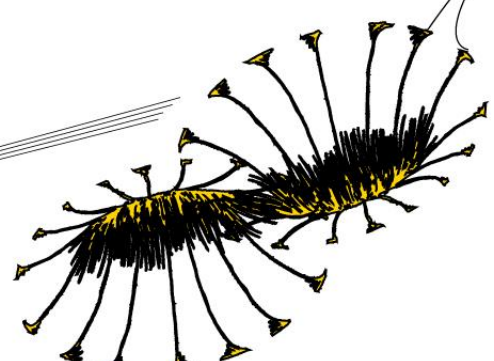



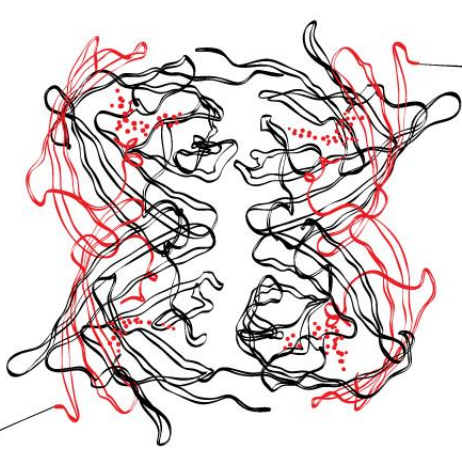


MASTER THESIS

THE PERCEPTION OF STANDARDS AND
INTERPRETATION OF RATING SCALE IN EMPLOYEE
PERFORMANCE ASSESSMENT



Graduation committee
T. Hirschler, MSc
Dr. M.A. Hendriks



A. Hordijk

26-6-2016

UNIVERSITY OF TWENTE.



Table of Contents

Acknowledgement.....	3
Abstract	4
Introduction	5
Exploration and definition of the (research) problem	7
Description of the organisational context.....	11
Research approach.....	13
Results	15
Conclusion and discussion	22
Recommendations and implications for practice.....	25
Reference list.....	27
Appendices	30
Appendix A: Overview of the organisation.....	30
Appendix B: Job description	31
Appendix C: Interview questions and topics.....	32
Appendix D: Analysis for the second interview question	33

Acknowledgement

Through this acknowledgement I would like to express thanks to the University of Twente and specifically the teachers and supervisors connected to this research. Tim, as my final first supervisor in particular for helping me get started again and providing me with the deadlines I needed to finish this project and Maria as my second supervisor. Also, I would like to express thanks to the municipality of Lelystad (and specifically Claudia) for allowing the research to take place within their organisation and not only cooperation but also stimulating the participation of employees in this research. On a more personal note, thanks are in order to Koen who has been by my side during the entire process of writing this thesis and Marjan & Martin who let me stay with them during my time in Lelystad. Finally, thanks to every employee of the P&O department at the municipality of Lelystad for welcoming me into their team as if I was one of their own and the ladies of Switch and Huub in particular who helped me find my footing, guided me through the first couple of weeks but mostly made me feel at home and made me laugh.

Abstract

In recent years, organisations became convinced that not the physical capital, but the employees working in an organisation (the human capital) are their most important asset. Identifying the human capital of an organisation is most often done through an employee performance assessment (EPA). In every form of EPA it is very important to reduce subjective judgement, especially when only one assessor is involved. In this study, similarities and differences were examined between managers and employees in the perception of standards and the interpretation of rating scale for employee performance assessment. Managers and employees turned out to have a different perception of standards for EPA. Where managers are familiar with the standards for EPA and focus on interpersonal relations, employees are often unfamiliar with EPA standards and are not aware managers focus on interpersonal relations. Managers also use standards other than those described in the official standards where most employees do not think extra standards are used. Results show a difference in not only the interpretation between manager and employee, but also a difference between individuals. The majority of interviewees were unable to completely describe desired behaviour for each of the six rating scale categories. Despite the differences, there are also some similarities between managers and employees when it comes to the interpretation of the rating scale. The topics 'happy customer' and 'communication' were both named by multiple managers and employees in the description of desired employee behaviour. These findings suggest that the differences in perceived standards and definition of rating scale are issues in which clear communication is of the essence. Also it should be clear in advance which standards are used during EPA and which are excluded from assessment.

Keywords: Employee performance assessment, job description, perception, interpretation, standards, rating scale, municipality

The value of an organisation can be expressed in many different ways. During the 1950's, the focus was placed mainly on factors such as physical capital, labour, land, and management (Mincer, 1962 in Nafuhko, Hairston & Brooks, 2004). Since then, it became clear that these factors were not always sufficient to explain the growth of the economy (Denison, 1962 in Nafuhko et. al., 2004). In recent years, organisations became convinced that not the physical capital, but the employees working in an organisation (the human capital) are their most important asset (Nafuhko et. al., 2004). It turned out that investing in employees positively influences productivity and return on investment for organisations (Psacharopoulos, 1985).

The idea that the human capital is the most important asset in an organisation, is introduced in the human capital theory. It perceives schooling and education of employees as deliberate investments of an organisation in its human capital and states that these investments will increase the productivity of the employees (Nafuhko et. al., 2004). Doing this, the productivity of the entire organisation will increase (Nafuhko et. al., 2004). Of course, not all organisations have unlimited resources to invest in their employees. To ensure they invest in their employees efficiently, organisations need to be aware of the employees (human capital) most likely to return the investment. Identifying the human capital of an organisation is most often done through an employee performance assessment (EPA) (Judge & Ferris, 1993).

Jacobs, Karfy & Zedeck (1980) identify six purposes for employee performance assessment: 1) disciplinary action; 2) feedback / employee development; 3) promotion; 4) selection; 5) training / supervision; and 6) organisational diagnosis and development. Besides the first goal, the goals two to six of EPA are all directed towards identifying the human capital of an organisation. Because EPA can be used for so many different purposes, it is an important instrument serving as input for many kinds of human resource decisions (Judge & Ferris, 1993). Therefore, it is essential that it is conducted in a valid (Rothstein, 1990) and reliable (Viswesvaran, Ones & Schmidt, 1996) way. Valid and reliable performance assessments will provide a high quality overview of the performance of employees in the entire organisation. This overview can be used in future policy changes, providing an indication of unused knowledge and skills of employees in the organisation.

Reliability and validity will likely vary for each kind of employee performance assessment. Nonetheless, in every form of EPA it is very important to reduce subjective judgement, especially when only one assessor is involved (Berendonk, Stalmeijer & Schuwirth, 2013) which is often the case in EPA. According to Ogunfowora, Bourdage & Lee (2010) there are two categories of assessor bias: errors of omission and errors of commission. Errors of omission occur when assessors omit aspects included in the performance goals when they assess employee performance (Ogunfowora et. al., 2010). Errors of commission occur when assessors include information in the performance assessment which is not specified in the performance goals (Ogunfowora et. al., 2010). Next to bias through assessors, it is also possible the performance goals are not clearly or completely defined (Urdan & Mestas, 2006). This may result in difficulties interpreting these goals by both the assessed employee and the assessing manager (Urdan & Mestas, 2006).

The employee performance assessment usually starts with the setting of performance goals. These performance goals thus function as standards for the desired behaviour of employees. Errors of commission and errors of omission are also related to these standards since they describe assessors adding to or omitting parts of these standards when conducting an EPA. To ensure a valid and reliable (and with that a high quality) performance assessment, it is essential that the standards are clearly and completely defined and that the assessor uses these standards without adding or omitting aspects (Ogunfowora et. al., 2010).

Finally, the word 'assessment' in employee performance assessment indicates that a manager somehow passes judgment on the behaviour of an employee. In the Netherlands, this judgment is most often quantified using a rating scale with categories (van den Broek et. al., 2000). The amount of

categories can differ for each rating scale, numbers between two and ten exist (van den Broek et. al., 2000). One of the causes of problems in employee performance assessment is the inability to thoroughly understand the meaning of rating scales (Campbell, Dunnette, Lawler & Weick, 1970 as mentioned in Borman, 1978). An unclear definition of a rating scale may lead to different interpretations of the scale by different people (for example managers versus employees or managers versus managers), thus causing differences in quality of the employee performance assessment.

Several authors already conducted research in this field. Sulsky & Balzer (1988) found that it is essential to make sure an appropriate standard is used, otherwise, the EPA will be rendered meaningless. Hongguang & Lehong (2011) found that each standard should have clear content in order to make sure it is clearly defined. Brewer & Wilson (2013) found that the greater the precision of the definition of the standards, the more effective EPA becomes. According to Hongguang & Lehong (2011) the standards should match the work environment of an employee as closely as possible. Scullen, Judge & Mount (2003) concluded that it does not matter from which position (manager, employee, etc.) one views standards, the perception remains the same. It is therefore not necessary to create multiple sets of standards of use a different kind of performance appraisal. However, according to Scullen et. al. (2003) both managers and employees prefer a type of EPA that matches their own views and beliefs as close as possible. If this is not the case, the chances of EPA being performed as envisioned by upper management decrease. Regarding rating scales, it is best if a rating scale is described using desired behaviours and if the goal of EPA lies in development of employees (contrary to administrative goals where EPA is used solely for promotion/degradation purposes) (Fay & Latham (1982, in Heneman, Moore & Wexley, 1987). However, raters who are accurate at the behaviour description can still be poor raters (Murphy, 1991). The quality of the assessment depends on the behaviour and the dimensions being rated and behaviour observation and recall is not sufficient for an accurate EPA (Murphy, 1991). Finally, descriptions of a rating scale using adjectives such as satisfactory, very low, low, etc. are subjective because their interpretation can mean different things to different raters (Brewer & Wilson, 2013).

Based on the research mentioned above, the aim of this research can be drafted. Since all studies above are conducted in countries other than the Netherlands, it would be interesting to see how the statements of above authors (for example regarding the perception of standards) hold up within a Dutch organisation. Although Scullen et. al. (2003) mention the perception of standards does not change according to the position one is in, they do not state the same for the perception of rating scale. This study will therefore look into the differences between manager and employee in rating scale perception. Above, it is also described that the best rating scales use desired behaviours. This study focused on whether a rating scale using desired behaviours ensures a good assessment, thereby hoping to confirm the statement presented above in a Dutch organisation. Summarizing, the aim of this research is to explore whether (and if so, what) similarities and differences exist between manager and employee in the perception of standards and interpretation of rating scale for employee performance assessment.

The following chapters will cover the conceptual framework and research question, context description, research approach, results, conclusion and discussion eventually leading to the answering of the research question and providing recommendations for further research and implications for practice.

Exploration and definition of the (research) problem

This chapter covers the exploration and definition of the research problem. To provide the reader with more information concerning relevant terms and topics, the subjects mentioned in the introduction will be explored in more detail thus forming the conceptual framework of this research. The topics employee performance assessment, standards and rating scale are studied in more detail and the interpretation of these standards and perception of the rating scale by manager and employee is introduced. The relation between all these constructs will be summarized in a model. By zooming in more and more on the terms and topics relevant for this research the chapter works its way up to the research question and sub questions presented at the end of this chapter.

Wherry and Bartlett (1982) describe employee performance assessment (EPA) as an attempt of the assessor to report the behaviour of an employee over a set period in time, (usually the past six months or year). The EPA can cover a lot of aspects regarding employee performance, from the general added value of an employee to the organisation to whether he starts on time (Wherry and Bartlett, 1982). The EPA is thus an interaction between two parties: assessor and employee. Evidence suggests that managers are the most reliable source for EPA (Conway et. al., 1997; Viswesvaran et. al., 1996). Therefore, this study will focus on the manager as assessor.

The EPA is often part of a performance management cycle which usually consists of three aspects, either planning-coaching-assessing or target setting – progress reviewing – assessing (van de Broek, van der Giessen & van Oerst - van Dorst, 2000). The process of assessing is predominantly based on the recall of observations on the part of the assessor (Wherry and Bartlett, 1982). The accuracy of this rating depends on three aspects: 1) actual performance of the employee; 2) observation of this performance by the assessor and 3) the assessor's recall of these observations (Wherry and Bartlett, 1982). This process is visualised in Figure 1.

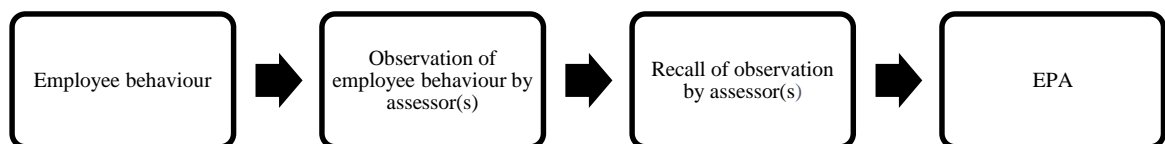


Figure 1. Process of assessing employee behaviour

As mentioned in the introduction, the EPA usually starts with the setting of performance goals. The three types of bias (definition errors, errors of commission and errors of omission) can be translated to defining/interpreting, inclusion and exclusion of standards. Standards for EPA need to be defined simply and clearly in order to avoid different interpretations by different people (van den Broek et. al., 2000). A very clear way to describe these standards, for example, is by using the SMART method (van den Broek et. al., 2000). SMART stands for Specific, Measurable, Acceptable, Realistic and Time (van den Broek et. al., 2000). Besides an unclear definition of EPA standards, it is also possible assessors do not only include the standards described, but include other standards as well; the errors of commission (Ogunfowora et. al., 2010). Finally, assessors can choose to purposefully exclude aspects of the standards that should be included in EPA; the errors of omission (Ogunfowora et. al., 2010). Due to these errors, there is a difference between the standards that should be used in EPA and those that are actually used. The combination of 'official' standards and the interpretation of these standards by both employee and manager lead to the perceived standards used in performance assessment (as depicted in Figure 2).

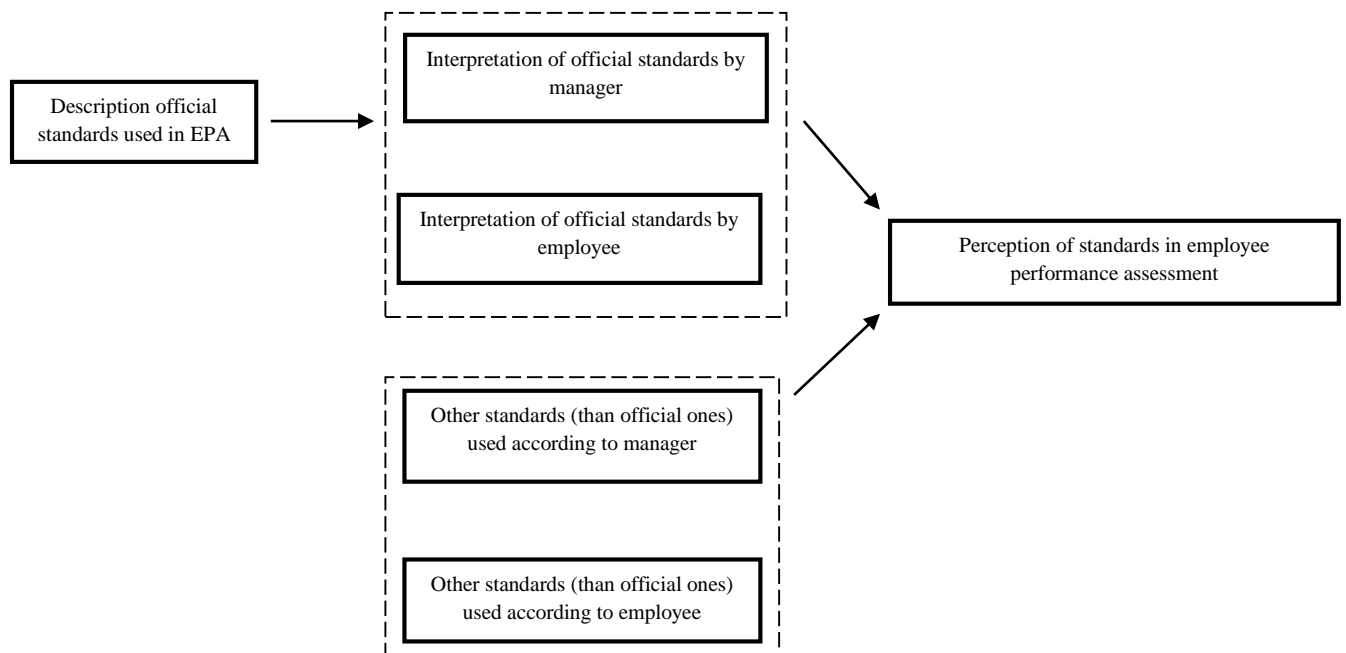


Figure 2. Overview of relations between aspects of standards used in EPA

At the end of EPA, a manager passes judgment on the behaviour of an employee. In the introduction it was stated that an unclear definition of a rating scale may lead to different interpretations of the scale by different people. Since this research focuses on the manager as the assessor, the difference in interpretation will most likely exist between managers and employees (however, differences in manager versus manager or employee vs employee are also possible). The rating scale as it is described is first interpreted by both employee and manager before it is put to use. This is depicted in Figure 3.

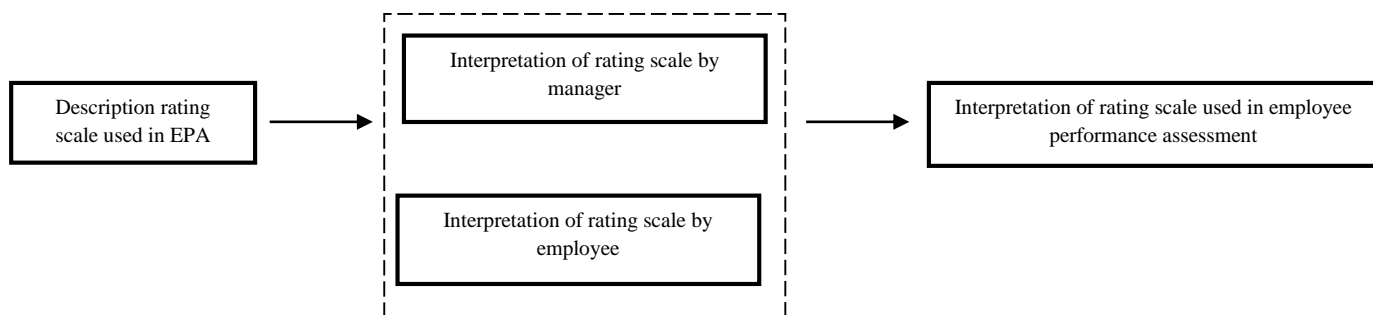


Figure 3. Overview of relations between aspects of a rating scale used in EPA

Combining the information depicted in the text and models above, an overarching model can be created representing all of the influences on EPA focused on in this research. This model is presented in Figure 4.

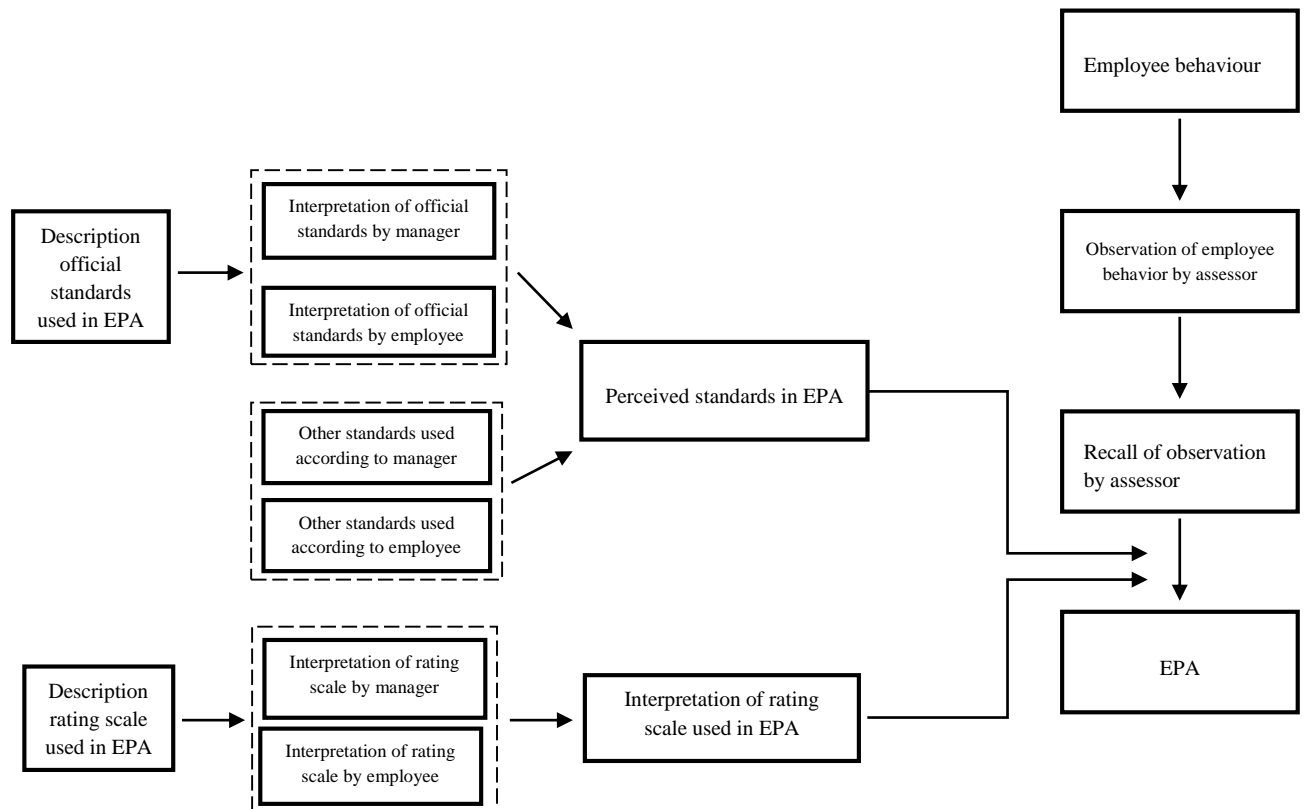


Figure 4. Influences on employee performance assessment (EPA) by manager

The two general aspects of standards and rating scale influence the EPA to great extent. Since the way a manager assesses is largely based on the standards and the rating scale that are applied, the focus of this research will be placed on those two aspects. It is assumed that by providing good quality rating scales and standards, a manager has a clearer frame of reference during observation, helping him to view employee behaviour more accurately. An overview of the variables included in this research is depicted in Figure 5. The grey areas indicate the link to EPA, but will not be directly included in this study.

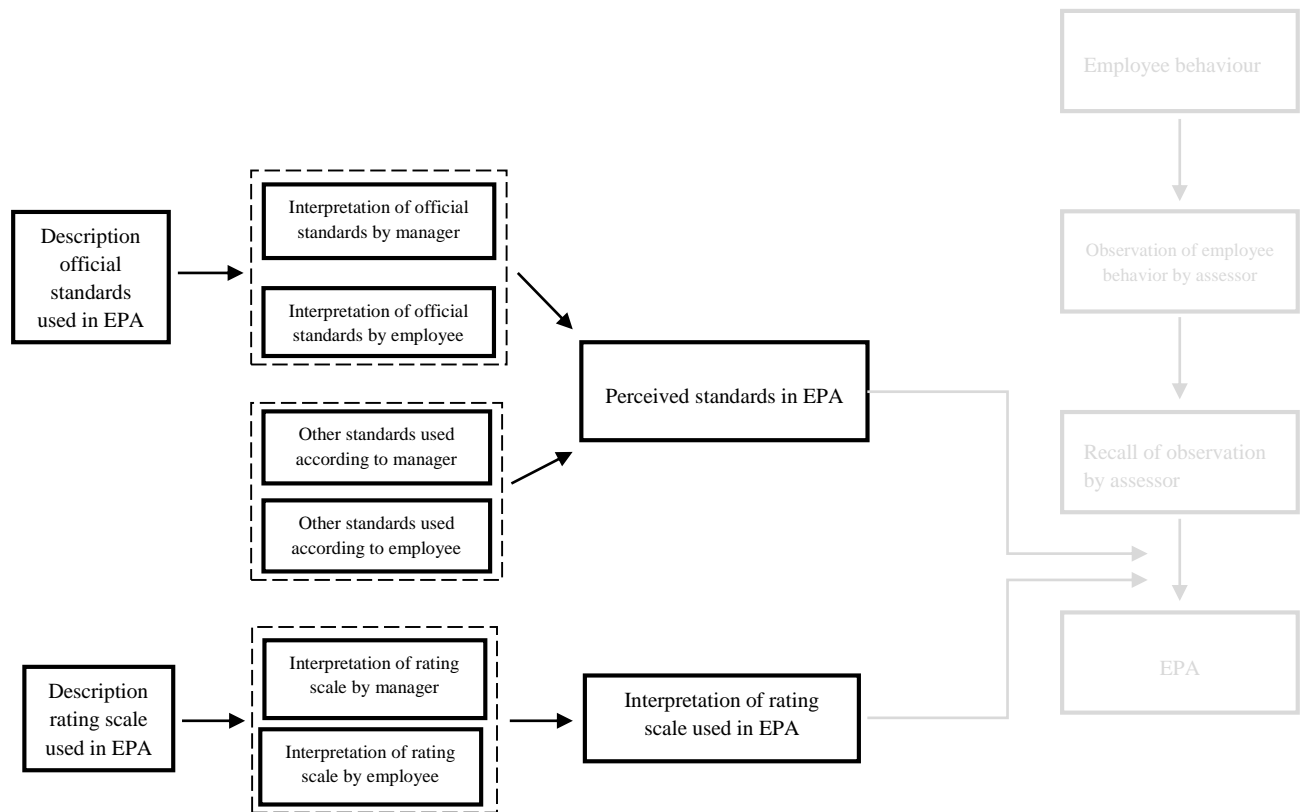


Figure 5. Overview of variables included in this research

The conceptual framework combined with the information presented in the introduction, leads to the following research question:

“What are the similarities and differences between managers and employees in the perception of standards and interpretation of rating scale for employee performance assessment?”

The sub questions are:

- What are the perceived standards for employee performance assessment of both manager and employee?
- How are the different categories of the rating scale interpreted by both manager and employee in terms of expected employee behaviour?

Description of the organisational context

This chapter covers the context of the organisation where this research took place. The context description provides an impression of the organisation in which this research was conducted. It also assists in clarifying why certain methods and instruments were used. Information was gathered in the years 2013-2014 and all organisation related documents were up to date at the time of data collection.

Approximately 681 people work at the municipality of Lelystad. The top of the municipality is formed by the council, which oversees the mayor and aldermen, audit and registry. The mayor and aldermen in turn oversee two directors who manage eight departments: *Beleid* [Policies], *Beheer Openbare Ruimte* [Management of Public Space], *Concernstaf* [Corporate Staff], *Dienstverlening* [Client Services], *Economie en Vastgoedontwikkeling* [Economics and Real Estate Development], *Facilitair, Advies & Beheer* [Facilities, Advice and Management], *Ingenieursbureau en Projectmanagement* [Bureau of Engineers and Project Management] and *Werk, Inkomen & Zorg* [Work, Income & Care]. Each department is led by a department head and consists of multiple teams which in turn are managed by a team leader. A complete overview of the organisational layout is presented in appendix A.

Currently, the municipality faces continuous budget cuts due to the economic crisis, resulting in a vacancy stop. Due to this vacancy stop, the importance of an adequate human resource development (HRD) program becomes more evident. The majority of the municipality employees are fifty years of age and older, meaning a large amount of employees will retire within fifteen years. Due to budget cuts these vacancies have to be filled by employees already working at the municipality. Thus, the recognition of employees' potential for development and talent is crucial in the light of changing job demands.

In order to ensure the vacancies within the municipality can adequately be filled with current employees, it is essential to know the strengths and weaknesses of employees as well as their talents. One way to measure strengths and weaknesses is through the employee performance assessment. The outcome of this assessment provides an overview of what the employee achieved in the last year, where his points for improvement are and hopefully also provide insight in other (relevant) talents. For example, a legal advisor who spends most of his time helping out colleagues with computer issues might also be a good fit for the IT department.

To ensure the EPA provides a good overview of the strengths, weaknesses and talents of an employee the EPA needs to be conducted in such a way that an employee will be assessed the same way throughout the municipality. For example, if one manager would find the employee to perform excellent and another manager finds the same employee insufficient, the EPA would not be a reliable source of input regarding the strengths, weaknesses and talents of an employee. It is thus important for EPA to be conducted as valid and reliable and with as little error as possible.

Currently, the only way in which the behaviour and achievements of employees is documented (for every employee in the municipality) is through the HR-cycle (also known as performance management cycle). The HR-cycle consists of at least three meetings: 1) A planning meeting where the plans and goals for the year are set; 2) at least one follow-up meeting to review progress, and; 3) the employee performance assessment (EPA) determining the extent to which the agreements are kept and grading them on a six category rating scale (excellent, good, sufficient, almost sufficient, insufficient and malfunctioning). *Functiewaarderingen* [job descriptions] provide the standards for the HR cycle. They consist of descriptions of general tasks and responsibilities for each group of functions in the municipality. An example of a job description is presented in appendix B. The job description is the only available document that provides these sort of standards for each employee within the municipality. Besides the job description, departments and teams are allowed to add their own performance

agreements. Since the standards and the rating scale are the two most important pieces of EPA, it is decided to focus on these two throughout the study.

Research approach

In this chapter, the theory of EPA and the reality of the municipality of Lelystad are combined into the research approach. The methods used and types of data collected are presented as well as the methods for data analysis. Following this chapter, the results of the study will be presented.

For data collection and analysis, a qualitative research method was used. In order to measure perception and interpretation as accurately as possible, an interview study was conducted. To limit the span of this research, the focus was placed on three departments of the municipality.

Both sub questions (as presented in earlier) were answered using an interview study. One interview was conducted to answer both sub questions in one meeting. In order to do this, the interview was divided into two sections, a section focusing on the standards and a section focusing on the rating scale. In the analysis of perception of standards, a distinction was made between the standards described in the job description and other standards included in EPA. Regarding the standards described in the job description it was investigated whether interviewees are familiar with the job description, how the job description is currently used, whether there is an emphasis on one or more topics mentioned in the job description and how this document should be used in future EPA's. Regarding the standards used other than those described in the job description it was investigated if there are other standards and if so, which standards. In the analysis of the perception of rating scale, the interpretation of the different categories of the rating scale in terms of expected employee behaviour by both manager and employee was studied. This was assessed by asking each interviewee to read the section 'contact' in the job description and henceforth to provide examples of behaviour that fit each of the categories of the rating scale for the contact standards as described in the job description. The 'contact' section is described as follows:

"Contact is an essential part of the job. The employee has external contacts on a regular basis where they can act independently, even in case of multiple stakeholders. This might occur during contact with both employees and customers but should always happen within the set boundaries of conduct. In many cases, the employee can revert to a form of back-up and / or there will be a possibility of feedback present, in such a way that the employee can choose to consult his superior or others. In case of conflicting interests, it is key to explain rationales and obtain cooperation." Based on this description, interviewees provided the desired employee behaviour for each of the six rating scale categories (excellent, good, sufficient, almost sufficient, insufficient and malfunctioning). An example of an answer would be that employees need to be logged into their desktop phones at all times or that they would have to make sure to behave professional around customers.

The interviews were semi-structured interviews, using a topic list and questions based on literature (Baarda, de Goede & Teunissen, 2009). The main topics and questions were specified and follow-up topics were indicated ahead of the interview. The sequence of the questions was set before the interview but was not set in stone and changed according to the 'flow' of the interview (Baarda et. al., 2009). The interview questions and topics are presented in appendix B, a short summary of the main topics and questions is presented in Table 1.

Main topic	Questions
Current use of standards	In what way is the job description currently used?
	Is this the way it is supposed to be used?
	Is there an emphasis on certain point in the job description during EPA?
Use of other standards	Are there standards used during EPA that are not in the job description?
Use of rating scale	Describe for each of the six rating scale categories what behaviour fits that category regarding the contact section.

Table 1. Overview of main topics and questions for the interview

Guest, Bunce, & Johnson (2006) state that for a research in which the aim is to understand perceptions and experiences among a group of relatively homogenous individuals twelve interviews should suffice. In this study two groups of respondents were identified: managers (team leaders) and employees. To ensure accurate sampling, this research focused on the departments *Werk, Inkomen en Zorg (WIZ)*, [Work, Income and Care], *Dienstverlening (DVL)* [Client services] and *Ingenieursbureau en Projectmanagement (IBP)* [Bureau of Engineers and Project Management]. It was attempted to select a consistent number of employees for each team to ensure equal representations. To ensure that data from this research is comparable, only employees adhering to the “*Vaktechnische* [Technical] job description” were asked to participate in the research. The technical job description was selected based on the amount of employees adhering to this job description (it is the largest job description group in the municipality, thus ensuring a large amount of respondents to select from). Within each participating team, there were at least four employees adhering to the technical job description. The department heads each urged their team leaders and teams to participate in the study and supported the data collection in their departments. They did so due to personal conversations with the researcher of this study which convinced them of the added value of the study.

Both managers and employees were selected using multi-stage purposeful sampling (Onwuegbuzie et. al., 2007). First, criterion sampling was used. Managers were included based on the amount of employees adhering to the technical job description in their team. Within these teams, only employees adhering to the technical job description were asked to participate in the project. Next, convenience sampling (Onwuegbuzie et. al., 2007) was used to select which specific managers and employees were asked to participate in the project. In case any of the employees declined, a new employee was selected from that same team and technical job description. Twenty-five invitations were sent, from which only twelve respondents were interested in taking part in the study. Reasons for not participating ranged from ‘no time’ to ‘not interested’. Eventually, five team leaders were selected and agreed to participate, as were seven employees from the combined teams resulting in a total amount of twelve respondents. For each team leader there was at least one employee belonging to his/her team.

Introducing the interview to the managers and employees, a letter was distributed, covering seven aspects as presented by Seidman (2006): 1) An invitation to participate (specifying in what, how and how long it will take; 2) Risks; 3) Rights; 4) Possible benefits; 5) Confidentiality of the records; 6) Dissemination, analysis of data; and 7) Contact information. Also, an informed consent section was added, enabling the participants to indicate their willingness to take part in the research. After piloting the interview (on two employees involved in the municipality, but not included in the respondents), a total of twelve individual interviews (c.f. Seidman, 2006) were conducted. The interviews were recorded using a voice recorder (with permission of the participants) and were held in one session with a duration of approximately 30-40 minutes depending on the length of the responses from the participants. At the end of the interview, opportunity was provided for questions and participants were informed regarding contact opportunities, member-check and anonymity issues (Baarda et. al., 2009).

The interviews were transcribed verbatim based on the recordings (Baarda et. al., 2009). To ensure anonymity, names were replaced by ‘employee 1, 2’ and ‘manager 1, 2’ etc. during data analysis. Names and numbers were saved in a password protected document. Participants only have access to their own interview data (on request). The interviews then were analysed using open coding / axial coding (e.g. Burnard, 1991). Using the ‘axial coding’ technique mentioned by Baarda et. al. (2009) labels were allocated to relevant aspects of interview transcript. Second, labels were sorted into categories. Following Baarda et. al. (2009), the categories were described to such extent that it is clear for all readers what quotes were ranked into that category.

Results

It was the objective of this study to explore whether (and if so, what) similarities and differences exist between manager and employee in the perception of standards and interpretation of rating scale for employee performance assessment. This chapter covers the results of this research. The topics mentioned in the previous chapter are run through one by one. For each topic, the general findings are reported as well as the differences (or similarities) between managers and employees.

The first sub question focused on the perceived standards for employee performance assessment of both manager and employee. This was covered using several interview questions. The first question asked during the twelve interviews conducted was related to the familiarity of both employees and managers with the job description. Interviewees were asked whether they know what the job description is and if they are familiar with the content of the document. Six of the total of twelve interviewees were familiar with both the existence and content of the job description. However, the six interviewees familiar with both the existence and content of the document existed of four managers and only two employees. Four out of five managers were thus familiar with both the existence and content of the document versus only two out of seven employees. Of the employees, four out of seven were familiar with the existence of the document, but were unaware of the content. Table 2 provides a summary of these results.

	Manager	Employee
Familiar	4	2
Unfamiliar	1	1
Only familiar with existence of document	0	4

Table 2. Familiarity of managers and employees with the job description

The second question posed during the interview concerned the current use of the job description in employee performance assessment (EPA). Interviewees were asked how the job description is currently used during EPA. The categories, descriptions and labels are presented in Table 3. Table 3 indicates four different categories of answers provided by interviewees when asked about the current use of the job description. Table 4 provides an overview of the amount of managers and employees that provided answers fitting each category. Note that is it possible for an interviewee to mention more than one current use of the job description.

Category	Description	Label
Side issue	All quotes adhering to this category describe the use of the job description as a side issue, having little to none direct relevance during EPA	Not actively, only passively / to refer to Side issue
No use	All quotes adhering to this category describe the use of the job description as none or irrelevant.	Unknown / None
Point-by-point walkthrough	All quotes adhering to this category describe the use of the job description as an overview of points that are walked through one-by-one during EPA.	Point-by-point walkthrough
Basis	All quotes adhering to this category describe the use of the job description as a basis, starting point or input for the aspects covered during EPA.	Theoretic basis General starting point Basis for performance agreements Avoiding discussion Input for competencies Assumed to be known Basis for expectations

Table 3. Full list of categories, descriptions and labels regarding current use of the job description

As seen in Table 4, all managers indicate they use the job description as a basis for EPA. They do not talk about or refer to the document itself during EPA, but rather use the content of the document as starting point for the conversation topics. Employee quotes are divided across the categories.

	Manager	Employee
Basis	5	3
Side issue	2	1
No use	1	2
Point-by-point walkthrough	0	3

Table 4. Current use of the job description by managers and employees

Subsequent to the second interview question, the third question covered the content of the job description and the extent to which it is currently used during EPA. Managers and employees were asked to read through the job description and point out whether some topics were more important during EPA (as currently conducted) than others. Table 5 provides an overview of the categories, descriptions and labels found for the answers to this interview question.

Category	Description	Label
Interaction with others	All quotes adhering to this category place the emphasis of the EPA on interpersonal relations and working in teams	Cooperation with colleagues
		Respect towards others
		Able to have decent conversations
		Spot within the team
		Respect qualities of others
		Accepting differences in others
Competencies	All quotes adhering to this category place the emphasis of the EPA on competencies named in the job description or name function content as most important	Not passing judgment on colleagues
		Proactivity
		Initiative
		Updating theoretical knowledge
		Job content
		Creativity
Perception of the manager	All quotes adhering to this category say that the emphasis is placed where needed based on the view of the employees' current performance by the manager	Independence
		Contact
		Perception of employee
No emphasis	All quotes adhering to this category say that there is no specific emphasis during EPA	Perception of manager
		No emphasis

Table 5. Full list of categories, descriptions and labels for the current use of the job description

Table 6 depicts the distribution for managers and employees among the categories. Here also, the numbers represent the amount of interviewees that provided answers fitting each category. Again, it is possible for an interviewee to mention more than one emphasis on topics in the job description. Interaction with others was named by three managers but not by employees. Competencies and the perception of the manager were mentioned by an equal number of managers and employees whereas three employees indicated they did not feel there was a specific emphasis on a topic in the job description.

	Manager	Employee
Interaction with others	3	0
Competencies	3	3
Perception of the manager	1	1
No emphasis	0	3

Table 6. Emphasis on topics in the job description by managers and employees

The final question of the first section of the interview covered the future use of the job description as envisioned by managers and employees. Interviewees were asked if and how the job description should be used during future EPA meetings. Table 7 depicts the categories, description and labels regarding this question.

Category	Description	Label
Does not meet present demands	All quotes adhering to this category indicate the job description is no longer relevant and should not have a place within EPA	Does not meet present demands
Starting point for EPA	All quotes adhering to this category indicate the job description is a good point to start from when conducting EPA but should not be used literally	Job description To refer back to Starting point for EPA
Facilitate decisions regarding promotion and degradation	All quotes adhering to this category indicate the job description is only relevant to facilitate decisions relating to promotion/ degradation issues	Facilitating flexibility in appointment Facilitating decisions regarding promotion

Table 7. Full list of categories, descriptions and labels regarding future use of the job description

Table 8 depicts the distribution for managers and employees among the categories. It is noticeable that five out of seven employees and three out of five managers feel the job description should still be used in the future. The document should be used as a starting point for EPA, a basis to ensure everyone knows what is expected of employees. Two employees feel the document does not meet the demands of the present (and therefore future) time and therefore should not be used at all. Two of the five managers feel the main use of the document should be to facilitate decisions regarding promotion and degradation. Employees should be evaluated based on the content of the job description to review whether they meet the standards for the position they aim to fill.

	Manager	Employee
Does not meet present demands	0	2
Starting point for EPA	3	5
Facilitate decisions regarding promotion and degradation	2	0

Table 8. Future use of the job description as envisioned by managers and employees

To investigate whether the criteria mentioned in the job description are the only ones applied during EPA a question regarding the use of other criteria was posed. Interviewees were asked if other standards than those specified in the job description are used during EPA. From the twelve interviewees, seven felt that during EPA other criteria besides those mentioned in the job description were used. Noticeable was that the five interviewees that felt there were no other criteria were all employees, meaning that all managers interviewed use other criteria in EPA besides those provided in the job description.

Besides investigating whether extra criteria are used, the interviews also covered the kind of extra criteria used in EPA according to managers and employees. Since the labels were relatively hard to place in categories on this question, the decision was made to not create categories but analyse the labels instead. Table 9 provides a description for each of the labels.

Label	Description	Manager	Employee
Performance agreements	All quotes adhering to this label indicate that performance agreements are used as extra criteria during EPA	4	2
Competencies	All quotes adhering to this label indicate that competencies are used as extra criteria during EPA	4	0
Individual differences	All quotes adhering to this label indicate that individual differences are used as extra criteria during EPA	1	0
Behaviour and attitude	All quotes adhering to this label indicate that behaviour and attitude are used as extra criteria during EPA	0	1
Chores and projects	All quotes adhering to this label indicate that extra projects and jobs are used as extra criteria during EPA	1	0
Attendance	All quotes adhering to this label indicate that attendance is used as extra criteria during EPA	1	0
Social skills	All quotes adhering to this label indicate that social skills are used as extra criteria during EPA	3	1
Motivation	All quotes adhering to this label indicate that motivation is used as extra criteria during EPA	1	0
Job content	All quotes adhering to this label indicate that being up-to-speed on relevant knowledge is used as extra criteria during EPA	1	0

Table 9. Full list of label description of the interview question regarding extra use of criteria

Performance agreements appear to be the most used extra criterion during EPA (four managers and two employees). These performance agreements are drawn up during the first meeting (planning) of the EPA cycle, then evaluated during the second (functioning) and third (assessment) meeting. Competencies (other than those mentioned in the job description) are also an important extra criterion for managers during EPA (four managers named these). Managers mostly use competencies to help them name the skills they would like to see from their employees. These competencies are also used to draw up and clarify performance agreements so it is not surprising that both are named relatively often. However, none of the employees mentioned competencies when it came to the use of extra criteria. Finally, social skills are named by three managers as an extra criterion in EPA. The fact that there are only four answers from employees regarding this question is related to the previous statement that five out of seven employees do not feel extra criteria are used during EPA. Another noticeable result is that the same manager accounts for all of the answers in the categories individual differences, projects, attendance and motivation.

The second and final research sub question focused on the interpretation by both manager and employee of the different categories of the rating scale in terms of expected employee behaviour. Interviewees provided the desired employee behaviour for each of the six rating scale categories (excellent, good, sufficient, almost sufficient, insufficient and malfunctioning). One manager was not able to answer this question, so the data will consist of four managers and seven employees. Six other interviewees were unable to provide a description for all six rating scale categories, so instead they provided answers for the ones they could come up with. These answers were summarized into keywords (labels), of which an overview is presented in appendix D. Appendix D also shows which interviewees were unable to provide an answer for which category. Analysis was conducted based on the keywords and nine categories of answers were found. The keywords and their description are presented in Table 10.

Description		Managers		Employees	
		Quotes	People	Quotes	People
Happy customer	Quotes related to customers leaving feeling satisfied	5	2	6	4
Communication	Quotes related to the communication skills of employees	9	2	6	3
Listening	Quotes related to the extent to which employees listen to customers	2	1	3	2
Integrity	Quotes related to the integrity of the employees	3	1	2	1
Independence	Quotes related to the extent to which employees do their job independently	8	3	1	1
Interest in others	Quotes related to the extent to which employees express interest in others	0	0	8	5
Knowledge	Quotes related to the extent to which employees keep their knowledge updated	1	1	4	2
Availability	Quotes related to the extent to which employees are available for customers	0	0	3	1
Exceeding expectations	Quotes related to the extent to which employees' performance exceeds expectations	2	2	3	3

Table 10. Keywords, description relating to the interpretation of the rating scale. Also presenting the amount of quotes and people belonging to each category

Table 10 also represents the amount of quotes that fall into each category, separated for managers and employees. Finally, this table presents the amount of managers or employees that provided the number of quotes mentioned directly to the left. There are several notable results that can be derived from this table. First, two categories are named often by both managers and employees; 'happy customer' (five quotes by two managers and six quotes by four employees) and 'communication' (nine quotes by two managers and six quotes by three employees). Second, the category 'independence' is named eight times by three managers, but only once by an employee. This possibly indicates a difference in focus between managers and employees. The reversed situation can be seen for the 'interest in others' category which was named eight times by five employees, but not once by managers. Finally, although

the category 'exceeding expectations' was named by two managers and three employees, this category was only used when it came to the keywords in the 'excellent' category of the rating scale.

Conclusion and discussion

This chapter answers the research questions of this study. The main results will be summarized and the topics for discussion will be introduced. Recommendations and implications for practice are listed in the next chapter.

The first sub question focused on the perception of standards for EPA by both manager and employee. This was assessed through multiple interview questions. The most notable conclusion from the first interview question is the difference between employees and managers in their familiarity with the job description. Four out of five managers (performing the employee performance assessment) are familiar with both existence and content of the document, but this goes for only two out of seven employees. Four out of seven employees were familiar with the existence of the document, but unaware of the content and thus the official requirements stated for their job and EPA. The fact that so many employees are not aware of the content of the job description and almost all managers are, is a conclusion that possibly has multiple causes. First, the job description is construed in 2008 making it between five and six years old at the time of data collection. If this document is not actively communicated, employees that were not employed in 2008 might not be aware of its existence. Secondly, since managers tend to only use the job description as a basis without naming or actively using the document, it is quite possible employees will not come into contact with it, even during EPA.

Zooming in on the aspects mentioned in the job description, three out of five managers (but no employees) look at interpersonal relations during EPA. This may have to do with the overarching position of a manager versus an employee. A manager tends to look more at a team as a whole (and therefore, interpersonal relations are important) where employees might focus more on their place in the team or their own performance.

Apart from the standards mentioned in the job description, all managers indicate the use of other standards where five out of seven employees indicate no other standards are used. Most often extra standards are used in the form of performance agreements (four managers and two employees) or competencies not named in the job description (four managers but no employees). Most noticeable here is that all managers indicated the use of extra standards during EPA. These standards are not specified in the job description and are most often not communicated otherwise. These are the errors of commission (Ogunfowora et. al., 2010) mentioned in the exploration and definition of the research problem. Managers purposefully use standards other than those mentioned in the job description. This might in part have to do with the fact that the document was created in 2008 and that it does not contain the most recent job demands for most jobs. Also, due to the ambiguous description of the standards in the job description, it is possible managers interpreted the meaning of the standards differently or decided to add their own (more specific) standards. The use of performance agreements and (in case of managers) competencies can be explained by the need to set some kind of specific goal for each employee which cannot be described in the general job description. These specific goals may consist of competencies but may also be more personal. Although this explains why both managers and employees indicate performance agreements as an extra criterion, it does not explain why competencies are not. However, during the interviews and the presence of the researcher in the municipality of Lelystad, it became apparent that employees tend not to look at themselves in terms of competencies but rather describe themselves and their goals in behaviour or personality traits.

Summarizing the conclusions mentioned above, the first sub question can be answered. It appears that managers and employees have a rather different perception of standards for EPA. Whereas managers generally are familiar with the job description, employees are not. Where managers focus on interpersonal relations, employees do not, thus being unaware of a part of the standards used by managers. Finally, managers use standards other than those described in the job description where five out of seven employees do not think extra standards are used. Employees thus have a very limited

knowledge of the standards that are applied by managers during EPA and there is a gap between the perception of managers and employees where it comes to standards for EPA.

The second sub question related to the interpretation of different categories of the rating scale by both manager and employee in terms of expected employee behaviour. This turned out to be a tough question to answer during the interviews. One manager was not able to answer this question as a whole and six other interviewees were unable to provide a description for all six rating scale categories. This mostly had to do with the presence of six rating scale categories. Interviewees for example struggled with the difference between almost sufficient / insufficient or sufficient / good. Most interviewees also indicated the use of six categories was no necessity for them, they would prefer three or four. Despite these difficulties, all but one interviewee provided at least three descriptions for the rating scale categories and thereby provided a description of behaviour for all the categories they deemed essential.

Two topics are named often by both managers and employees; 'happy customer' (five quotes by two managers and six quotes by four employees) and 'communication' (nine quotes by two managers and six quotes by three employees). Since interviewees were asked to provide descriptions based on the contact section of the job description, communication is not odd to find in both managers and employees. Although no immediate explanation for the presence of 'happy customer' was found, it is possible the municipality as a whole focuses on this topic a lot, thus influencing the thought pattern of interviewees. 'Independence' is named eight times by three managers, but only once by an employee. The reversed can be seen for the 'interest in others' topic which was named eight times by five employees, but not once by managers. It is possible managers focus more on independence because they mostly feel the effects when employees cannot work independently. Managers will be the ones to answer questions, provide back-up in conversations and are generally most often interrupted in case of a lack of independence in employees. Since colleagues possibly experience this less, it is possible they do not emphasize this as much when it comes to EPA. The 'interest in others' category that was found in employees but not in managers can be explained by Conway (1999, in Scullen, Judge & Mount, 2003). Colleagues have a different relationship with other employees than managers. Managers can always revert back to some form of power or authority in order to get what they want from an employee but colleagues are dependent on interpersonal relationships (Conway, 1999 in Scullen, Judge & Mount, 2003). This is why teamwork and interpersonal relations would receive more focus from employees than from managers (Conway, 1999 in Scullen, Judge & Mount, 2003).

The information presented above leads to the answer to the second sub question of this study. This question focused on the interpretation of the different categories of the rating scale by both manager and employee in terms of expected employee behaviour. There appears to be a difference in not only the interpretation between manager and employee (for example in the topics independence and interest in others), but also a difference between individuals. This can be seen in the seven interviewees who were unable to answer this question completely and indicated not to use all six of the rating scale categories. Despite the differences, there are also some similarities between managers and employees when it comes to the interpretation of the rating scale. This is shown in the topics happy customer and communication which were both named by multiple managers and employees.

Besides the answers to the two sub questions of this study, there are some general points of discussion. In this study, the decision was made to focus on employees belonging to the 'Technical job description' only. As it turns out, even within the 'Technical' group, there still are a lot of different positions ranging from for example administrative positions to customer care to legal advice. It is possible this wide range of positions has had an influence on the results of this study, mostly due to the fact that it is possible that part of the differences found are explainable by the difference in positions and thereby the views on the municipality and EPA. However, this grouping into the 'Technical' group is how the municipality envisioned the job description to work so this approach does seem to have been an adequate way to test that.

Another general remark is related to the importance of EPA within the municipality as a whole. During several non-recorded conversations with municipality employees, it became apparent that most see the EPA (and HR-cycle as a whole) as a mandatory aspect of their work. The added value of receiving feedback, target setting and assessment is not clear to all and certainly not experienced. This also showed during the recruitment of participants for this study. Even with the support of department heads, only twelve out of twenty-five people responded positively when asked to participate in the study. The declining answers ranged from 'no time' (where time investment was supported by department heads) to 'not interested', thus supporting the hypothesis that EPA is no more than a mandatory process. This may be strengthened by the fact that there are relatively no consequences (as indicated during non-recorded conversations) to a negative EPA outcome. Should an employee be evaluated negatively, a plan for improvement is set up. This is as far as the consequences go, there is no pay cut, no (immediate) degradation risk or likewise consequence. This possibly does not encourage the commitment of both managers and employees to participating in a meaningful employee performance assessment.

Recommendations and implications for practice

This chapter covers the recommendations and implications for practice. First, several recommendations for further research will be made.

During this study, it became apparent that the work environment and organisational culture might be of influence on the way EPA is carried out in an organisation. The link between EPA and organisational culture is supported by Hofstetter and Harpaz (2015) who found that a performance assessment cannot occur if it is not supported by the norms and values of an organisation. The fact that (at least some) managers and employees view the EPA as a process with little added value; that the stakes, gains and interests are not always clear to participants and; that there are virtually no consequences to a negative outcome of EPA are all aspects found in this study that determine the organisational culture of the municipality. This would be a good starting point for all future research regarding EPA, to look into the relation between organisational culture and the employee performance assessment, specifically focusing on the relation between experienced added value of EPA and the way EPA is conducted and the relation between the consequences to a negative EPA outcome and the way EPA is conducted and followed up in the months after.

It would also be interesting to research which organisational culture is successful in supporting EPA and which organisational culture works counter supportive. Cravens, Oliver, Oishi and Stewart (2015) found that a positive organisational culture (where employees feel they are being valued by the organisation) is key to achieve the goals of EPA. The findings of this study can provide a starting point for research regarding counterproductive organisational cultures and possible solutions.

As a final recommendation it is important to keep in mind that accuracy of a rating scale is not the only aspect to focus on in EPA. Sulsky and Balzer (1988, in Murphy (1991)) mentioned that aspects like inter-rater agreement or ease of use might be more important than accuracy. It would be good to study the relationship between accuracy and the topics mentioned by Sulsky and Balzer.

Next, some implications for practice will be covered. As mentioned in the exploration and definition of the research problem, standards for EPA need to be defined simply and clearly in order to avoid different interpretations by different people (van den Broek et. al., 2000). The example of these standards provided in the research approach regarding contact shows that a clear and unambiguous description is not the case. The sentences are long, words are open to interpretation and it is not specified when an employee meets the criteria for a good assessment. The use of SMART standards might bring some solace here. Rubin (2002) notes that SMART goals assist in increasing employee motivation and improve the odds that employees will meet the goals set for and with them. It is however recommended to do further research in order to investigate which method of description will be the most effective.

Each organisation can take from this research to closely look into their standards and description of rating scale. Are they clearly defined and do both managers and employees agree on the definition and meaning? If not, the SMART method mentioned above may prove useful. Also it should be clear in advance which standards are used during EPA and which are excluded from assessment. Creating a culture supportive of EPA is recommended here since a culture where all levels of managers are involved directly in creating and maintaining the EPA decreases the risk of errors of commission (Hofstetter and Harpaz, 2015). Here lies a responsibility with the HR department, the managers, and the employees themselves.

For larger organisations, it is recommended to consider abandoning the single-assessor EPA or to ensure that managers are consulting with either an EPA expert or each other on a regular basis to ensure quality and generalisability across the organisation regarding the rating scales. It is perhaps a possibility to explore other methods of EPA than those performed by just one manager, for example 360° feedback. In 360° feedback, not only managers assess employees, but employees also assess themselves and peers. In order to do this, they need to be acquainted with the skills and behaviours (thus

standards) that are important to the organisation (Garavan, Morley and Flynn, 1997). This method of assessment might not only help to create a more complete and honest picture of the employees' performance (by eliminating the single assessor EPA) but it will also ensure all employees are familiar with the standards that are expected of them.

Finally, a comment regarding generalisability; this study focused on a sample of twelve interviewees, five managers and seven employees. Since the municipality as a whole has approximately 681 employees, the sample size was only 1,76%. This is not enough to generalise the findings for the entire municipality. Also, this study focused on the technical job description only. This is one of the job descriptions with mostly lower- to medium educated jobs. It is possible that other job descriptions (with other educational backgrounds) provide different results. This too is a reason not to generalise the findings for the entire municipality. Since the results are not generalizable for the municipality as a whole, it is not possible to indicate whether results can be generalized to other organisations.

Even though the results are not generalizable throughout the municipality, there is still some things to take into account regarding the daily practice of the municipality. In order to improve the engagement of managers in the EPA process, it is essential that all managers are involved in the creation of the new version of the EPA within the municipality. This can be done by using focus groups, interviews or general meetings. Through this, all managers will become involved and have the opportunity to be heard thus improving their engagement and the value they attach to the EPA process. When managers start to see the EPA process as an added value, they will convey this to their employees. In order to make sure the employees are heard as well, it is advised to conduct a focus group with employees when the new version of EPA reaches the testing phase. This group will then be able to transfer their findings and (assumable) enthusiasm to other employees.

For the P&O department, a point to review for the new EPA version is the amount of categories on the rating scale since most of the interviewees were unable to provide a detailed description for all categories. Also, it is essential to look into the definition of the rating scales: when is a category awarded to an employee, what behaviour or values are attached to that category? This links directly to the job description document which should possibly be redefined, making it so that it is clear for all employees what is expected of them and which behaviour will result in which rating. The complete content of the job description as a standard for EPA might also be reviewed. All managers indicated they use standards besides the one mentioned in this document. A focus group with managers will help to determine what managers truly deem important in their employees and these topics might be considered to be added to the standards for EPA.

Reference list

- Baarda, D.B., De Goede, M.P.M. & Teunissen, J. (2009). *Basisboek kwalitatief onderzoek*. Groningen, the Netherlands: Noordhoff Uitgevers
- Berendonk, C., Stalmeijer, R.E. & Schuwirth, L.W.T. (2013). Expertise in performance assessment: assessor's perspectives. *Advancements in Health Science Education*, 18, 559-571. doi: 10.1007/s10459-012-9392-x
- Brewer, N., & Wilson, C. (2013). *Psychology and policing, Chapter 11. Performance Appraisal*. Psychology Press
- Broek van den L., Giessen van der, R., & Oers - van Dorst van, A. (2000). *Performance management*. Kluwer
- Burnard, P. (1991). A method of analyzing interview transcripts in qualitative research. *Nurse Education Today*, 11, 461-466.
- Campbell, J. P., Dunnette, M. D., Lawler, E. E., & Weick, K. W. (1970). *Managerial behaviour, performance and effectiveness*. New York: McGraw-Hill.
- Conway, J. M., & Huffcutt, A. I. (1997). Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance*, 10, 331-360
- Cravens, K. S., Oliver, E. G., Oishi, S., & Stewart, J. S. (2015). Workplace Culture Mediates Performance Appraisal Effectiveness and Employee Outcomes: A Study in a Retail Setting. *Journal of Management Accounting Research*, 27(2), 1-34
- Denison, E.F. (1962). *The sources of Economic Growth in the United States and the Alternatives Before Us*. New York: Committee for Economic Development. In Nafuhko, F.M., Hairston, N., Brooks, K. (2004). Human capital theory: implications for human resource development. *Human Resource Development International*, 7(4), 545-551. doi: 10.1080/1367886042000299843
- Evers, G.H.M., van Laanen, Ch.C.M., & Sipkens, G.J.J. (1993). *Effectieve personeelsplanning: een praktische kijk op personeelsstromen*. Kluwer Bedrijfswetenschappen, Deventer
- Garavan, T. N., Morley, M., & Flynn, M. (1997). 360 degree feedback: its role in employee development. *Journal of Management Development*, 16(2), 134-147
- Gemeente Lelystad: Concernplan (2014). Retrieved June 10, 2014, internal document
- Gemeente Lelystad: Geschiedenis (2014). Retrieved June 10, 2014, from <http://www.lelystad.nl/Gemeente>
- Gemeente Lelystad: Geschiedenis (2014). Retrieved June 10, 2014, from <http://www.lelystad.nl/4/Toekomst/ToekomstHistorie-Geschiedenis.html>
- Gerring, J. (2004). What is a case study and what is it good for? *American Political Review*, 98(2), 341-354

- Guest, G., Bunce, A., & Johnson, L. (2006) How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59-82. doi: 10.1177/1525822X05279903
- Heneman, R. L., Moore, M. L., & Wexley, K. N. (1987). Performance-rating accuracy: A critical review. *Journal of Business Research*, 15(5), 431-448
- Hofstetter, H., & Harpaz, I. (2015). Declared versus actual organizational culture as indicated by an organization's performance appraisal. *The International Journal of Human Resource Management*, 26(4), 445-466
- Hongguang, D., & Lehong, G. (2011, August). Analysis & Optimization on Employee Performance Assessment Indicator System in Xuzhou Coal Mining Group. In *Management and Service Science (MASS), 2011 International Conference on* (pp. 1-6). IEEE
- Jacobs, R., Karfy, D. & Zedeck, S. (1980) Expectations of behaviorally anchored rating scales. *Personnel Psychology*, 33, 595-640
- Judge, T. A., & Ferris, G. R. (1993). Social context of performance evaluation decisions. *Academy of Management Journal*, 36, 80-105.
- MacLeod, L. (2012). Making SMART goals Smarter. *Physician Executive Journal*, March-April
- Mincer, J. (1962). On-the-job training: costs, returns and some implications. *Journal of Political Economy*, 70(5), 50-79. In Nafuhko, F.M., Hairston, N., Brooks, K. (2004). Human capital theory: implications for human resource development. *Human Resource Development International*, 7(4), 545-551. doi: 10.1080/1367886042000299843
- Murphy, K. R. (1991). Criterion issues in performance appraisal research: Behavioral accuracy versus classification accuracy. *Organizational behavior and human decision processes*, 50(1), 45-50
- Nafuhko, F.M., Hairston, N., Brooks, K. (2004). Human capital theory: implications for human resource development. *Human Resource Development International*, 7(4), 545-551. doi: 10.1080/1367886042000299843
- Psacharopoulos, G. (1985). Returns to education: a further international update and implications. *The Journal of Human Resources*, 10(4), 583-604. doi: 0022-166X/85/0004-0583
- Ogunfowora, B., Bourdage, J. & Lee, K. (2010) Rater personality and performance dimension weighting in making overall performance judgments. *Journal of Business Psychology*, 25, 465-476. doi: 10.1007/s10869-009-9144-y
- Onwuegbuzie, A.J., Leech, N.L. (2007). A call for qualitative power analyses. *Quality and Quantity*, 41, 105-121. doi: 10.1007/s11135-005-1098-1
- Rothstein, H.R. (1990) Interrater reliability of job performance ratings: growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology*, 75(3), 322-327
- Rubin, R. S. (2002). Will the real SMART goals please stand up. *The Industrial-Organizational*

- Scullen, S. E., Mount, M. K., & Judge, T. A. (2003). Evidence of the construct validity of developmental ratings of managerial performance. *Journal of Applied Psychology*, 88(1), 50
- Seidman, I. (2006) *Interviewing as qualitative research. A guide for researchers in education and the social sciences*. Teachers College Press, New York.
- Smith, E. R., Mackie, D. M., & Claypool, H. M. (2007). *Social psychology*. Psychology Press.
- Sulsky, L. M., & Balzer, W. K. (1988). Meaning and measurement of performance rating accuracy: Some methodological and theoretical concerns. *Journal of Applied Psychology*, 73(3), 497
- Taken van een gemeente (2014). Retrieved June 10, 2014, from <http://www.rijksoverheid.nl/onderwerpen/gemeenten/taken-gemeente>
- Viswesvaran, C., Ones, D.S., & Schmidt, F.L. (1996) Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 1996, 81(5), 557-574
- Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies*. Upper Saddle River, NJ: Prentice Hall

Appendices

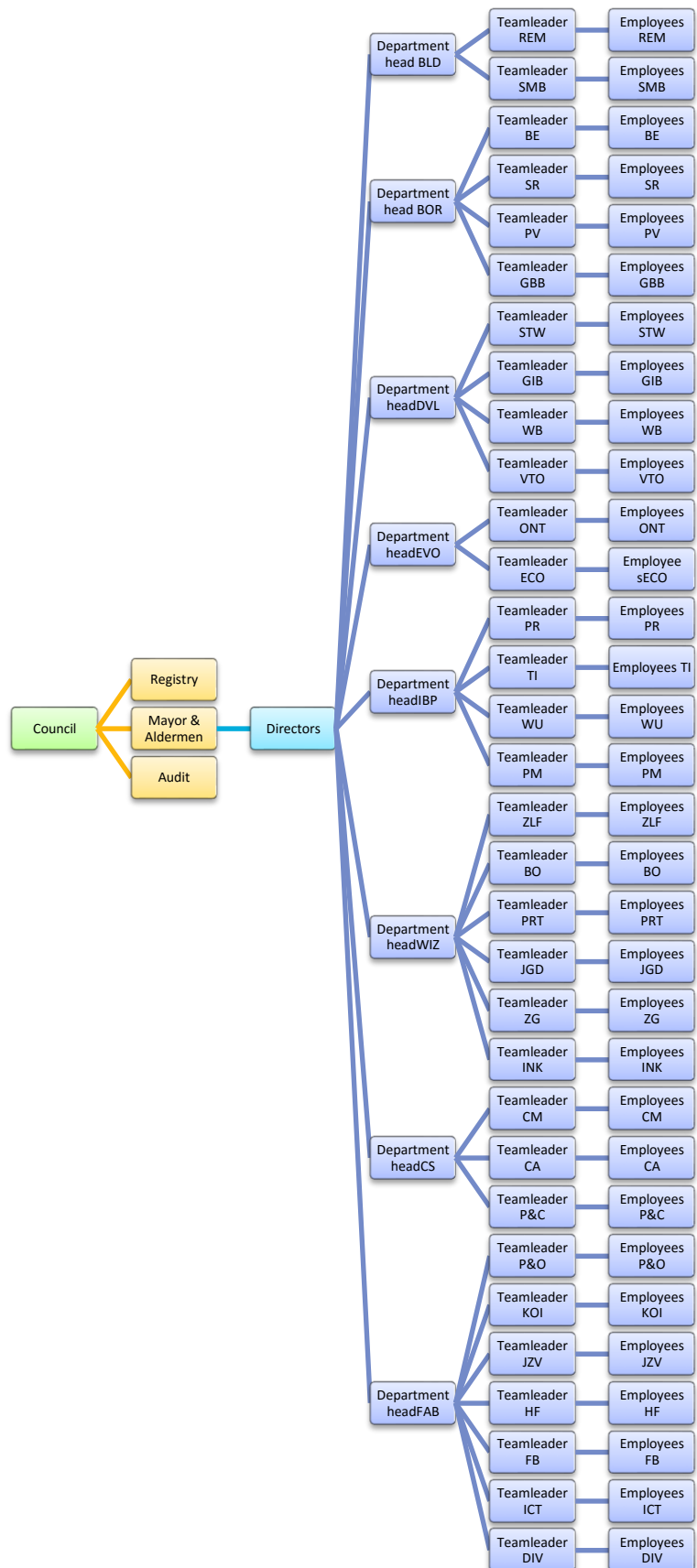
Appendix A: Overview of the organisation

Departments

BLD	Beleid
BOR	Beheer Openbare Ruimte
CS	Concernstaf
DLV	Dienstverlening
EVO	Economie en Vastgoedontwikkeling
FAB	Facilitair, Advies en Beheer
IBP	Ingenieursbureau en Projectmanagement
WIZ	Werk, Inkomen & Zorg

Teams

BE	Beheer
BO	Bedrijfsondersteuning
CA	Concernadvies
CM	Communicatie
DIV	Documentaire Informatie Voorziening
ECO	Economie
FB	Financieel Beheer
GBB	Gebouwenbeheer
GIB	Geo Informatie en Belastingen
HF	Huisvesting en Facilitair
ICT	Informatie, Communicatie, Technologie
INK	Inkomen
JGD	Jeugd
JZV	Juridische Zaken en Veiligheid
KOI	Kwaliteit, Onderzoek en Informatiemanagement
ONT	Ontwikkeling
P&C	Planning & Control
PM	Projectmanagement
P&O	Personeel & Organisatie
PR	Projecten
PRT	Poort
PV	Proces en Voorbereiding
REB	Ruimtelijk Economisch Beleid
SMB	Sociaal Maatschappelijk Beleid
SR	Service
STW	Stadswinkel
TI	Technische Installaties
VTO	Toezicht, Veiligheid en Ondersteuning
WB	WABO en Bestemmingsplannen
WU	Werkvoorbereiding en Uitvoering
ZG	Zorg
ZLF	Zelfstandigen Lokaal Flevoland



Appendix B: Job description

Motiveringsrapport Functiewaardering

Pagina 1 van 1

Organisatie:	Gemeente Lelystad
Naam functie:	Vaktechnisch B
Functiecode:	22
Organisatie-onderdeel:	Vaktechnisch
Datum waardering:	18-12-2008
Datum beschrijving:	18-12-2008

Hoofdgroep:

III

Het gaat in de functie om werkzaamheden waarbij een breder inzicht in de samenhang is vereist. De werkzaamheden als zodanig zijn bepaald tot een afgebakend werkgebied, maar vereisen eigen analyse, interpretatie, conceptie en aanpak c.q. optreden en gedragbepaling. De functie voldoet aan de omschrijving van hoofdgroep III. Uitgegaan is van een opleiding op MBO-niveau 4. Hoofdgroep III impliceert twee jaar ervaring om de theoretisch opgedane kennis in de praktijk te leren toepassen.

Functionele Vorming

Score: 4

Aanvullend is nodig: 4 à 5 jaar opleiding en/of ervaring, onder meer in verband met het verrichten van de complexere vakinhoudelijke werkzaamheden op basis van zelfstandige vakuitoefening en werkkuitvoering variërend van meervoudig tot complex van aard, waarvoor ruime aanvullende vaktechnische kennis nodig is, op één of meerdere vakgebieden.

Handelingsvrijheid

Score: 3

De functie vereist een zelfstandig optreden. Gelet op de eisen die de organisatie stelt aan de functie (plaats in de organisatie) is in het algemeen evenwel geen sprake van een feitelijke inhoudelijke controle van de (niveaubepalende) werkzaamheden. Gegeven de aard van de werkzaamheden mag worden verwacht dat het werk/resultaat of de werksituatie slechts aan een eindcontrole/-toetsing/-beoordeling kan worden onderworpen.

Keuzemogelijkheden

Score: 3

In de functie is creativiteit en initiatief nodig bij het oplossen van zich voordoende praktische problemen. Het betreft veelal eenmalige oplossingen.

Leidinggeven

Score: 0

Geeft geen leiding.

Contact

Score: 3

Contact is een essentieel onderdeel van de functie. Betrokkene heeft regelmatig externe contacten, waarbij zelfstandig wordt opgetreden, ook wanneer er sprake is van tegengestelde belangen, welke kunnen voorkomen in de contacten met medewerkers en klanten, dit evenwel binnen vastgestelde marges. Veelal kan worden teruggevallen op een vorm van een machtsmiddel en/of zal een terugkoppelingsmogelijkheid aanwezig zijn, in die zin, dat ruimte kan worden geschapen voor nader overleg met de leidinggevende of met anderen. Bij belangentegenstellingen gaat het erom het hoe en waarom te verklaren en medewerking te verkrijgen.

Totaalscore:

III - 13

Salarisniveau:

8

Appendix C: Interview questions and topics

Categorie 1: Huidige standaarden

- Op welke manier wordt deze standaard op het moment ingezet / gebruikt bij de beoordeling van het functioneren van medewerkers?
Antwoorden in de richting van: 'een richtlijn', 'een startpunt voor verdere invulling' of 'voorschrijvend'
- Komt dit overeen met hoe de standaard volgens jou bedoeld is?
Antwoorden: Ja/ Nee
 - o Kun je dit verder toelichten?
- Zijn er punten waar je voornamelijk op let / waar voornamelijk op gelet wordt bij de beoordeling van het functioneren van de medewerker?

Categorie 2: Andere standaarden

- Zijn er volgens jou, naast de standaarden beschreven in de functiefamilie en de gemaakte afspraken in de planningsgesprekken nog andere standaarden / criteria die bij een beoordeling worden gehanteerd?
Antwoord: Ja / Nee
 - o Welke standaarden / criteria zijn dit en zijn deze vooraf bekend?
 - o Vraag voor leidinggevend: welke overwegingen heb je om deze standaarden extra te gebruiken?
 - o Vraag voor medewerkers: Vind je dit een waardevolle toevoeging voor de standaarden die tijdens een beoordeling gebruikt worden?
Antwoord: Ja / Nee
 - Waarom is dit wel of niet een waardevolle toevoeging?

Categorie 3: Beoordelingsschaal

De standaarden voor een bepaalde functie zijn organisatiebreed vastgelegd in de functiefamilies (presenteert voorbeeld). Naast de individueel gemaakte afspraken vormen deze standaarden de basis voor de beoordeling. Tijdens een beoordeling wordt er beoordeeld op een zes-puntsschaal: 3 onvoldoende scores (slecht, onvoldoende, bijna voldoende) en 3 voldoende scores (voldoende, goed en uitstekend).

- Kun je aan de hand van de standaarden in de functiefamilie aangeven wanneer elk van de zes scores volgens jou wordt toegekend aan een medewerker? Probeer hierbij een voorbeeld te geven aan de hand van één standaard waarin het onderscheid tussen de scores naar voren komt.
- Vraag voor leidinggevende: Welke score ken jij het vaakst toe en geef je aan wat men moet doen om een score te verbeteren?
Vraag voor medewerker: Welke score heb jij als laatst ontvangen en werd daarbij aangegeven wat je moet doen om een hogere score te krijgen?

Appendix D: Analysis for the second interview question

	Malfunctioning	Insufficient	Almost sufficient	Sufficient	Good	Excellent
Manager 1	Not communicating, no listening	No listening, just sending information	Missing empathy	Mutual understanding, empathy	Knowing when limits are reached, calming customers down	People leave feeling good, employee shows understanding
Manager 2	Discussion with customer, always in need of back-up	Not being independent enough		Independent, customers feel satisfied	Provides good information, customers leave satisfied, independent	Helping colleagues, unburden manager, shows initiative
Manager 3						
Manager 4	Not communicating, bringing the team down	Negative communication	Missing empathy	Making contacts on a wait-and-see basis	Communicating, good choice of words	Exceeding expectations
Manager 5	No external contacts, integrity	Only using known contacts instead of making new ones		Does what position requires, independent, integrity	Integrity, manager can lean on employee	Contacts on a different level than necessary for the position
Medewerker 1	Providing incorrect information, unwilling to cooperate, no necessary knowledge	Customer does not leave feeling satisfied	No adequate basic knowledge, no clear communication towards the customer	Clear communication towards the customer	Customer understands employee, conversations end satisfactory	Continuous high performer, keeps up with relevant developments
Medewerker 2	Not listening, valuing yourself more than others	Disinterest	Interaction is not deemed pleasant by customer			Natural
Medewerker 3	Always locking yourself in a room			Not always interested in colleagues	Display interest in others	
Medewerker 5	Not listening	No listening, disinterest	Trying to conceal disinterest for the customer	Employee feels he's doing his job well	Other people indicate they benefit from employee's communication	Has the sway in the conversation
Medewerker 7	Very cynical, negative attitude	Not representing the municipality	Personal circumstances affect performance	Representing the municipality well, providing objective information, integrity	Municipality and customer are satisfied	Continuous good = excellent
Medewerker 10	No open conversation possible		Not being independent enough	Fulfilling the position from a perspective of common interests		
Medewerker 13	Bad availability by phone	Not always available by phone	Available more than 50% of the time			Being there for each other