



Performance appraisal – how to improve its effectiveness

Subject: Arbeids- en organisatiepsychologie

Student: Anne von Elverfeldt

Student number: s0048348

First teacher: B.P. Veldkamp

Second teacher: R.R. Meijer

University of Twente, Enschede

20-06-05

Summary

Performance appraisal systems need to be effective in improving or sustaining employee performance, otherwise they are a tremendous waste of time and money spend on development and implementation. From literature analyses it became clear that the most significant factor in determining performance appraisal system effectiveness is the acceptance of its users. Thus, a questioning was conducted in a target organisation to test how the users perceive their current performance appraisal system. It was found that factors as *360-degree appraisal*, *procedural justice*, *goal-setting* and *performance feedback* scored relatively high, while *performance-based pay* received the worst score. The only demographic variable that partly accounted for the variance in opinion about factors was age.

Content

1. Introduction	5
2. Performance appraisal system characteristics	7
2.1 Rating approaches	7
2.1.1 360-degree performance appraisal	8
2.2 Rating techniques	10
2.3 Performance-based pay	12
2.4 Rating accuracy, errors and bias	13
2.5 Performance feedback	16
2.6 Training	17
2.7 Employee participation	18
3. The target organisation	20
3.1 The target organisation and its products and services	20
3.2 Historical background of the target organisation	20
3.3 Implications and relevance for this research	21
4. Performance appraisal at the target company	22
4.1 Rating approach	22
4.2 Rating technique	23
4.3 Rating scale	23
4.4 Participants	24
4.5 Raters	24
4.6 Performance appraisal process	24
4.6.1 Preparation	25
4.6.2 Achievement of objectives	26
4.6.3 Agreement of SMART objectives for the coming year	26
4.6.4 Performance feedback	26
4.6.5 Summary - Overall Assessment	26
4.7 Performance-based pay	27
4.8 Performance appraisal training	28
5. Method	29
5.1 Procedure	29
5.2 Participants	29
5.3 Measurement instrument	29
5.4 Analyses	31
6. Results	32
6.1 Results of different factors	32
6.2 Impact of demographic variables	34
6.2.1 Gender	34
6.2.2 Age	35
6.2.3 Tenure	35
6.3 Open question	36
7. Discussion	38

8. Conclusions and implications	40
9. Reflections	41
9.1 Working alone	41
9.2 Support of teachers	41
9.3 Contact with target organisation	42
References	44
Appendix A: Performance appraisal form Germany	47
Appendix B: Performance appraisal questionnaire	55

1. Introduction

Formal performance appraisal has become a widespread instrument of human resource management. Surveys reported in the 1970s and 1980s already indicated that between 74 percent and 96 percent of U.S. organisations, and a comparable proportion of British firms had a formal performance appraisal system in place. Large, complex organisations are especially likely to conduct formal appraisals (Berry, 2003).

According to Jacobs et al. (1980) performance appraisal can be described as a systematic attempt to distinguish the more efficient workers from the less efficient workers and to discriminate among strengths and weaknesses an individual has across many job elements. In short, performance appraisal is a measurement of how well someone performs job-relevant tasks (Parrill, 1999). These measurements are normally done by the direct supervisor of the ratee and can serve different organisational purposes. Examples are employee selection, disciplinary action, development / feedback, promotion, training / supervision and personnel planning.

Execution of performance appraisal means that underlying assumptions to performance appraisal exist. According to Reinke (2003) one of the most basic assumptions is that employees differ in their contribution to the organisation because of individual performance, and that supervisors are actually able and willing to distinguish between employees. Furthermore, for development purposes one assumes that accurate and timely feedback can change behaviour (Tziner et al., 1992) in a way that the organisation as well as the individual is profiting. According to Tziner and Kopelman (2002) this is fostered through the following mechanisms: raters' identification of employees' strengths and weaknesses, the provision of feedback and the facilitation of communication with supervisors. Another assumed aspect is concerned with the practicality of performance appraisal: Time and costs for development and execution phases of the process do not outperform the organisational win which is reached by appraising performance (Jacob et al., 1980).

But there are also some methodological assumptions that are made by those applying formal performance appraisals. Jacobs et al. (1980) describe them: The first is that equivalence exists. This means that the situations under which all ratees are evaluated and the ways different raters actually evaluate ratees are comparable. Second, there are uniformed interpretations of standard expectations and forms among raters. Furthermore, the rater must have the possibility of direct observation plus additional data as for example attendance rates.

Unfortunately, the performance appraisal self and the process are not without flaws. According to Kondrasuk et al. (2002) these problems can be categorized into three areas: (1) the process and format, (2) evaluators role and (3) problems involving the evaluatees.

An example for the first category is the issue of perceived fairness described by Rarick and Baxter (1984) which significantly influences performance appraisal system effectiveness. Fairness is made up by the three different concepts of distributive fairness, procedural fairness and interactional fairness. Distributive fairness is the degree to which rewards and punishments are actually related to performance inputs. According to Swiercz et al. (1999) distributive fairness is the most important predictor of job performance. Procedural fairness describes the degree to which procedures and policies which determine the performance appraisal score are perceived as fair. Davis and Landa (1999) found that the absence of fair procedures increases distress because the results of performance appraisal are essentially outside the control of the employee. But if employees are confident in the fairness of performance appraisal process, they are more likely to accept performance ratings, even adverse ones (Roberts, 2003). Furthermore, procedural fairness is a significant predictor for pay and job-satisfaction (Swiercz et al., 1999). Interactional fairness refers to fair treatment of employees by agents of the organisation. Together, all three types of fairness are predictors of

commitment (Swiercz et al., 1999). Thus, employees' attitudes towards the system will predict how willing they are to buy into the goals they are expected to meet (Harris, 1988).

The second category of problems deals with the evaluator role. According to Kondrasuck et al. (2002) these problems emerge in particular because of conflicting roles of being coach and judge at the same time, lack of rater training or personal bias as favouritism, subjectivity or leniency.

The third category has mainly to do with dissatisfaction with type and amount of feedback and uncomfortable feelings because of lack of control. Especially the way of feedback giving has the potential to arouse negative emotions which in turn negatively affect the acceptance of the performance appraisal system. This partially stems from the fact that performance appraisal has an impact upon an employee's sense of self-worth (Rarick and Baxter, 1984). Combined with the fact that employees tend to overrate their own performance and may feel resentful when receiving appraisals which are lower than they expected (Harris, 1988) it is an explanation why acceptance might be lowered. The lack of user acceptance engenders resistance and a reduction in user motivation (Roberts, 2003) and can result in the undesirable closure of communication between leader and employee around the performance issue (Davis and Landa, 1999).

In summary, it can be concluded that performance appraisal systems become useless if they do not elicit positive reactions among raters and ratees (Tziner and Kopelman, 2002). Generally, this mainly deals with the performance appraisal system being accepted because it is perceived as being distributional and procedural fair and being a valid measure for the position at hand. Thus, in recognition of the large amounts of time and money that need to be invested to develop and implement an appraisal system, an ineffective appraisal system would be a severe threat and loss of resources to an organisation.

On basis of these facts it seems important for each organisation to regularly check if their performance appraisal is perceived as intended and if users still support system and process. Thus, the leading research question is as follows:

Which factors are critical for the successful execution of the performance appraisal system in the target organisation?

In the following, further specific characteristics of performance appraisal systems which according to the literature in particular seem to influence the effectiveness of such a system are discussed. At the hand of a target company and its current performance appraisal system an analysis of performance appraisal effectiveness will be done. Because the effectiveness is dependent on users' perceptions and acceptance, the analysis bases on how raters and ratees actually perceive the system in praxis. If proven necessary, on basis of these results implications for improvement can be given.

2. Performance appraisal system characteristics

Performance appraisals need to be effective; otherwise they are a waste of time and money. What makes them effective is their potential to improve employee performance. But performance appraisal will only lead to behavioural change if its users accept the system. In the following the characteristics of performance appraisal which have proven to partly determine the effectiveness by affecting acceptance are introduced and discussed. First, it is started with more general aspects like rating approaches and rating techniques. Second, the impact of tying performance to pay will be presented. Next, the accuracy of ratings and the existence of errors and bias in ratings will be discussed. The fourth part deals with the way performance feedback is given. Fifth, the influence of training will be analyzed. The last section shows how participation of system users affects the later effectiveness.

2.1 Rating approaches

According to Latham and Wexley (1977) there are mainly three different rating approaches existing: (1) examine cost-related variables, (2) judgements on traits or attitudes and (3) observe and record behaviour. The cost-related approach means that profits, product quantity or quality or return on investment is taken as criteria to appraise an individual. Latham and Wexley (1977) judge this approach as a generally inadequate measure of individual job performance for several reasons: First, it gives no information to the employee why he is effective or ineffective. Thus, there is also no clue how to improve performance. This obstacle is also identified by Harris (1988). He emphasized that a system must be designed in such a manner that it provides guidance for employees relative to how increased performance is to be achieved. Second, these measures are also contaminated by other factors over which the individual has little or no control. The second approach introduced by Latham and Wexley (1977) also does not indicate what exactly the individual has to do differently to improve performance. The interpretation of the feedback is left to the employee. The behaviour approach is a more direct measure of what the employee does or has to change to become more effective. Furthermore, it is less influenced by factors not under control of the employee.

Berry (2003) distinguishes between objective and subjective measures of performance. Objective data can be obtained through production measures and personnel-related measures. The rubric of production measures is similar to the cost-related variables of Latham and Wexley (1977). Additional to the disadvantages identified by Latham and Wexley (1997), Berry (2003) adds that these measures can be deficient in assessing performance on certain jobs. Personnel data for example can include information on various work behaviours such as unruliness, theft or simply attendance. But most of these measures are deficient because they do not address job-specific performance. Berry (2003) thus recommends using this information only as one part of the performance appraisal. Subjective measures are based on human judgement (Berry, 2003). Even though most commonly used, these ratings are far from being perfect measures of performance and they should not be substituted for objective measures when these are available. As indicated by a mean correlation of 0.39, performance ratings are not interchangeable with objective measures.

A third way to distinguish rating approaches is by the sources appraising performance. Generally, the rating task requires someone with special knowledge and ability, and someone who can observe the performance of the employee (Berry, 2003). There are four different single appraisal sources possible: First, the most common source of performance feedback is the immediate supervisor of the employee because the supervisor is expected to have the basic knowledge and ability to conduct performance appraisal on the subordinates (Berry 2003). A second possible source are co-workers or peers of the employee in question. According to Berry (2003) they are in a position to observe and often are aware of how well their

colleagues perform on the job. This feedback source might be especially useful when the supervisor does not have much direct contact with an employee and cannot observe the individual's work. In comparison with the accuracy of supervisory ratings there seems to be a quite strong agreement as indicated by an average correlation of 0.62 (Berry, 2003). Third, performance can be evaluated by the employee in question. This process is called self-appraisal and is often conducted when employees work alone or are relatively independent of others (Berry, 2003). But unfortunately, self-perception differs from how the performance is viewed by others. In particular, employees' self-ratings tend to be more positive than the ratings they receive from others. The fourth possible feedback source are subordinates and is thus only suitable for employees who actually have subordinates. Even though this system has the potential to improve managerial skills it is not without problems. For example, managers might be concerned that some subordinates will rate them negatively because the manager has disciplined the employee in the past. Also, subordinates are likely to fear reprisals if they give negative ratings. The presented single sources can be combined to multi-source or 360-degree performance appraisal systems.

First, it can be concluded that the behaviour approach is most suited for appraising performance. But when available, objective data should be included to give a more complete picture of the employees' performance. Second, performance can be appraised by a supervisor, co-workers, the employee self or subordinates. But to guarantee the acceptance of users, before implementing it must be determined which approach is preferred in the organisation at hand. If only some sources are combined, there is a multi-source performance appraisal system at hand. If there is a complete combination of upward, downward, lateral and self-appraisal then this system is called 360-degree performance appraisal. This special rating approach will be discussed in detail in the following separate section.

2.1.1 360-degree performance appraisal

A quite young variant of performance appraisal is 360-degree performance appraisal. According to Wise (1998) in the typical 360-degree process, supervisor(s), subordinates, peers and (less frequently) internal or external customers provide feedback on performance for each target ratee, using some type of standardised instrument. The ratee is then expected to use the data, along with a self-rating, to make appropriate changes to improve performance.

The purpose of 360-degree performance appraisal is generally the same as for normal performance appraisal, but it is assumed that the new process offers some advantages: Dalton (1996) argues that it provides people with information about the effect of their action on others in the workplace. From the viewpoint of the ratee 360-degree appraisal thus provides one with a rich, textured and multifaceted opportunity to see oneself as others do. Stark et al. (1998) also believe that a notion of behavioural change might be elicited through a process of enhanced self-awareness. Individuals are forced into a cognitive process of reflection that ultimately results in greater levels of awareness of their own actions and the consequences those actions have on others across various levels in and out of the organisation. One example for the usefulness of 360-degree appraisal stems from Raju and Collins (1998). They propose that teambuilding skills of managers, which are one of the most effective means of changing employees' attitudes or satisfaction, might be evaluated by relevant others. Managers may receive feedback from peers and subordinates and make comparisons how their skills are perceived by the two constituencies within a team.

Furthermore Martell and Leavitt (2002) point at the zeitgeist surrounding flatter organisations with the ideal of team-oriented, egalitarian and collaborative culture. This in turn leads to a belief in the practice of fairness in the appraisal process – that is, peers and subordinates should be a part of the process.

In terms of methodology, 360-degree appraisal might improve the subjective measurement of performance by supplementing supervisory ratings with those of multiple raters. Because of the maximum reliability found of supervisory ratings are only 0.6, Rothstein (1990) concludes that ratings by one rater will not provide adequate assessment of performance. Especially the employees possess valid, unique and relevant performance information and insight that is unavailable or unobservable by the rater (Roberts, 2003). This argument is supported by many authors. Raju and Collins (1998) state that peer and subordinate ratings are particularly useful because they provide two different and important perspectives on ratee skill and behaviour. Fecteau and De Vries (2001) even go one step further and admit that supervisory ratings are often plagued by a host of potential problems including bias which might be minimized by addition of alternative sources. The use of multiple rates can reduce bias and errors in performance appraisal because multiple raters can bring non-redundant information to the judgement task (Martell and Leavitt, 2002). But in believing in the error-reducing capacity of 360-degree appraisal one automatically assumes that all rater groups define each performance dimension similarly and raters calibrate rating scale points similarly (Wise, 1998). Wise (1998) found support for this underlying assumption: well-constructed scales can exhibit construct validity across different rater groups. Fecteau and de Vries (2001) also analyzed if different groups of raters are able to share a common conceptualization of performance dimensions underlying the items. They found that the relationship between the items and the constructs they measured were close enough across raters to be regarded as equivalent.

Unfortunately, not all optimistic expectations concerning 360-degree appraisal have become reality. For example Stark et al. (1998) found reactions to peer ratings to be generally negative and significant differences in rating accuracy across sources. Even more important, the perhaps most consistent findings in the empirical literature on performance appraisal systems is that the ratings obtained from different sources generally do not converge. The intercorrelations among the ratings provided by different types of raters tend to be moderate at best (Fecteau and DeVries, 2001). Self-ratings tend to be higher than supervisory ratings, which in turn tend to be higher than subordinate ratings (Fecteau and DeVries, 2001). Even though self-appraisal seems to be especially prone to wrong evaluations, Roberts (2003) suggests that it is anyway useful because it increases preparation and readiness for the interview enhances overall satisfaction and increases perceived fairness. According to Roberts (2003) the ultimate goal is not absolute agreement, but a process directed towards achieving consensus over time.

Fecteau and DeVries (2001) hypothesized which reasons could be responsible for the differences in performance evaluation. First, it might be possible that different rater groups may have different conceptualisations of what constitutes effective performance in a particular job. Second, raters differ in their opportunity to observe any given individuals' work behaviour and raters are exposed to only moderately overlapping sets of ratee behaviour. Third, motivational and informational differences between rating sources, such as self-raters need for self-enhancement and differences in social comparison information available to self-raters and their supervisors, might exist. Fourth, well-established attributional tendencies, such as the self-serving attributional bias and the actor-observer effect might be at hand. Fifth, it is possible that differences in observed ratings from different sources may be substantive differences between the rater groups.

In summary, it can be concluded that 360-degree appraisal might be a useful tool in enriching performance appraisal and enhancing its acceptance. But this will only be the case if raters and ratees generally perceive the additional feedback as relevant and favourable. Thus, it is also possible that even though not the whole package of available feedback sources is wanted, some indeed are. This has to be figured out specifically for the organisation at hand.

2.2 Rating techniques

Rating techniques can be distinguished on basis of several factors (Berry, 2003): use of a rating scale, ease of development, amount and kind of information which is yielded and the purpose of the rating. Because of the variety of different techniques here it is concentrated on the most often researched ones: graphic rating scale, behaviourally anchored rating scale, behavioural observation scale and mixed standard scale.

The most common way for a rater to express a judgement of a ratee's job performance is with a graphic rating scale (GRS). Such scales provide a continuum from high to low performance levels concerning an overall performance or specific performance dimensions (Berry, 2003). Anchors which can be verbal or numerical are placed at the mid- and/or endpoints on the scale. The rater then has to indicate with either on-point or between point responses how the individual has performed. According to Parril (1999) GRS has three advantages: First, this procedure is simple, easily constructed and implemented what makes it a cost-effective method for evaluating employees. Second, the results from that method are standardized what allows comparisons to be made between ratees. Third, because of the ease of use GRS are appealing to evaluators. A possible disadvantage is stated by Tziner and Kopelman (2002). They believe that a ratee is likely to experience fuzziness regarding the activities to pursue in order to accomplish improvements in performance. Furthermore, Berry (2003) points at the problems raters might be confronted with: GRS provide little information to define different performance levels on a certain dimension. Thus, the response options provide little meaning which can result in faulty performance evaluations.

The behaviourally anchored rating scale (BARS) was developed to make the rating task easier what in turn is expected to result in more accurate ratings. BARS use behavioural statements or concrete examples to illustrate multiple levels of performance for each element of performance (Tziner and Kopelman, 2002). The rater then acts as an observer who indicates which behavioural description most closely resembles the ratee's work behaviour instead of requiring the rater to act as a judge who decides whether the ratee's performance on each element is excellent, average or below average (Harrell and Wright, 1990). Rarick and Baxter (1986) summarized the potential advantages: First, raters get a clearer idea of what constitutes good job performance. Thus, the ambiguity concerning expectations is reduced. Second, BARS result in more accurate measurements because of a better understanding of the requirements for good job performance. This opinion is supported by Tziner and Kopelman (2002) who found BARS to be less susceptible than GRS to both halo and leniency effects. Third, a better performance feedback can be given because BARS provides guidelines for improving work performance. Harrell and Wright (1990) add that BARS help rater focus on specific desirable and undesirable incidents of work behaviour which can serve as examples in discussing a rating. This in turn increases the ratee's perception of the feedbacks accuracy. Fourth, there is a better consistency in terms of interrater reliability. This was also found by Tziner and Kopelman (2002). According to their results BARS yielded marginally higher interrater-agreement than GRS: Parrill (1999) identified two more advantages if BARS are developed by the same people who will eventually use them: First, this results in a heightened understanding, awareness and insight. Second, anchors are formulated in language or terminology of raters. This ensures similar interpretations by raters. But BARS is not without disadvantages (Rarick and Baxter, 1986): First, development and use are more costly in terms of time, effort and money. According to Jacobs et al. (1980), studies do not offer evidence that the tremendous amount of time and effort involved in constructing and using BARS are worth the outcome. Second, there is a risk that rater and ratee become more concerned with activity performance rather than accomplishing actual results. Third, rater may not be able to match the observed behaviour with the stipulated anchor. As Tziner and Kopelman (2002)

describe it, raters often have difficulties selecting anchors that most accurately reflect a ratee's performance.

The behavioural observation scale (BOS) is a procedure that was also based on BARS rationale for reducing subjectivity and error in performance appraisal (Latham and Wexley, 1977). This technique asks raters to report the frequency of certain behaviour. BOS results in several advantages: According to Tziner et al. (1992), first BOS produces higher levels of employee satisfaction with appraisal process. Second, because it pinpoints the precise course of action needed to improve performance, BOS seems superior in fostering behaviour change. Tziner and Kopelman (2002) additionally state that BOS appears more likely to minimise barriers in the communication process between superiors and subordinates because it pinpoints for both the specific organisational expectations and performance requirements. This way, role ambiguity and role conflicts are likely to be reduced. Furthermore, BOS decreases raters' fear of possible confrontations with their subordinates because this technique directs the discussion toward the frequency of specific behaviours, rather than toward the raters' evaluation of those behaviours. From the viewpoint of the ratee the feedback is seen as more factual, objective and unbiased. What is problematic about BOS is described by Kane and Bernardin (1982): A given occurrence rate interval does not, in fact, connote a constant level of performance satisfactoriness for all job behaviours. This means that substantial differences potentially exist between the magnitude and satisfactoriness of the occurrence rates of different job behaviours. As Kane and Bernardin (1982) put it, the inevitable result of using a rating scale that ignores such differences between the occurrence rate-satisfactoriness relationships of different job behaviour will be serious distortions in the resulting appraisal.

In the mixed standard scale (MSS) three performance standards are developed per behaviour dimension. These standards reflect average performance, superior and inferior performance. The standards for all behavioural dimensions are then randomly sequenced to form a MSS (Benson et al., 1988). Rater than can indicate whether the focal ratee is better, equal or worse than standard. According to Berry (2003) this technique aims at ensuring that the rater does not simply use an overall impression of the ratee and produce a rating that contains error, such as leniency and halo. But unfortunately the weaknesses of MSS are too obvious: It leaves the rater not only without any clear sense of how good a rating has been given, but also with very little information that the rater can use for feedback to the ratee (Berry, 2003). Benson et al. (1988) add that MSS are difficult to score and only slightly easier to develop than behavioural scales. This comes back in the fact that raters prefer BARS over MSS (Benson et al., 1988). And without acceptance of the system, there will be no acceptance of the data. Furthermore, the validation of MSS may prove difficult because of the extent that MSS are less accurate descriptions of true behaviour any resulting data would be expected to have a substantial error component, attenuating subsequent validity coefficients.

In summary, research indicates that there is no clear picture of which type of scale is the best. Because different scale formats elicit judgement processes, the acceptability and effectiveness of various formats varies across individuals (Härtel, 1993). The influence of the rater's characteristics on the appraisal process was for example identified by Härtel (1993). According to this research it must be distinguished between raters who are field dependent or -independent. Field dependence/ independence are the poles of a cognitive style continuum. Field dependence refers to cognitive dependence on the external organisation of information while field independence refers to the ability to impose organisation independent of the form in which it is perceived. It has been proved that these two cognitive styles are stable personality characteristics. Concerning performance appraisal Härtel (1993) found these styles to influence how raters appraise performance dependent on the rating format. For example it was found that field independent raters are more accurate raters than field dependence raters when performance formats are holistic. Using more highly structured scale

formats led to ratings of field dependent raters that were more accurate. Furthermore, field independent raters also reported greater confidence in their ratings and less confusion and frustration with the rating task. Organisations are in a certain way left to decide for themselves what constitutes the best method of performance appraisal (Parrill, 1999). But to equalise the negative influence of field dependence on rating accuracy it is recommended that the performance appraisal instrument has highly structured scales. According to Jacobs et al. (1980) it is further especially important that the job behaviours included in the rating process are relevant to successful job performance, evaluate the magnitude of importance and frequency of occurrence for each behaviour rated.

2.3 Performance-based pay

One widely accepted notion for improving individual performance is tying pay to performance in order to increase productivity (Swiercz et al., 1999). Performance based pay is a system which specifically seeks to reward employees for their contribution as individuals or as a part of a group, or to reward employees on account of the organisations overall positive performance (De Silva, 1998). There are various types of schemes which fall within the description of performance-based pay. But all of them are designed to share with or distribute to employees the financial results of organisational performance. The schemes fall into the following broad categories: individual-based incentive schemes, profit sharing, gain sharing, employees share option scheme or skill/competence based pay.

The performance-based pay approach has proven to be effective in improving an organisations success. According to Banket et al. (2001) for example the implementation of a performance based incentive plan proved to lead to the attraction and retention of more productive employees. This selection effect occurs because a performance-based compensation contract can act as a screening device that encourages less productive employees to leave and that motivates more productive employees to join or remain with the organisation. Furthermore, the plan motivated remaining employees to continually improve their productivity. This effort effect occurs because a performance based incentive plan motivates employees to learn more productive ways to perform their tasks. De Silva (1998) adds that further benefits of performance-related pay to management and employees are that: (1) where performance/profits increase, higher earnings accrue to employees, (2) employees identification with the success of the business is enhanced, and (3) variations in pay lead to employees becoming more familiar with the fortunes/misfortunes of the business.

Many other authors are not that convinced of the effectiveness of performance-based pay. Critiquers argue that performance-based compensation programs encourage competition rather than collaboration (Solmon and Podgursky, 2000). Because everyone is concerned to secure his own success and thus his own pay, helping others to succeed is not advantageous for oneself. A related point concerning the tendency to undermine teamwork is recognized by De Silva (1998). He argues that individual performance is often difficult to measure objectively, and an exclusively individual performance-related system can damage teamwork. Instead, he proposes team-based criteria in cases where individual performance is difficult to measure, or where there is a need for a corporate culture to promote team values and cooperation, or where the roles of individuals are more flexible, or where the expected performance depends more on team, rather than individual efforts.

Furthermore, Davis and Landa (1999) state that money will buy only a minimum level of commitment. They distinguish between external and internal commitment or motivation. The externally committed employee operates at the level necessary to satisfy the demands of their leaders. Internally committed or motivated employees operate at a level of engagement necessary to provide individual- or self-satisfaction. Thus, they argue that internal commitment is the key factor in an organisations' success because internally committed or

motivated employees are most likely to make significant contributions to the success of an organisation. Internal commitment occurs for example when employees are committed to a particular project, person or program and can be boosted by bottom-up communication.

Also De Silva (1998) argues that performance related pay, if used in isolation, has little impact on motivation or performance. According to him improvement in performance has to be secured through behavioural change effected through a range of measures, such as training and better information/consultation/communication mechanisms. Thus, critical to the success of any performance-related pay system is the practical recognition that it is only one part of a reward system which consists of both financial and non-financial rewards (De Silva, 1998). The non-financial part of a reward system would typically address individual needs such as working in a team, recognition, opportunity to influence decisions, skills development, career opportunities, and a sense of achievement.

Anyway, it is clear that performance-based pay is a highly emotive subject to employees, since it raises subjective perceptions of fairness, indicates worth as an individual to an organisation and may have significance as indicator of social status as well as determining a standard of living. Unfortunately no particular model can be recommended: The system introduced must be conditioned by a variety of factors such as the nature of the business, its business and human resource management strategy (De Silva, 1998). But special attention while implementing a performance-based pay system must be paid to procedural justice. It became obvious that the procedure is very important to employees even when the outcomes for the individual are positive (Swiercz et al., 1999). This means that employees are more concerned with fair procedures than with the outcome of the appraisal process. But distributional justice is far from being irrelevant. The absence of distributional equity increases anxiety over the fairness of the compensation system. This leads to distrust and is thus lessening the productivity because employees become passive. Thus, if the system is not perceived as either being procedural or distributional fair, users' acceptance is in danger.

Other factors why performance-related pay fails are summarized by De Silva (1998): (1) inadequate criteria to measure performance, or criteria which are not easily understood, communicated and accepted, (2) inappropriate performance appraisal systems in that the objectives of the appraisal system do not match the objectives of the reward system, (3) absence of regular feedback on performance, (4) the absence of a right mix of intrinsic and extrinsic rewards, (5) the lack of an appropriate quantum of pay which should be subject to performance criteria, and (6) non-recognition of the fact that performance, especially profit, is sometimes dependent on factors outside the control of employees.

2.4 Rating accuracy, errors and bias

The accuracy of ratings is determined by the reliability and validity of the measurement at hand. Reliability refers to the relative absence of random measurement error in a measurement instrument or precision of a measurement instrument (Harrell and Wright, 1990). According to research theory a measurement is always formed out of a true score plus some error score. But the goal is of course to keep the error component minimal. To measure reliability mainly three different methods exist (Jacobs et al., 1980): The first is the interrater-reliability which assesses the consistency of ratings across different raters. Consistency over time is referred to as retest-reliability. Internal reliability describes if statements are consistent. Rothstein (1990) suggests that the reliability of ratings may be increased by providing sufficient opportunity to observe which will also improve the accuracy of ratings.

Rating accuracy is an important, albeit insufficient condition for feedback to positively affect future performance (Jelley and Goffin, 2001). Unfortunately, there are several different error phenomena which all poses a threat to the accuracy of ratings (Jacobs et al., 1980). In

general, these errors can be differentiated as being related to (1) inadequate observations, (2) faulty standards or expectations about performance, and (3) difficulties in using a rating scale (Berry, 2003).

Recency error is an example for observational errors. It refers to situations in which rater may neglect to pay much attention to an employee until just shortly before the performance appraisal is due. The caused evaluation is likely not to fully represent the ratee's performance (Berry, 2003). Harrell and Wright (1990) examined the influence of cognition on performance ratings as an explanation for such observational error. They found that the fact that individuals have cognitive limits to information processing forms a great obstacle for raters. To deal with it raters often form their own reality on the basis of information available to them, selectively attending to some behaviour while ignoring others.

Leniency and central tendency errors are examples of distributional errors. Leniency error is a tendency to rate higher than ratees deserve (Jacobs et al., 1980). The resulting mean rating score is high, variance among scores is low, and scores are concentrated at the high end of the distribution (Berry, 2003). Other things being equal, the more severe the perceived consequences of a negative rating, the greater the incentive for the rating to be lenient (Dalton, 1996). Evaluators with too lenient ratings are called easy evaluators or "Santa Claus" (Hamman et al., 1999). They are mostly found among groups of evaluators who do not want to put forth the effort to understand the performance standards, or among individuals who have been evaluators for an extremely long time. Severity error describes the opposite phenomena: A rater appears to have excessively high standards which results in a low mean score, and the distribution of scores is skewed toward the low end of the rating scale (Berry, 2003). Such an evaluator is called hard evaluator or "ax man" (Hamman et al., 1999). They often have the problem of being strongly biased by one event, thereby causing their assessments to be extremely harsh. This rater group will typically respond poorly to training. But among new raters there is also a tendency to rate harsher as a result of their application of existing standards. This group will respond well to training and will typically become excellent raters. Central tendency refers to giving no extreme ratings in either the positive or negative direction (Jacobs et al., 1980). Everyone is considered about average, and ratings are concentrated in the middle of the score distribution. Hamman et al. (1999) call this type of rater midline evaluator. According to them this error is due to the rater's feelings of unease with the assessment criteria, and the aversion to make mistakes. With appropriate training this tendency can be eliminated.

Some errors appear to be a result of using a rating scale. Halo error describes a tendency to rate a person in about the same way on all traits or dimensions because of a general, overall impression (Jacobs et al., 1980). The tendency to give similar ratings for dimensions that seem logical related is called logical error (Jacobs et al., 1980). The proximity error refers to the tendency to rate similarly those dimensions which are adjacent on the evaluation form (Jacobs et al., 1980).

Other errors are due to personal biases. If one uses for example oneself as a referent and evaluates the ratee opposite to the way the self is perceives there is a contrast error at hand (Jacobs et al., 1980). The similar-to-me error is the tendency to evaluate the ratee more positive if the ratee is perceived to be similar to the rater (Jacobs et al., 1980). Stereotyping means that impressions about an entire group alter the impression about a group member (Rudner, 1992). If the viewpoint or past experiences affect how behaviour is interpreted one speaks of perception differences (Rudner, 1992).

Even though one can conclude that lack of error implies a certain level of accuracy (Parrill, 1999), there are additional factors as opportunity to observe and delays between observing and giving feedback which have the potential to enhance or deflate the accuracy of ratings. Rothstein (1990) for example found that there is a strong relationship between the

opportunity to observe the ratee and interrater-reliability. It seems that this relationship is strongest in the first year. According to this result, organisations should not routinely use the ratings of one supervisor to evaluate the performance of employees with less than one year on the job or which are observed less than one year by their current supervisor. In general, enhanced opportunities to observe and record behaviour lead to more accurate ratings. But usually delays between observation and performance ratings are imposed by the limited time and resources that organisations are willing to invest in performance appraisal (Sanchez and De La Torre, 1996). This may have severe consequences: Delays may affect the availability of behavioural memories and, therefore, attenuate their relationship with ratings produced under such high memory demands (Sanchez and De La Torre, 1996). Martell and Leavitt (2002) in contrast found that temporal delay affected behaviour recognition more than rating accuracy. Ratees forget specific behaviours over time but retain their initial evaluative impression. Lee (1995) recommends anyway minimizing the need for unaided recall and thus minimising the likelihood of memory decay by steadily recording incidents of good and poor ratee behaviour.

Smith (2001) describes bias as an over- or underidentification of evaluations relative to performance for some groups because of identification that the rater may have toward his or her own or toward a reference group. He argues that these systematic distortions result from limitations in cognitive processing, from individual motivation, or from a combination from both. Ridge (2000) gives an example of biased ratings. According to his research there are favourable evaluations of in-groups and unfavourable evaluations of out-groups existing, even though not necessarily combined. Cook (1995) supports this finding. He found that in-group members achieve their position not by better work, but by some other path. But there are also other examples of bias presented in research of Cook (1995): Performance appraisal shows substantial bias against older persons. Furthermore, there is a small but consistent “own race” bias. Performance appraisals are also strongly biased by appearance, and this bias is not confined to young female ratees.

Two other types of bias are the correspondence and the encoding bias. The correspondence bias refers to the tendency of perceivers to ignore compelling situational explanations when determining the cause of a person’s behaviour and instead wrongly assumes that the behaviour is a direct manifestation of a person’s underlying dispositional state (Martell and Leavitt, 2002). A rater compares perceived ratee features with schemas or prototypes. When the ratee features resemble a conceptual category the ratee is assigned to this category. At this point the risk of encoding bias exist because accessible but arbitrary constructs can bias encoding because ratee behaviour often has ambiguous meaning that makes it open to varied interpretations and extraneous, contextual influences (Kinicki et al., 1988).

Beside errors and bias another type of factors risk the accuracy of ratings. In an in-depth interview for example three quarters of raters freely admitted that they allow liking to inflate performance ratings and dislike to lower performance ratings (Cook, 1995). Even 83% of interviewed raters said that being in a good or bad mood shifts the performance ratings they make. Political considerations or private agendas nearly always influence ratings, too (Banket et al., 2001). Examples are to maximize merit increases, to protect persons whose performance is suffering because of personal problems or just to avoid confrontations. Furthermore, it became obvious that subordinates who helped managers to feel proud of being a good manager, or helped them to make them feel fulfilled, received better performance ratings. Reinke (2003) found that such favouritism undermines employees’ perceptions of procedural fairness and therefore acceptance. Related to this subject, ingratiation and other impression management techniques also contaminate performance appraisal, but also has negative effects on the morale of the staff (Cook, 1995).

In summary, it can be concluded that even though bias and errors never can be totally deleted there is a chance to minimise them. To do so these conditions must be met: (1) make sure that raters have sufficient opportunity to observe their ratees; (2) give raters techniques at hand how to minimise the cognitive demands in evaluating behaviour and the risk of decay; (3) train raters sufficiently in understanding and using the appraisal standards; and (4) design the performance appraisal form and scales in such a way that they are comfortable and easy to use for raters. Concerning the acceptance of a performance appraisal system it is furthermore necessary that appraisees perceive the rating as relatively unbiased, while appraisers must feel able to implement the system as intended.

2.5 Performance feedback

To feedback to the employee generally aims at improving performance effectiveness through stimulating behavioural change. Thus is the manner in which employees receive feedback on their job performance a major factor in determining the success of the performance appraisal system (Harris, 1988). Hearing information about the self discrepant from ones self-image is often difficult and painful. Thus, because feedback may strike at the core of a person's personal belief system it is crucial to set conditions of feedback so that the ratee is able to tolerate, hear, and own discrepant information (Dalton, 1996). Only if conditions facilitate the acceptance of feedback information then the likelihood of change increases. Dalton (1996) further specifies these conditions: The feedback event should be a confidential interaction between a qualified and credible feedback giver and ratee to avoid denial, venting of emotions, and behavioural and mental disengagement. In such an atmosphere discrepancies in evaluations can be discussed and the session can be used as a catalyst to reduce the discrepancies (Jacobs et al., 1980).

Because employees and their supervisors often find appraisal both painful and demotivating Davis and Landa (1999) argue that practice of informal, regular communication between supervisor and employee are far more desirable and effective than formal performance appraisal process. Kondrasuk et al. (2002) also propose to integrate the process of feedback into the daily interactions of supervisors and subordinates in a way that more frequent but less formal meetings. So called achievement updates on a weekly basis then touch upon good and bad issues, while so called achievement assessments take place bi-monthly, are more formal and aim at getting a more clear depiction of issues troubling both sides.

Roberts (2003) instead concludes that effective feedback is timely, specific, and behavioural in nature and presented by a credible source. Tziner et al. (1992) were able to prove that when performance feedback is precise and timely it may result in behaviour change, even though job behaviours are generally difficult to modify. And if during the interview is adequate time for a full discussion of the issues and counselling it will enhance perceived system fairness, system satisfaction, acceptance and supervisory support (Roberts, 2003). Furthermore performance feedback alone generates improvements to ratees' organisational commitment, and particularly to work satisfaction (Tziner and Kopelman, 2002). But performance feedback combined with goal-setting contributes the most strongly to ratees work satisfaction; possibly since goal-setting fosters feelings of participation in work-related issues and meaningfulness at work. Tziner and Kopelman (1992) also found that the process of goal-setting gives the appraisee a broader picture of the work unit and the organisations' objectives. Harris (1988) supports the findings about the positive effects of goal-setting. She recommends an evaluative interview for providing feedback which focuses on problem-solving and goal-setting and which has high employee involvement. Done this way it is more likely to be satisfying to employees than retrospective, subjective interviews.

Roberts (2003) analyzed why goal-setting is so effective and proposed that its effectiveness derives from its ability to focus employees' effort and attention on the critical task at hand, enhancing employee persistence and reducing the likelihood of being distracted. It thus focuses attention and effort on the future which can still be changed. The judgemental performance appraisal process emphasizes past behaviour which cannot be altered anymore. Roberts (2003) also gives instructions how to set goals effectively so that employee performance and satisfaction are enhanced. According to these rules performance goals must be specific, moderately difficult and accepted.

Beside this focus on the future, two more things need to be taken into account. First, according to Wise (1998) it is very risky to give too complex feedback. The more complex the feedback, the more likely recipients will distort it by focusing on results that match their self-perceptions and ignore contradictory ones. Second, Roberts (2003) states that, to be maximally effective, there must be an ongoing formal and informal performance feedback.

In summary, it can be concluded that giving feedback in an appropriate manner is a key factor in determining the employee's willingness to adapt behaviour. First, it is important that the feedback is given in a confidential atmosphere and that the appraiser is perceived as being in state to give useful feedback. Second, the feedback should be precise and detailed and if possible contain examples of the behaviour at hand. Third, feedback giving should leave room for discussion of important aspects, thus involving the ratee in giving opportunity to state his opinion and referring to his problems. Fourth, goal-setting should be part of the feedback. Goals should be clear, as well as the way how to achieve these goals, relevant, specific and moderately difficult. Furthermore, both parties should accept the goals. Fifth, it might be useful to give beside the annual performance review regular informal performance feedback. If all these conditions are met, the acceptance of the feedback will be enhanced and behaviour change will be more likely.

2.6 Training

Nearly all authors agree in the fact that rater must be trained to observe, gather, process, and integrate behaviour-relevant information in order to improve performance appraisal effectiveness. Rudner (1992) for example proposes that training should aim at three goals: First, it should familiarize judges with the measure they will be working with. Second, it must ensure that judges understand the sequence of operations that they must perform. Third, it should explain how the judges should interpret any normative data that they are given. Because errors are well-ingrained habits, Tziner and Kopelman (2002) state that extensive training is necessary for avoiding such errors. Therefore, the training should provide trainees with broad opportunities to practice the specified skills, provide trainees with feedback on their practice appraisal performance, and that a comprehensive acquaintance with the appropriate behaviours to be observed. Harris (1988) also points at the necessity of training: Continued training is needed in areas such as goal-setting and monitoring performance on a frequent basis, and personal and interactional skills. She proposes that an organisation could provide training as this on a regular basis in such a manner that it becomes an accepted part of the supervisor's position and thus becomes a part of the organisations culture. According to Reinke (2003) there needs to be an increased focus on the interpersonal issues surrounding appraisal. The reason is that especially interpersonal issues such as trust are important in the performance appraisal process and should thus be a part of the training program.

Kondrasuk et al. (2002) also aid that the condition of training for all involved individuals must be fulfilled. According to them this means that training is frequently updated and involves appraisal aspects as for example give and take feedback, personal bias, active listening skills and conflict resolution approaches. To build trust and thus enhance acceptance of the performance appraisal process Reinke (2003) adds that a broad understanding of the

system is essential. Harris (1988) also emphasized the importance of training. If implemented this way, employees are less confused, less disappointed concerning measures and are more aware about the intentions of performance appraisal. This also means that they will be capable of useful critique and feedback concerning the appraisal process.

There are several training methods existing: Lee (1985) introduces the term frame-of-reference training (FOR). This kind of training attempts to develop and establish detailed performance scripts. These work as norms or standards for effective performance behaviours and can be used to organise information. In research of Lee (1985) FOR has been found to produce more accurate performance ratings. In contrast, rater error training (RET) is more focused on rating errors frequently committed, and the need to avoid these (Hedge and Kavanagh, 1988). But according to research of the authors RET actually introduces new, and usually inaccurate rating behaviour and should be called inappropriate response set training. Hedge and Kavanagh (1988) conclude that RET is nonsensical if one is interested in improving rating accuracy. Instead they recommend a combination of observation and decision-making training because this might lead to even more improvement in rating accuracy than would either alone.

But even if training improves via several factors performance appraisal effectiveness, Rothstein (1990) emphasizes that it is not likely that training raters could compensate for the lack of opportunity to observe. Good observational skills or a better understanding of common rater errors, coupled with inadequate exposure to the ratee will not lead to reliable ratings.

In summary, it is a necessary condition that raters are sufficiently trained so that they: (1) understand the appraisal process; (2) are able to use the appraisal instrument as intended which includes interpreting standards and use of scales; and (3) are able to give effective feedback including goal-setting. Acceptance will only be gathered if ratees and raters perceive these conditions to be fulfilled. Furthermore, these skills need to be updated or refreshed on a continuing basis. But also ratees should receive a certain form of appraisal training to introduce them to the appraisal system. To attain their acceptance and support of the appraisal system also employees must understand the appraisal system as a whole as well as the behavioural aspects and standards that are used to evaluate their performance.

2.7 Employee participation

Roberts (2003) proposes genuine employee participation in several aspects of the appraisal process because it has the potential to mitigate many of the dysfunctions of traditional performance appraisal systems as well as to engender a more human and ethical human resource management decision-making process. The first participation should according to him take place during the development of reliable, valid, fair and useful performance standards. Second, there should be employee participation during designing the rating format and measurement scales.

The results concerning participation are according to Roberts (2003) constantly positive: Employee participation is a key element of intrinsic motivational strategies that facilitate worker growth and development. Furthermore, employees attain ownership over the performance appraisal process and employees' acceptance is enhanced that way. Third, it generates an atmosphere of cooperation and employee support which reduces appraisal related tension, defensive behaviour and rater-ratee conflict.

Cox (2000) adds that these positive effects are especially generalisable to the design and implementation of pay systems. She suggests that systems implemented following meaningful consultation with employees are more effective than those which are implemented unilaterally by managers or with less employee involvement. Positive impacts were found on a number of factors, including absence and labour turnover rates, throughput and productivity.

Even more important she argues that some of these potential failure factors can be diminished by consultation of employees. First, beside increasing the chance of resolving problems of the current system, seeking information from employees may ensure that the rewards offered are commensurate in timing and kind with the kind that of rewards employees desire. Second, the consultation process may allow the opportunity to identify any individuals or groups likely to be adversely and unfairly affected and to take action to prevent this before the scheme is implemented. Third, involving as many parties as possible in the development of a payment scheme makes them more committed to its success and makes them more likely to accept the system.

The participation of employees functions most effectively in an atmosphere of trust, open communication and equal employee treatment. Therefore, it requires conceptual, affective and experiential education which can be reached by means of training (Roberts 2003). But Roberts (2003) also points at the need to execute regular employee attitude surveys and focus groups to systematic evaluate performance system participation effectiveness.

In summary, it can be concluded that, given the appropriate atmosphere and culture in an organisation, employee participation will enhance motivation, feelings of fairness and overall acceptance of the performance appraisal process. Thus, to attain these positive effects it must be determined if employees in a certain organisation would actually perceive participation as an enriching factor.

Now it became clear theoretically, which performance appraisal factors influence the systems' effectiveness and how these characteristics have an impact on the acceptance of users. In the following part the target organisation and its current performance appraisal system will be presented. Furthermore it will be analyzed how theoretically the effectiveness influencing characteristics are implemented.

3. The target organisation

This part has the goal to describe the target organisation. Thus, some important facts about the business of the target organisation, its structure and its historical background will be presented. Then the relevance for this research will be discussed.

3.1 The target organisation and its products and services

The objective company is a European commercial bank with firm roots in North Rhine-Westphalia, Germany's largest federal state. With Group total assets of Euro 253.8 billion as of December 31, 2005, it is one of Germany's leading financial service providers. It is the central institution for the saving banks in North-Rhine Westphalia and Brandenburg, and as an internationally operating commercial bank it acts as their link to the global financial markets. Working in close partnership with the saving banks, the target organisation offers the full range of products and services of a universal bank, focusing on lending, structured finance, capital market and private equity products, assets management, transaction services and real estate finance. The bank employs 7,154.

As a partner in the Saving Banks Financial Group, the bank provides the regionally operating saving banks with products and services as well as execution. It is also financial partner to major cooperations and larger medium sized companies, institutional clients and public-sector clients. Thanks to its presence in the most important financial centres around the world, the target organisation provides capital market expertise and access to the international capital markets for its clients and the saving banks.

The regional focus of its activities is on Europe and on North-Rhine Westphalia in particular. In a close business partnership based on shared cooperation, individual agreements govern the joint market approach of the target organisation and the saving banks. Medium-sized companies, in particular, benefit from this. For "Mittelstand" clients, the bank concentrates its product mix on export financing, leasing, forfeiting, securisation and selected capital market and equity products.

By providing efficient financial solutions, the target organisation aims to be the premier partner of its clients. In doing so, it seeks to ensure the long-term success of the company for shareholders and employees alike. In accordance with the principles of the Corporate Governance Code, the bank behaves in a responsible and transparent manner towards clients, shareholders, employees and the public. The target organisation is aware of its social responsibility and its actions are guided by the objectives of sustainable development.

In addition to its head office in Düsseldorf and Münster, the target organisation has branch offices in Berlin, Cologne, Dortmund, Frankfurt am Main, Hamburg and Munich. In Europe, the bank maintains branches in Istanbul, London, Madrid, Milan and Paris and outside Europe in Hong Kong, New York, Shanghai, Singapore, Sydney, Tokyo and Toronto. The bank also has representative offices and subsidiaries in all the major trading centres of the world.

3.2 Historical background of the target organisation

The target organisation was created in 1969 by the merger of its old-established predecessor's institutions, Landesbank für Westfalen Girozentrale, Münster, founded in 1832, and Rheinische Girozentrale und Provinzialbank, Düsseldorf, founded in 1854. On August 30, 2002, the bank was transformed into a joint stock company after the public mission activities were integrated into Landesbank NRW, which was established on August 1, 2002. The entire commercial business activities of the former bank are now concentrated in an "Aktiengesellschaft". The public-law guarantee mechanisms (institutional liability and

guarantor liability) continue to apply for the bank until the transition period expires in mid 2005.

3.3 Implications and relevance for this research

First of all it becomes clear that the target organisation is big, very professional, serious and goal-oriented. Organisations of this kind and size usually have a specific Personal Department. This is also true for the target company which employs a business unit Human Resources which is responsible for all aspects of personal related issues. Furthermore, this professional attitude is likely to be also obvious in the way performance appraisal is designed and implemented.

The historical background has as well large implications for appraising employees' performance. Banks are rated by certain instances concerning their liability. The better this rating turns out to be the better are the conditions to which a bank can lend money on the global financial markets. Because of the guarantee mechanisms provided by the federal state the target organisation received the best rating possible. Being able to rely on this fact the target organisation was incredible successful especially in the 1980s. For a very long time there was thus economical security and stability and the bank did not have to face real competition and was even able to choose its customers. For example, there were nearly no private customers accepted except from people with a portfolio worth minimum 500.000 Euro. Thus, just because the bank could afford it, for several decades employees were spoiled and got thus used to a "snobbish" attitude. One example is that for some years payments were raised every single year eight percent. Another example is that each employee received a yearly bonus of a month salary over a period of twenty years. This has become a kind of tradition and employees rely on the payment of the bonus.

Now, because of the judgment from the European court of justice in Brussels the bank is no longer allowed to rely on the guarantee mechanism provided by the federal state because it is against the competition law. Thus, after a transition period the target organisation will have to face "normal" economical circumstances and competition for the first time since foundation. The rating the bank will receive is now only dependent of the own performance of the bank. How good this rating will be will decide about the future survival and success of the bank. It is thus very important for the target company to motivate its staff to improve performance or sustain a high level. Performance appraisal is one of the tools the target organisation has to measure performance, modify behaviour and reward high performance.

The next important step is now to give an overview how performance is currently appraised in the target company. This will be done by describing all aspects relating to the literature analyses: rating approach, rating technique, rating scale, participants, raters, process, performance-based pay and training.

4. Performance appraisal at the target company

The current performance appraisal system aims to provide structured feedback on results achieved and behaviour and skills demonstrated on the basis of objectives agreed. Furthermore, the performance appraisal provides the basis for individual personal development and decision-making for performance-related remuneration in the bank and is intended to facilitate an effective dialogue between employees and managers.

The first of the coming sections gives overview over the chosen rating approach, – technique and –scale. Furthermore, it is described who takes actually part in the performance appraisal and by whom the participants are rated. The next section explains how the performance appraisal process is ideally performed according to the target organisation. The last parts deal with issues as how performance is tied to pay and if or how users are prepared to implement the performance appraisal system with the help of training. At the end of each section a short theoretical evaluations of the characteristics of the current performance appraisal system is provided which leads to expectations how users might perceive they current system.

4.1 Rating approach

The performance appraisal in the target organisation is based on a combination of the cost-related approach and the behaviour approach. The cost-related approach is used in part A of the appraisal, while behavioural aspects are integrated in both parts, A and B.

In part A of the appraisal, the definition of objectives takes place for the following year. The objectives refer to five broad categories: (1) management, (2) clients, (3) projects, (4) P&L and (5) processes/objectives relating specifically to the function. The category management includes all objectives related to the management of a team, for example motivation of employees, recruitment or reduction of staff, team formation, conducting performance appraisal, development and coaching of employees, pooling of teams. The second category clients deals with all objectives related to the provision of services for internal and external clients, for example for acquiring new clients, development and maintenance of client relationships, quality/number of business transactions concluded, opening-up of new business fields. Within the category projects, all objectives related to the management of or participation in projects, for example development of strategies, development of information systems, and definition of new standards are contained. The category P&L refers to all objectives related to revenue generation and cost saving. Processes/objectives relating specifically to the function as the last category describes all objectives contributing to the improvement of processes, for example pursuit of new ways of working or changes in process organisation. Furthermore, all objectives related to the performance of tasks generally involved in the position. Thus, where possible within the objectives it is made use of objective data as criteria of successful performance.

Part B of the appraisal is based on observing and recording job-relevant behaviour, thus exclusively subjective data. There are five groups of competencies to be appraised which are a combination of specific behaviour and skills: (1) specialist and methodological skills, for example “the employee plans and organises working procedures in such a way that tasks are completed efficiently and on time”, (2) personal skills, for example “the employee questions his/her own working methods/results so that he recognises mistakes and learns from them”, (3) interpersonal skills, for example “the employee is aware of conflict and seeks to achieve clarification”, (4) entrepreneurial skills, for example “the employee recognises and exploits profit opportunities that are in accordance with bank interests”, and (5) management skills, for example “the manager delegates authority and responsibility to allow employees to work autonomously”. With the exception of management skills, they apply to all employees. If

requested, comments can be made for each competency in the relevant field, for example regarding specific skills.

The rating approach of the target organisation is in line with the suggestions from literature. There it was stated that generally it could be concluded that the behavioural approach is most suited for appraising performance. But when available, objective data should be included to give a more complete picture of the employees' performance. By this the employee gets clues how to improve performance, but at the same time the subjectivity is tried to be minimised. The performance appraisal system in the target organisation includes evaluation on objective, cost-related data, for example number of acquired new clients, but also on subjective, behavioural data, for example personal skills. Thus, one could expect that employees are satisfied with the rating approach, perceive it as being fair, objective and relevant for the individual job at hand.

4.2 Rating technique

The rating technique of the behavioural part of the performance appraisal can be described as a graphic rating scale. This theoretically means that behavioural dimensions are presented on which the performance of the employee in question has to be indicated with the help of a scale. In the performance appraisal at the target organisation, the objectives as well as the five groups of competencies to be appraised are a combination of specific behaviour and skills. With the exception of "Managing the Team", they apply to all employees. If requested, comments can be made for each competency in the relevant field, for example regarding specific skills.

From literature the strength and weaknesses of this technique became obvious: The advantages can be summarized as (1) procedure being simple, easily constructed and implemented, thus cost-effective; (2) results being standardized, thus allowing comparisons between ratees; and (3) being easy to use, thus appealing to evaluators. The disadvantages are (1) ratees' fuzziness regarding the activities to pursue in order to accomplish improvements in performance; and (2) lack of information how to define different performance levels on a certain dimension.

Research results were not able to prove the superiority of one rating technique over the other. As a consequence, the actual choice is less important, as long as the performance appraisal instrument has highly structured scales. Furthermore, it is especially important that the job behaviours included in the rating process are relevant to successful job performance, evaluate the magnitude of importance and frequency of occurrence for each behaviour rated. These factors seem to be partly realized: The scales are highly structured, to assure relevance objectives can be customized to the job at hand and the magnitude of importance can be assigned in part A, while frequency of occurrence is evaluated in part A and B. Thus, one could expect that ratees perceive the performance appraisal instrument as valid for their individual job, while raters feel comfortable with the use of the instrument.

4.3 Rating scale

The graphic rating scale used in the performance appraisal of the target organisation has five verbal described points: A, B, C, D and E. An "A" indicates aspects of conduct and performance which do not (yet) meet the requirements of the position, for example where there is still need for improvement. This may be perfectly normal, for example, in a job-training phase. A "B" is described as generally meeting the expectations, but still having a need for improvement in individual situations. A "C" means overall good performance. A very good performance, worthy of special recognition, partly beyond expectations of the position is coded in a "D". If one wants to indicate specially identified strengths which go well beyond what is expected of employees in this position one uses an "E". Thus, the five

points are grouped in three blocks. An “A” means (still) unsatisfactory performance, “B”, “C” and “D” are in line with the requirements and “E” means outstanding performance. The scale allows only for on-point indications.

With five possible points the scale is quite compact and should as a consequence contribute to the ease and comfort of use which was emphasized in literature. This leads to the expectation that raters probably are satisfied with the use scale. On the other hand, because of the possibility to chose the mid point, one could assume that occurrence of central tendency error is heightened. If so, this in turn could have a negative impact on if ratees perceive their ratings being free from bias and error. This influences negatively the perceptions of distributive justice, too.

4.4 Participants

All employees with an unlimited employment contract with the target company participate in performance appraisal. Temporary employees who have been working for the bank for more than 6 months are appraised in freely formulated terms. Employees who were granted leave of absence in the course of the year are appraised for the period of their active employment. Apprentices, trainees and employees who are employed within the framework of “personnel transfers” are appraised according to a different procedure. In case of longer leave or secondment the applied rules of transfer are also applying.

From literature it became clear that performance appraisal system users emphasize procedural fairness. Because generally everybody gets appraised, but for special conditions special procedures are applied, one could expect that users perceive this system as being procedural fair.

4.5 Raters

The line manager who has disciplinary responsibility in the month of December of the year under review shall carry out the performance appraisal process. If several managers are responsible for a unit, the leading manager may delegate this task to team leaders, for instance, on condition that the employee to be appraised agrees and that the appraisers have participated in a performance appraisal training seminar. The period of appraisal corresponds to the calendar year. Deadlines for completion are published each year.

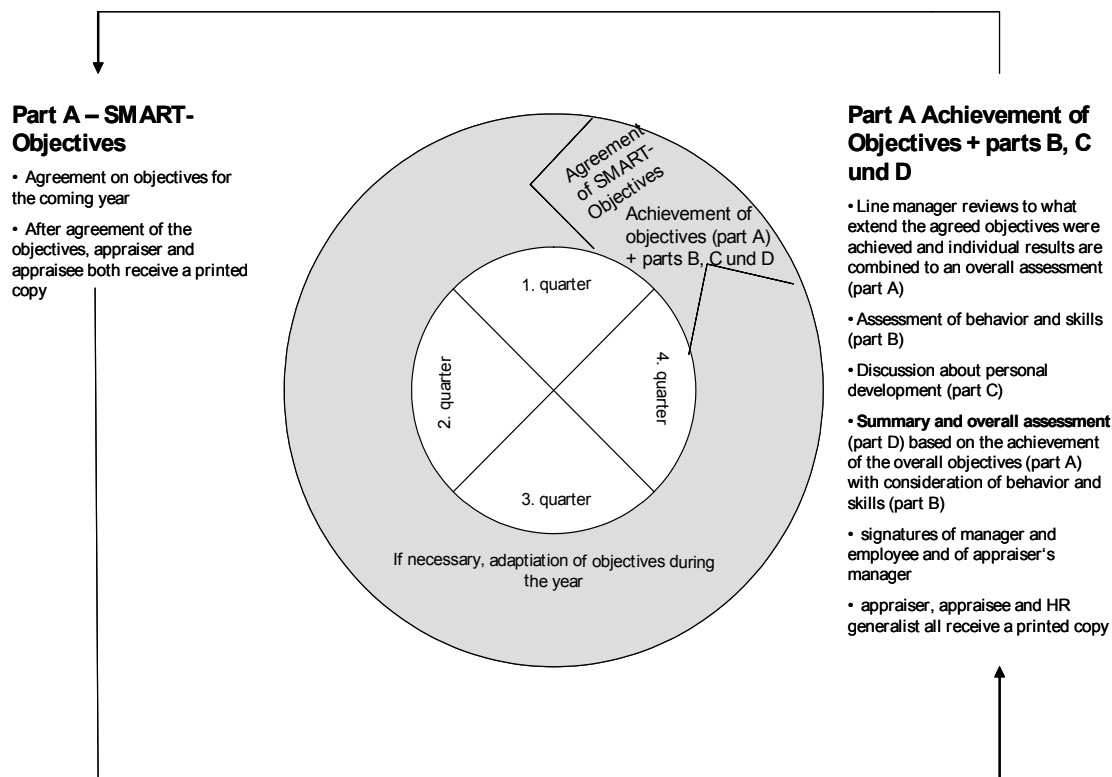
Employees working permanently in projects of at least 6 months duration are appraised by their project managers upon termination of a project or at the time of the annual performance appraisal round. Relevant objectives are agreed at the start of the project. Project managers pass the performance appraisal forms to the manager with disciplinary responsibility who takes account of them in the appraisal at year-end and coordinates the “Overall Assessment“ with the individual project managers. For employees who work partly but predominantly in specific projects, the line manager responsible for appraisal requests information on performance, behaviour and skills demonstrated and includes it in the appraisal.

One important point following from literature was that it is absolutely necessary for accurate evaluations that they are done by someone who has sufficient opportunity to observe and record behaviour. The guidelines for performance appraisal are aligned with this finding. Furthermore, it is guaranteed that employees get only rated by persons who have finished the according training. Thus, these facts should lead to high perceptions of the rater’s ability to evaluate behaviour accurately and to high perceptions of procedural and distributive justice.

4.6 Performance appraisal process

Prior to the appraisal discussion, the appraiser should complete the following sections of the performance appraisal form for the past year if available: (1) part A “Achievement of

objectives previously set”; (2) part B “Skills assessment”; and (3) part D “Overall performance assessment”. The following sections should be prepared prior to and completed after the appraisal discussion: (1) part C “Personal/further development” and (2) part A “Agreement of SMART-objectives”. After the appraisal discussion, the line manager should complete part D “Summary”. The process is in short visualized in figure 4.2. The important parts of the process will be described in the following sections.



4.6 Performance appraisal process in brief

4.6.1 Preparation

It is emphasized that it is important for both the appraiser and the appraisee to prepare for the discussion on the agreement on objectives and on feedback so that both parties can gain maximum benefit. Thus, preparation is the first step in the performance appraisal process.

It is recommended that the appraiser should prepare by: (1) reviewing how the individual has performed against the objectives set and the skills required; (2) discussing the individual's performance with the previous line manager in case of internal transfer; (3) reviewing the feedback given during the year; (4) reviewing whether the employee has received the level of support required for fulfilling the tasks and whether they have adhered to previously agreed personal development measures; (5) formulating individual objectives for the next period of appraisal clearly and understandably; (6) planning individual development measures; (7) taking account of the whole appraisal period to give a balanced appraisal; (8) not placing undue emphasis on recent events; and (9) not dwelling on isolated incidents.

For the appraisee it is recommended that he prepares by: (1) reviewing how the agreed objectives were achieved; (2) reflecting why objectives were not/not fully met or exceeded; (3) thinking about possible objectives for the next year; (4) thinking about personal

development and relevant measures; (5) thinking about feedback and requests to be addressed to the appraiser; and (6) making proposals for improvement and further development of the unit/department.

4.6.2 Achievement of objectives

The second step is that the line manager reviews to what extent the agreed objectives were achieved. He/she should try to explore the reasons for under/over-performance (for example change in market conditions, lack of resources, legislative changes, organisational changes etc.). Then the appraiser should select a rating for each objective using the known scale and add comments to explain the decision. If an objective is no longer applicable the rater has to assign a (N/A).

If an objective is not achieved or exceeded due to circumstances beyond the control of the appraisee, the assessment should be adapted accordingly. The individual results are then combined to an overall assessment. A balanced assessment should be made which is not merely the arithmetic average.

4.6.3 Agreement of SMART objectives for the coming year

Next, within five defined categories, objectives derived from the business objectives of the unit should be determined for the coming year and given a percentage weighting. Objectives agreed on should be “SMART“: (1) Specified – The objective is clearly defined, (2) Measurable - The objective is measurable, and results can be verified; (3) Accepted – The objective is the result of a clear agreement on operating and business objectives; (4) Realistic – The objective can be achieved in the current business context and is part of the overall strategy; and (5) Time-related – The period for completion of the objective is determined. The agreement of objectives should be accepted by both parties. Only if conflicts arise that cannot be resolved, the line manager decides on the strength of his/her overall responsibility for the organisational unit and makes the final decision.

4.6.4 Performance feedback

The fourth step is giving performance feedback in the appraisal discussion. The performance appraisal discussion is intended to create a constructive dialogue between line manager and employee based on mutual respect. In order to provide maximum benefit for both parties, the target organisation thus suggests the following: (1) Rater and ratee should agree on a mutually convenient time and place in a quiet and undisturbed atmosphere; (2) Rater and ratee should agree the date a few days in advance and request the appraisee to prepare, for example by reading this guidelines, reviewing personal performance etc.; (3) It should be taken time for the discussion; (4) It should be taken into account that approximately half of the time is required for setting the objectives for the following year and planning the personal development measures; (5) It should be an open and two-way discussion in which both parties divide the time equally; (6) Raters should express both positive and critical feedback. Positive feedback reflects appreciation and provides motivation. Critical feedback helps to describe problems, examine reasons and agree on possibilities for improvement; (7) Feedback should be confined to the behaviour demonstrated and the objectives achieved rather than to features of personality such as intelligence; (8) Raters should give examples, because they help to explain assessments; (9) It should be considered that people react to feedback differently. Being sensitive to this fact and formulating feedback to take account of the person being assessed.

4.6.5 Summary - Overall Assessment

The last step is the forming of an overall assessment. This is based on the achievement of the overall objectives under section A (SMART-Objectives). However, as behaviour and skills

(section B) are of great importance, these features should also be considered in the overall assessment. Hence, the line manager decides whether the “Overall Performance Assessment” taking into account the behaviour as shown under part B differs from the assessment under part A, and determines the assessment on the basis of the known scale (A through E).

Here, an “A” means that objectives are not achieved which is an unsatisfactory performance. A “B” indicates that objectives are partially achieved, but there is scope for improvement in individual situations. If objectives are achieved and one can speak of a good performance within the framework of the requirements, a “C” has to be chosen. A “D” means that the objectives are exceeded, thus a very good performance which partially exceeds requirements. An “E” marks an outstanding performance which exceeds requirements by far, thus, the objectives are substantially exceeded, too.

In the field „Comments“ the line manager can add comments to the overall assessment. The assessment serves as a basis for the decision on performance-related remuneration.

In summary, the performance appraisal process is highly structured and there are clear guidelines established what ratee and rater should do to make the performance appraisal successful. Especially for performance feedback and goal-setting a lot of recommendations are given which match the findings from literature. Thus, if the guidelines are actually put into praxis as intended this should result into high perceptions of procedural fairness, satisfaction with performance feedback and goal-setting. For the forming of an overall assessment rather vague guidelines are established. Thus, one might expect that ratees often do not understand how the assessment is actually formed.

4.7 Performance-based pay

Seniority or length of service used to be the main criterion in pay determination in the target organisation. Now, with Vorstandsbeschluss of August, 2004, a more performance-based payment process was requested in the target organisation to emphasize the employees' contribution to the overall performance. This process is based on the resulting performance evaluations of the annual performance appraisal and concerns the bonus or win participation paid in April each year.

The process is as follows: First, the bonus pool is defined each year on level of the whole bank. The bonus pool is a certain percentage of the organisations win determined at the end of the business year. Then, this bonus pool is spread over the different segments of the bank. This is done with respect to for example the history, segments performance and contribution to win. Third, on basis of the overall assessment of the performance appraisal a certain bonus is then allocated to the individual employee. Thus, this performance-based pay system is a mix of profit share scheme and performance bonus. Employees are given a share of the profit, but the amount of share is in the end determined by the individual performance. The amount of bonus varies from a half-month salary to two-month salaries.

Because of the influence of the worker's council it is not allowed that the overall assessment is formed by arithmetic approach. This is because it is tried to avoid that a direct comparison of employees' appraisal evaluations is made possible. Accordingly, only the evaluations “A”, “B”, “C”, “D” and “E” are allowed. A further partition may not be done, even though it is well possible to allocate different boni to employees with similar position and same scale grade if: (1) they are working in different branches, (2) or they differ within a certain scale grade.

To get paid for performance is thus a very young concept in the target organisation. This combined with the fact that before for years everybody received a month salary as a bonus independent of actual performance, one can expect that most ratees are not satisfied with the new approach. Furthermore, it is questionable if the amount of pay employees can

earn is effective in motivating. Another problem is that the bonus is defined by many different instances and is thus influenced by circumstances that are out of control for the individual employee. So it is likely that ratees experience fuzziness regarding process and lose motivation because of their lack of control. In summary, it is to expect that users are dissatisfied with the performance-based pay system.

4.8 Performance appraisal training

In the target organisation official performance appraisal training is organised for supervisors and team- or project leaders who have to rate employees as a part of their supervising position. This training takes two days. During the first day the meaning of performance appraisal to the organisation is introduced. Then, it is trained how to use the performance appraisal system of the target company. The last goal of the first day is to prepare supervisors to determine SMART-objectives and how to prepare the talk about them from content. During the second day supervisors learn how to check their own appraisal behaviour. Furthermore, they get trained in preparing performance reviews. Then, they will get training how to form the overall assessment and how to explain it. The last goal is that supervisors are able to detect developmental potential in employees and to plan interventions accordingly. For each of the training days there are four possible dates per year. A refreshing or updating of the skills is not planned.

For employees who will only get rated there is no special training organised. When they enter the organisation they will get short information about the existing of performance appraisal in the target organisation. They actually get the system to know when goals are set for them or they get appraised for the first time.

If something changes in how performance is appraised then information will be spread by the bottom down communication process, articles in the quarterly magazine of the target organisation or information meetings.

In summary, raters receive an introductory training. Thus, ratees should perceive the raters as being able to conduct the performance appraisal. Raters themselves should feel able to carry out the process as intended, to use the instrument as intended, to set goals as intended, to evaluate ratees accurately and to give useful feedback. Because of the missing training for ratees it is possible that they do not know or understand the instrument and standards used to evaluate them.

The next part describes how the measurement to test the expectations concerning the different factors was conducted in the target organisation. Thus, it gives an overview over the procedure, the participants, the measurement instrument and the statistical analyses.

5. Method

This part aims at giving an overview over the four aspects of the measurement conducted: First of all, a description of the measurement procedure is given. Then, it is shown who actually took part in the measurement. The third section explains the measurement instrument and its factors. This part ends with naming the statistical analyses that were conducted.

5.1 Procedure

A questioning was conducted in the targets organisations segment “Banking Operations IT”. The segment Banking Operations IT cares for the information technology systems of the bank during their whole life cycle which means from introduction until closure. These IT systems typically are trade systems or transaction systems. The segment has approximately 110 employees who are spread over the branches of Düsseldorf, Münster and London. Questionnaires were given to individuals who voluntarily wanted to take part.

5.2 Participants

43 employees took part in the research which results in a response rate of 47,3%. Of the 43 participants 15 were supervisors or team leaders with rating tasks while all of the 43 employees get annually rated. 4 raters were female (40%) and 11 raters were male (60%). The raters ranged in age from 27 to 57 years ($M = 36,67$; $SD = 8,27$) and in tenure in organisation from 4 years to 22 years ($M = 8,47$; $SD = 5,33$). 9 ratees were female (20,9%) and 34 ratees were male (79,9%). The ratees ranged in age from 21 to 60 years ($M = 38,4$; $SD = 9,1$) and in tenure in organisation from 2 years to 34 years ($M = 7,89$; $SD = 6,3$). The open question was answered by 13 respondents (30,23%).

5.3 Measurement instrument

A questionnaire was developed to measure the factors found to influence the effectiveness of performance appraisal instruments. The questionnaire consists of four separate parts:

In the first part “Demographics”, some demographic data like age, gender and tenure in organisation were gathered.

In the second part “Statements ratees”, 47 statements from the viewpoint of a ratee regarding performance appraisal were presented. There, respondents had to indicate their level of agreement with each statement. This was done with an five-point Likert-scale ranging from one to five, with one being “I absolutely disagree”, three being a neutral score and five being “I absolutely agree”. Eight influencing factors were measured: instrument validity, distributive justice, procedural justice, goal-setting, performance feedback, performance-based pay, employee participation and 360-degree appraisal.

Instrument validity refers to the degree to which the instrument measures what it is intended to measure. This factor was measured by seven statements like “My supervisor and I agree on what equals good job performance” and “The performance appraisal instrument accurately measures what I do on my job”. Reliability analyses revealed that all seven original items had a Cronbach’s alpha of 0,765. With elimination of item four (“I understand the measures used to evaluate my performance”) this value could even be increased up to an alpha of 0,776.

Distributive justice refers to the degree to which the outcomes of performance appraisal are perceived as being fair by employees. This factor was measured by four statements like “The best workers receive the highest evaluation scores” and “The performance appraisal process results in a clear and unbiased appraisal”. Reliability analyses revealed that all four original items together had a Cronbach’s alpha of 0,762. The elimination of any of the items would not have had a significant increasing effect on the reliability.

Procedural justice refers to the degree to which the appraisal process is being perceived as fair by employees. This factor was measured by seven statements like “My supervisor utilizes the evaluation system to assess my performance objectively and without bias” and “I feel treated fairly during the performance appraisal process”. Reliability analyses revealed that all seven original items had a Cronbach’s alpha of 0,792. With elimination of item five (“I know the standards used to evaluate my performance”) this value could even be increased up to an alpha of 0,847.

Goal-setting refers to the perceived quality of goals set and their acceptance by employees. This factor was measured by eight statements like “I agree that the performance goals set up for me are reasonable” and “I am determined to achieve my performance goals”. Reliability analyses revealed that all eight original items had a Cronbach’s alpha of 0,818. The elimination of any of the items would not have had a significant increasing effect on the reliability.

Performance feedback refers to the perceived quality of performance feedback provided to employees and their satisfaction with the feedback. This factor is measured by seven statements like “The performance feedback I receive is helpful in improving my on-the-job performance and in attaining my goals” and “Throughout my performance feedback I feel that I have the possibility to discuss work-related issues with the supervisor”. Reliability analyses revealed that all seven original items had a Cronbach’s alpha of 0,895. With elimination of item two (“I receive regular and timely performance feedback beside the annual performance review”) this value could even be increased up to an alpha of 0,910. Of all eleven factors this was the highest reliability.

Performance-based pay refers to the degree to which employees perceive performance-based pay as motivating and fair. This factor is measured by five statements like “The forming of a global evaluation is clear and valid” and “There is a clear, direct and compelling linkage between performance and pay in the performance appraisal system”. Reliability analyses revealed that all five original items had a Cronbach’s alpha of 0,490. With elimination of item four (“I believe the amount of performance-based pay I can earn through high evaluation ratings will make a noticeable difference in my future performance”) and item five (“Performance-based pay based on performance ratings is the most effective method for motivating employees to improve/sustain performance”) this value could be increased up to an alpha of 0,683.

Employee participation refers to the degree employees would appreciate the participation of employees in developing or designing an appraisal instrument or even themselves would be willing to participate. This factor is measured by three statements like “I would be willing to participate in developing a new performance appraisal system” and “Participation of employees in the development of performance standards leads to a better performance appraisal instrument”. Reliability analyses revealed that all three original items had a Cronbach’s alpha of 0,783. The elimination of item one (“I would be willing to participate in developing a new performance appraisal system”) would have had a significant increasing effect on the reliability (alpha = 0,926) but in regard of the small number of items this was not done.

360-degree appraisal refers to the degree to which employees would prefer to be additionally judged by different instances and if they think that this process would enhance the accuracy of performance appraisals. This factor is measured by five statements like “Being appraised by several sources (supervisor, peers, customers etc) would enhance the accuracy of performance appraisals” and “I would prefer my performance to be evaluated additionally by my colleagues because they have relevant performance information and insight”. Reliability analyses revealed that all five original items had a Cronbach’s alpha of 0,557. With elimination of item two (“I would like to give my supervisor feedback on his/her

performance”) and item three (“Appraising my own performance would enhance my awareness of my own performance”) this value could be raised up to an alpha of 0,710.

In the third part of the questionnaire “Open question” there was one open question which could be answered by respondents in case of concrete ideas or suggestions for improving performance appraisal.

The fourth part “Statements raters” had to be answered additionally by persons who also have to appraise others (commonly supervisors or team leaders). The procedure was the same as in the second part: 15 statements were presented and the level of agreement had to be indicated with five-point Likert-scale ranging from one to five, with one being “I absolutely disagree”, three being a neutral score and five being “I absolutely agree”. Three influencing factors were measured: rating techniques, rating accuracy and training.

Rating technique refers to the degree the current rating format is being perceived as easy and comfortable to use and if they allow accurate performance appraisals. This factor was measured by five statements like “I feel that the scales allow an accurate assessment of different dimensions of performance” and “The existing form is too complex”. Reliability analyses revealed that all five original items had a Cronbach’s alpha of 0,754. The elimination of any of the items would not have had a significant increasing effect on the reliability.

Rating accuracy refers to the degree to which raters feel able to make accurate ratings. This factor is measured by four statements like “I feel I have enough information regarding performance standards to make accurate judgements about employees on each performance dimension” and “I am motivated to correctly evaluate employees’ behaviour”. Reliability analyses revealed that all four original items had a Cronbach’s alpha of 0,338. With elimination of item two (“I regularly record incidents of good/poor behaviour relevant for the performance evaluation of employees”) this value could be raised up to an alpha of 0,512. Of all eleven factors this was the lowest reliability but elimination of further items would not have had an increasing impact on Cronbach’s alpha.

Training refers to the degree to which raters feel sufficiently trained in giving feedback and using the instrument. This factor is measured by six statements like “I am sufficiently trained in all skills needed in appraising performance” and “My appraisal skills are regularly refreshed and updated through training”. Reliability analyses revealed that all six original items had a Cronbach’s alpha of 0,448. With elimination of item one (“I need more training in conducting performance appraisal interviews”) and item five (“I am able to use the appraisal instrument as intended”) this value could even be raised up to an alpha of 0,806.

A factor analyses was done to test if the proposed eleven factors could be found back in answering structure of respondents. But unfortunately, the existence of the distinct factors as introduced could not be proved. This is possibly due to the small amount of respondents.

5.4 Analyses

All statistical analyses were conducted with SPSS version 12.0. In the beginning the instrument itself was the target of analyses. As above mentioned the reliability of the instrument was determined per factor. Where necessary items that were responsible for decreased reliability values were deleted from further analyses. In sum, ten items had to be eliminated. With factor analyses it was then tried to prove the existence of the eleven distinct factors as proposed in the instrument. Unfortunately, this could not be proven.

The second phase, three different analyses were conducted: (1) some descriptive statistics to explore the strength and weaknesses of the performance appraisal system as perceived by users; (2) independent samples t-test to control for impact of gender of respondents; and (3) linear regression to test if age and tenure had influence on perceptions. The corresponding results are presented in the following section.

6. Results

In this section the results of the different analyses are presented. The first part deals with the descriptive statistics which gives an impression how the several factors are perceived by the performance appraisal system users of the target organisation. Next, the results of the independent samples t-test and linear regression analyses are presented. From this, it can be concluded if demographic variables as age, gender or tenure influence the perceptions of the performance appraisal system. Third, it is summarized which ideas and suggestions respondents had to improve performance appraisal at the target organisation. This last part gives first of all a completion of the picture that users have and secondly inspiration for improvements.

6.1 Results of different factors

To find out how users of the appraisal system perceive the different factors the total mean scores of the factors and the mean scores of the individual items of a factor were compared. From this one can conclude where the respondents feel the system has strengths or weaknesses. The results are presented per factor.

The descriptive statistics show that the total mean score for *instrument validity* is $M = 3,26$ with a standard deviation of 0,66. Comparison of the mean values of the separate items revealed that respondents evaluated item seven (“Raters share a common conceptualisation of the performance standards”, $M = 2,63$) worse than all other items of this factor (ranging from $M = 3,07$ to $M = 3,65$).

For the factor *distributive justice* the total mean score is $M = 3,10$ with a standard deviation of 0,76. Comparison of the mean values of the individual items showed that respondents evaluated item two (“The performance appraisal process results in better communication between me and my supervisor, $M = 3,44$) better than all other items of this factor (ranging from $M = 2,95$ to $M = 3,02$).

The total mean score for *procedural justice* is $M = 3,88$ with a standard deviation of 0,68. Comparison of the mean values of the individual items showed that respondents evaluated item three (“If I have problems with my performance evaluation I can communicate my concerns openly to my supervisor”, $M = 4,07$), item four (“my supervisor will be ethical in how he/she scores my performance”, $M = 4,07$) and especially item six (“My supervisor takes the performance appraisal procedure seriously”, $M = 4,12$) better than all other items of this factor (ranging from $M = 3,56$ to $M = 3,88$).

Descriptive statistics revealed a total mean score for *goal-setting* of $M = 3,79$ with a standard deviation of 0,58. Comparison of the mean values of the individual items showed that respondents evaluated item three (“I accept the goals I have been assigned”, $M = 4,16$), item four (“I agree that the performance goals set up for me are reasonable”, $M = 4,05$) and item five (“I am determined to achieve my performance goals”, $M = 4,16$) better than all other items of this factor (ranging from $M = 3,21$ to $M = 3,88$).

The total mean score for *performance feedback* with $M = 3,74$ is quite similar to *goal-setting*. The standard deviation is 0,76. Comparison of the mean values of the individual items showed that all were answered quite similar (ranging from $M = 3,65$ to $M = 3,98$).

According to the descriptive statistics the total mean score for *performance-based pay* is $M = 2,89$ with a standard deviation of 0,76. Of all eleven factors this was the lowest total mean score. Comparison of the mean values of the individual items showed that respondents evaluated item three (“There is a clear and reasonable process established for grieving both evaluation and performance-based pay results”, $M = 2,72$) worse than all other items of this factor (ranging from $M = 2,93$ to $M = 3,02$).

Employee participation has a total mean score for of $M = 3,47$ with a standard deviation of 0,81. Comparison of the mean values of the individual items showed that

respondents evaluated item two (“Participation of employees in the development of performance standards leads to a better performance appraisal instrument”; M = 3,65) better than all other items of this factor (ranging from M = 3,33 to M = 3,44).

The total mean score for *360-degree appraisal* is M = 3,91 with a standard deviation of 0,70. Of all factors this was the highest total mean score. Comparison of the mean values of the individual items showed that respondents evaluated item four (“Being appraised by several sources (supervisor, peers, customers etc) would enhance the accuracy of performance appraisals”, M = 4,07) and item five (Being appraised by several sources (supervisors, peers, customer etc) would provide me with valuable information about different important aspects of my performance”, M = 4,19) better than all other items of this factor (ranging from M = 3,49 to M = 3,74).

With a total mean score of M = 3,17, the factor *rating technique* has the lowest total mean score of the three rater-factors. The standard deviation is 0,76. Comparison of the mean values of the individual items showed that respondents evaluated item one (“I feel comfortable with the scales used to evaluate performance”, M = 2,93) and item two (“I feel that the scales allow an accurate assessment of different dimensions of performance”, M = 2,80) worse than all other items of this factor (ranging from M = 3,20 to M = 3,67).

The total mean score for the second rater factor *rating accuracy* is M = 3,60 with a standard deviation of 0,62. Comparison of the mean values of the individual items showed that respondents evaluated item one (“I am motivated to correctly evaluate employees` behaviour”, M = 4,13) better than all other items of this factor (ranging from M = 3,20 to M = 3,47).

According to the descriptive statistics the total mean score for the third rater factor *training* is M = 3,23 with a standard deviation of 0,45. Of all eleven factors this is the lowest standard deviation found. Comparison of the mean values of the individual items showed that respondents evaluated item six (“My appraisal skills are regularly refreshed and updated through training, M = 2, 47) worse than all other items of this factor (ranging from M = 3,00 to M = 3,87).

In summary, it became obvious that the total mean scores of the factors are generally very close to each other on the middle of the scale. The factors which received the highest mean scores are *360-degree appraisal*, *procedural justice*, *goal-setting* and *performance feedback*. The lowest mean score received clearly the factor *performance-based pay*. The standard deviation of the mean scores is very close to each other, too. The highest standard deviation has the factor *employee participation*; the lowest has the factor *training*. For details see table 6.1.1 and table 6.1.2.

Factor	N	Minimum	Maximum	Mean	Std. Deviation
Instrument validity	43	1,67	4,67	3,2636	,66092
Distributive justice	43	1,00	4,75	3,1047	,75835
Procedural justice	43	2,17	5,00	3,8798	,67683
Goal-setting	43	2,00	4,63	3,7907	,58358
Performance feedback	43	1,00	5,00	3,7442	,76146
Performance-based pay	43	1,00	4,33	2,8915	,75845
Employee participation	43	1,00	5,00	3,4729	,80748
360-degree appraisal	43	2,00	5,00	3,9147	,70182
Rating techniques	15	1,20	4,40	3,1733	,75920
Rating accuracy	15	2,67	5,00	3,6000	,61978
Training	15	2,25	4,00	3,2333	,44788

6.1.1.1 Descriptive statistics of factors

Item	Mean	Std. Dev.	Item	Mean	Std. Dev.	Item	Mean	Std. Dev.
IV_1	3,35	,870	GS_3	4,16	,785	EP_3	3,44	,908
IV_2	3,44	,765	GS_4	4,05	,722	DA_1	3,49	1,077
IV_3	3,07	,936	GS_5	4,16	,998	DA_3	3,74	,790
IV_5	3,44	1,053	GS_6	3,81	,880	DA_4	4,07	,799
IV_6	3,65	1,021	GS_7	3,60	,903	DA_5	4,19	,732
IV_7	2,63	1,092	GS_8	3,21	1,059	RT_1	2,93	1,163
DJ_1	3,00	,873	PF_1	3,65	,997	RT_2	2,80	1,265
DJ_2	3,44	1,076	PF_3	3,74	,902	RT_3	3,67	,900
DJ_3	3,02	1,080	PF_4	3,70	,860	RT_4	3,27	1,033
DJ_4	2,95	,925	PF_5	3,74	1,002	RT_5	3,20	,941
PJ_1	3,56	,854	PF_6	3,98	,913	RA_1	4,13	,640
PJ_2	3,58	,906	PF_7	3,65	,813	RA_3	3,47	1,060
PJ_3	4,07	1,009	PBP_1	3,02	1,123	RA_4	3,20	,862
PJ_4	4,07	,768	PBP_2	2,93	,884	T_2	3,00	1,195
PJ_6	4,12	,931	PBP_3	2,72	,882	T_3	3,87	,915
PJ_7	3,88	,905	EP_1	3,33	1,107	T_4	3,87	,834
GS_1	3,88	,793	EP_2	3,65	,870	T_6	2,47	,915
GS_2	3,44	,854						

6.1.2 Descriptive statistics of individual items

6.2 Impact of demographic variables

This part shows the results of the independent samples t-test and linear regression. From these analyses one can interfere if demographic variables as gender, age and tenure have an impact on the perceptions of the performance appraisal system. The results are presented per demographic variable, thus first for gender, then for age and lastly for tenure.

6.2.1 Gender

The independent samples t-test was conducted to test if gender has an influence on how the performance appraisal system is perceived. No significant difference was found in the judgments of female and male respondents for all eleven factors. For details see table 6.2.1.

Factor	t	Sig. (2-tailed)
Instrument validity	,967	,339
Distributive justice	,094	,926
Procedural justice	1,160	,253
Goal-setting	-,483	,632
Performance feedback	1,084	,285
Performance-based pay	-,295	,769
Employee participation	-,650	,519
360-degree appraisal	-1,213	,232
Rating technique	-,684	,506
Rating accuracy	,061	,953
Training	-,978	,346

6.2.1 Results of independent sample test

6.2.2 Age

A linear regression analyses was conducted to test if age has an impact on how the factors of the performance appraisal system are perceived. Five of the eleven factors are slightly influenced by the age of the respondent: (1) Age accounts for 24% of the variance in opinion about *instrument validity* ($r^2 = 0,235$; $F = 7,262$, $t = 2,695$, $p = 0,001$). (2) Concerning *distributive justice* age also accounts for 24% of the variance in opinion ($r^2 = 0,238$; $F = 12,812$, $t = 3,579$, $p = 0,001$). (3) 10% of the variance in opinion about *goal-setting* ($r^2 = 0,102$; $F = 4,646$, $t = 2,155$, $p = 0,037$) can be explained by age. (4) Age accounts for 27% of the variance in opinion about *rating technique* ($r^2 = 0,265$; $F = 4,963$, $t = 2,166$, $p = 0,049$). (5) It also explains 38% of the variance in opinion about *rating accuracy* ($r^2 = 0,376$; $F = 7,835$, $t = 2,799$, $p = 0,015$). This means that the older the user is the better he perceives the factors to be.

For the other factors, *procedural justice* ($r^2 = 0,051$; $F = 2,198$, $t = 1,483$, $p = 0,146$), *performance feedback* ($r^2 = 0,014$; $F = 0,567$, $t = 0,753$, $p = 0,456$), *performance-based pay* ($r^2 = 0,013$; $F = 0,544$, $t = 0,737$, $p = 0,465$), *employee participation* ($r^2 = 0,049$; $F = 2,131$, $t = 1,460$, $p = 0,152$), *360-degree appraisal* ($r^2 = 0,022$; $F = 0,917$, $t = 0,958$, $p = 0,344$) and *training* ($r^2 = 0,246$; $F = 4,247$, $t = 2,061$, $p = 0,060$), no such influence could be found. For details see table 6.2.2.1.

Model		t	Sig.
Instrument validity	(Constant)	5,295	,000
	age	2,695	,010
Distributive justice	Constant)	3,447	,001
	age	3,579	,001
Procedural justice	Constant)	7,253	,000
	age	1,483	,146
Goal-setting	Constant)	8,032	,000
	age	2,155	,037
Performance feedback	Constant)	6,584	,000
	age	,753	,456
Performance-based pay	Constant)	5,613	,000
	age	,737	,465
Employee participation	Constant)	5,099	,000
	age	1,460	,152
360-degree appraisal	Constant)	8,859	,000
	age	,958	,344
Rating technique	Constant)	1,758	,102
	age	2,166	,049
Rating accuracy	Constant)	3,107	,008
	age	2,799	,015
Training	Constant)	1,885	,082
	age	2,061	,060

6.2.2.1 Coefficients of linear regression with age

6.2.3 Tenure

A linear regression analyses was conducted to test if tenure in organisation has an impact on how the factors of the performance appraisal system are perceived. But tenure does not significantly account for the variance in opinion about the eleven factors: *instrument validity* ($r^2 = 0,013$; $F = 0,531$, $p = 0,470$), *distributive justice* ($r^2 = 0,060$; $F = 2,626$, $p = 0,113$), *procedural justice* ($r^2 = 0,006$; $F = 0,233$, $p = 0,632$), *goal-setting* ($r^2 = 0,005$; $F = 0,207$, $p = 0,652$), *performance feedback* ($r^2 = 0,004$; $F = 0,182$, $p = 0,672$), *performance-based pay* ($r^2 = 0,005$; $F = 0,194$, $p = 0,662$), *employee participation* ($r^2 = 0,015$; $F = 0,610$, $p = 0,439$), *360-degree appraisal* ($r^2 = 0,000$; $F = 0,004$, $p = 0,953$), *rating technique* ($r^2 = 0,002$; $F =$

0,023, $p = 0,881$), *rating accuracy* ($r^2 = 0,002$; $F = 0,021$, $p = 0,886$) and *training* ($r^2 = 0,003$; $F = 0,045$, $p = 0,835$). For details see table 6.2.3.1.

Model	F	Sig.
Instrument validity	,531	,470
Distributive justice	2,626	,113
Procedural justice	,233	,632
Goal-setting	,207	,652
Performance feedback	,182	,672
Performance-based pay	,194	,662
Employee participation	,610	,439
360-degree appraisal	,004	,953
Rating technique	,023	,881
Rating accuracy	,021	,886
Training	,045	,835

a Predictors: (Constant), tenure
6.2.3.1 ANOVA of linear regression with tenure

6.3 Open question

The thirteen suggestions resulting from the “open question” part of the questionnaire mainly had two big issues: how goals are set and frequency of feedback. Concerning the goals that have to be set, respondents see the necessity to define them in a more measurable manner. They hope that by doing so the evaluation of performance becomes more objective: “The main focus of the appraisal in my experience has been future target-setting and the analysis of success / failure to meet previous targets. If those targets are set objectively then everyone can tell how well they are doing throughout the year and know when they have under- or over-performed”. Furthermore, they want the goals to be more relevant for the individual job. One respondent wrote: “Goals to which an employee is not able to contribute should not be taken into account e.g. headcount reduction or organisational win”.

A theme related to the frequency of feedback is the fast changing nature of the job and accordingly reasonable goals. As one person put it: “Often the job changes so much the goals set are obsolete.” But respondents propose a clear solution how goals can be controlled steadily for reasonableness: They would appreciate to receive more and regularly feedback, for example every three month. ” In my opinion only once a year is insufficient. It would be better to discuss the performance goals and evaluation quarterly, at least a feedback.” According to their opinion this could serve another objective besides keeping goals relevant: “That way the appraisal would be fairer and it would mean that it would cover the whole year better. Rather than at the moment as the appraisal is only annual. A good performance at the beginning of the year could be overshadowed by a bad performance at the end of the year (or vice versa). And it is human nature to concentrate more on recent events and so grades can sometimes be not as accurate as they should be.”

The results of the questioning can thus be summarised as follows: The factors which are perceived best are *360-degree appraisal*, *procedural justice*, *goal-setting* and *performance feedback*. The factor *performance-based pay* received the worst score. The only demographic variable which influences the perceptions of users regarding the performance appraisal system is age. *Instrument validity*, *distributive justice*, *goal-setting*, *rating technique* and *rating accuracy* are evaluated better the older the user is. Gender and tenure have no influence on perceptions. This is possibly due to the small amount of respondents. Larger groups of respondents have smaller variance in the mean score and receive thus more powerful results.

From the open question it results that respondents are concerned with the objectivity and relevance of goals. Furthermore, they would appreciate to receive feedback more often. In the next section the results are discussed.

7. Discussion

First of all, it is quite striking that the evaluations of the factors are positioned around the value of three which actually means a neutral assessment. Thus, although the theoretical evaluation led to expect high scores, these assumptions did not realise. Instead, all factors are more or less evaluated the same. It is thus in a way a pity that no factor yielded significant high scores, but on the other hand this result also means that no factor is judged totally bad. But besides reflecting the true score, a possible alternative explanation for this could be that respondents were, due to the five point Likert scale, prone to the central tendency error. To test this, the questioning should be repeated with for example a six-point Likert scale. If the values should represent the true score, interventions should be developed to enhance the overall satisfaction with the performance appraisal system.

Although no high total factor scores are yielded in the region of four or five, there are well differences detectable in the evaluation. The factors *360-degree appraisal*, *procedural justice*, *goal-setting* and *performance feedback* were judged higher than all other factors.

Concerning *360-degree appraisal* the results show that ratees would prefer to be evaluated by sources as peers and customers because they believe that this would enhance the accuracy of the appraisal and would provide them with valuable information about their performance.

The relative high score of *procedural justice* indicates that users perceive the appraisal process as being fair which a main predictor of performance appraisal acceptance is. The acceptance of the performance appraisal in turn is a necessary condition to elicit behavioural adaptation. Unfortunately, the evaluation of *distributive justice* which is the second important determinant for acceptance was not perceived that high. This reflects the opinion that the results of the performance appraisal lack fairness.

The third factor which received higher scores is *goal-setting*. This result shows that raters are able to clearly set up goals for employees. Furthermore, the findings from literature are supported: Employees are generally satisfied with goal-setting if the goals set are specified and accepted. From the open question can be concluded that raters sometimes lack the ability to develop goals which are objective and relevant for the employee in question and his job. Goals that cannot be controlled or influenced by the ratee are at best meaningless, but have the potential to harm motivation.

A related theme is *performance feedback*. Because the performance review can interfere with the self-image, it is especially important to present the feedback in a way that it can be accepted by the employee. This seems to be the case in the target company. But results of the open question revealed that employees would prefer to receive feedback more often. As proposed in literature it could be integrated into the daily interactions of supervisors and subordinates by meeting more frequently but less formal.

The factor *performance-based pay* received the lowest score. This is in line with the expectations which followed from the theoretical evaluation. Possible reasons for this result are various: First, it might be necessary that people get used to the system. The change probably resulted in fears that they could be disadvantaged by the new system. If the performance-based pay system is in use for a while, it could turn out that fears were unreasoned and users get as a consequence more satisfied. Next, it is questionable if the amount of pay employees can earn is effective in motivating. Maybe the limits for the bonus must be raised. Another point possibly responsible for the low score is that the bonus pool is defined by many different instances and is thus influenced by circumstances that are out of control for the individual employee. So it is likely that ratees experience fuzziness regarding process and lose motivation because of their lack of control. To find out which reason is actually caused the low score it is necessary to conduct a new questioning aiming at the

problems perceived with performance-based pay and how employees imagine an efficient link between performance and pay.

Another important finding is that demographic variables nearly do not influence the perception of performance appraisal users in the target company. The only impact was found for *instrument validity*, *distributive justice*, *goal-setting*, *rating technique* and *rating accuracy* which are evaluated the better the older the user is. Maybe the reason for the influence of age is that older employees take the performance appraisal system more for granted. Younger employees might still have the motivation to critically investigate and evaluate things. With age and experience people might learn that in organisations the individual opinion does not really make a change and come to accept things as they are.

8. Conclusions and implications

What became clear is that the performance appraisal system of the target company is not perceived as bad as it could have been, but unfortunately also not as good as it could have been. The effectiveness of the performance appraisal system is as a consequence probably limited. Thus, the results revealed at least that there is room for improvement in all factors. Furthermore, the results provide some ideas how these improvements could look like:

First of all, it should be decided if the integration of more appraisal sources could be useful. According to the findings peers and customers are seen as valuable information sources. How these sources are integrated into the appraisal process, for example if their rating should influence performance-based pay results, should be subject to further investigation. It is important to always keep in mind that the acceptance of users is the key determinant for performance appraisal effectiveness.

A second implication which is relatively easy to implement should be that raters receive more or better training in the development of objective and relevant objectives. What the exact reasons for this problem are must be analyzed by further research. The training material and-task should be investigated as well as the amount of time which is spent with training. At the moment, this part of the performance appraisal process is only one of three learning objectives for the first training day.

The third idea for improvement concerns the frequency of feedback. Informal, each rater can also easily implement more feedback. A solution could be that raters integrate so called achievement updates on a weekly basis which then touch upon good and bad issues, while so called achievement assessments take place bi-monthly, are more formal and aim at getting a more clear depiction of issues troubling both sides. More formally, it could be decided if quarterly a short performance review with objective updating is integrated into the performance appraisal process.

But the most important implication is that the performance-based pay system and the opinion of employees about it must be analysed urgently. Then of course the system needs to be adjusted according to the results. Otherwise the target organisation runs risk in the long term to destroy all positive motivational effects that such a system can have if it is well designed.

It might be important, too, that the target organisation eliminates the impact of age. This could for example be done by giving employees some form of performance appraisal training as they enter the organisation. This training must not be as detailed as for raters but should inform about the goals, the process and introduce the instrument. By doing this, the target organisation would make use of the chance that a completely new understanding and support for the performance appraisal system could grow from bottom upwards.

Furthermore, it should be thought about implementing regular feedback-checks with regard to performance appraisal. This is especially important before developing and implementing a new component. Otherwise one learns not until it is in praxis that money and time spend were actually useless. Such negative after-effects could be overcome with the help of surveys, good ideas and suggestions could be gathered and employees feel that the target organisation attaches great importance to their opinion. Thus, the quality of the performance appraisal and the acceptance of its users could easily be improved.

The study has two major shortcomings which limit the generalisability of the results: First, results may be specific for the department of the target organisation where the questioning was conducted. This department has a quite specific task and employees there often have a bank-untypical educational background. Thus, it is possible that the results are not representative for whole bank. Second, especially the raters' results may not be taken too serious because of the small amount of respondents. 15 respondents are far from being a reliable group.

9. Reflections

In this part some reflections about the past month are given room. Three issues are touched here: working alone, support of the teachers and contact with target organisation.

9.1 Working alone

The bachelor thesis was actually the first time during my study at University of Twente that I had to work totally on my own. Fostering the ability to work in groups seems to be one of prioritised objectives of this university. Thus, whenever possible, group projects are integrated into the content of each subject. In regard to the reality of working life where groups or teams are the major working unit it is probably very wise to staff students with the needed skills and to give them a prospect of their possible future working conditions.

But although I think of myself as being an adequate “team player” and actually enjoy working with peers, I was really pleased to be able to work on my own this time. Reasons for this are quite different:

First of all, in group projects it is much easier to avoid challenging oneself. Instead, one chooses tasks or gets assigned to tasks which one can handle competently. I for example usually did literature search and –summary because I enjoy this kind of work and probably am (due to the large amount of the training meanwhile) quite good at it. Of course this tactic has the goal to achieve the best possible results. But on the other hand it disturbs self-development in certain aspects of research. The statistic subjects are already some time ago and because of the delegation of task I did not have to deal with it either. As a consequence I thus really missed some skills in this section. But when working alone one has no possibility of avoiding the uncomfortable aspects of work.

Second, I have made the experience that group results are often worse than individual results. According to my opinion, the reason for this is not that group performance is the averaged performance of all group members. I believe that group processes as planning, coordination and information flow can have a negative impact on group output which overshadows the potential group performance. It is thus uncertain what the teacher actually wants to judge in a group project: the performance on subject-related knowledge and skills or the performance on group processes.

The third reason touches the issue of time. There is no doubt that it saves time to do the work with many people instead of doing it alone. But I have the feeling that a lot of time gets lost during group work because one has to meet to discuss, decide, share information etc. And it is not unusual that these meetings are inefficient because one talks about personal stuff or discusses the same things several times all over again. If one works alone all the deciding, discussing and information processing happens of course intra-individual. Thus, the time necessary to make and have appointments or to communicate and decide with others can directly be used to convert ideas into action. For me, this was thus the first time of time-efficient working.

Thus, all in all, in spite of all the advantages that group work may have, I really enjoyed the experience of working alone. Even though I had to face some problems for example the statistical analyses I felt it was a challenge. And not to forget, the whole bachelor thesis, from the beginning to its end, is the result of my own work. That is a great experience.

9.2 Support of teachers

The conditions to support me might not have been the best from both sides. From my side two major circumstances were problematic: First, as a German student I am living several kilometres away. Thus, my presence at the university is naturally limited. Second, the organisation where my research took place is even more far away. But from the teachers' side arose some disturbing aspects, too. For example, my first teacher was several weeks abroad,

while the other unfortunately had an accident and was sick at home. So, in comparison to another student who was literally working next door to the teachers, personal contact or visits were in a way limited.

But even though the circumstances were far from being perfect, I really felt such a friendly support by my teachers during the whole project. To communicate constantly and in case of problems we used e-mail. Although this tool combined with my language problems made it necessary to write more often or detailed what exactly one was talking about, we always managed to reach consensus. We also had meetings occasionally. Not too often, but well pointed at the most important points of this project. I believe I learned a lot useful things to organise such project. One good example is designing a plan with all the different steps that have to be done and the time necessary to fulfil the steps. Such a plan gives you a great guidance and reminds you of keeping working.

But the most significant moment was from my point of view as my first teacher asked if my contact person in the target organisation would be interested in meeting him. I never imagined that such an engagement is possible. This might stem from the stories I heard from friends studying at German universities. There it is sometimes even impossible to arrange an appointment with a teacher at his own office. Thus, I was really pleased that it was even possible that my teacher came to visit me in the target organisation. For all the help and support I really would like to thank them very much!

9.3 Contact with target organisation

There is no doubt about that I was in a very lucky position concerning the target organisation. The fact that I am working there beside my study makes so many things, that otherwise might be complicated, easy: I know the processes at the target organisation and, maybe most important, I know the people working there.

As from my teachers I also received extraordinary support from my supervisor and my colleagues. My supervisor literally fought for me so that I got the admission to do the research there. Without him, the whole project could not have been taken place as it did now. But my direct colleagues also showed engagement which researchers normally will not find. Some of them, for example, have sent the questionnaire to other be friend colleagues which I did not know.

But beside these positive aspects I again had to learn what it means to ask people for their participation in research, too. The people who actually did take part were more or less my direct colleagues, thus the people working on the same floor. If I was not personally known it would have been much harder to convince people to take part. Another issue was that although someone assured you participation it does not mean that you will receive an answered questionnaire. Because they are at work, people put the questionnaire away waiting for some better point of time to do it. But with some people, this point does not seem to come ever. After some days running around asking "Is the questionnaire ready?" I found out that it is probably easier to stand and wait directly next to the potential participant until he/she answered the questionnaire or to make a particular appointment.

What made it even more complicated for participants was the fact that the questionnaire was in English. Because the whole bachelor thesis is in English and I also wanted to reach some colleagues from London, I thought it would be clever to design only one questionnaire which is accordingly in English, too. But although my German colleagues are used to speak English, they use a totally different terminology. Thus, most of them had such big problems is answering the questionnaire that I needed to help them. Instead of a time-efficient questionnaire which was planned to take about 15 minutes, there was far more often a interview at hand taking about 30 minutes.

Thus, I was really reminded of the point that, in order to attract participants, you should make it as easy as possible for them to take part. First, translating a questionnaire would have been much faster than helping 30 people individually. Second, maybe some people who now did not take part because they were afraid of the foreign language and unknown vocabulary would have participated if the questionnaire would have been available in German as well.

References

- Banker, R.D. et al. (2001). An empirical analysis of continuing improvements following the implementation of a performance-based compensation plan. *Journal of Accounting and Economics*, 30, 315-350
- Benson, P.G. et al. (1988). The Impact of Rating Scale Format on rater Accuracy: An Evaluation of the Mixed Standard Scale. *Journal of Management*, 14 (3), 415-423
- Berry, L.M. (2003). Employee Selection. Thomson Wadsworth
- Cook, M. (1995). Performance appraisal and true performance. *Journal of Managerial Psychology*, 10 (7), 3-7
- Cox, A. (2000). The importance of employee participation in determining pay system effectiveness. *International Journal of Management Reviews*, 2 (4), 357-375
- Dalton, M. (1996). Multirater Feedback and Conditions for Change. *Consulting Psychology Journal*, 48 (1), 12-16
- Davis, T. & Landa, M. (1999). A Contrary Look at performance Appraisal. *Canadian Manager/Manager Canadian*, Fall, 18-28
- De Silva, S. (1998). Performance-Related and Skill-Based Pay: An Introduction. International Labour Office, Geneva
- Facteau, J.D. & DeVries, S.B.C.K. (2001). Are Performance Appraisal Ratings From Different Rating Sources Comparable?. *Journal of Applied Psychology*, 86 (2), 215-227
- Gabris, G.T. & Ihrke, D.M. (2001). Does Performance Appraisal Contribute to Heightened Levels of Employee Burnout? The Results of One Study. *Public Personnel Management*, 30 (2), 157-172
- Härtel, C.E.J. (1993). Rating Format Research Revisited: Format Effectiveness and Acceptability depend on Rater Characteristics. *Journal of Applied Psychology*, 78 (2), 212- 217
- Hamman, C.W.R. et al. (1999). Evaluating instructor/evaluator inter-rater reliability from performance database information. Proceedings of the Tenth International Symposium on Aviation Psychology
- Harrell, A. & Wright, A. (1990). Empirical Evidence on the Validity and Reliability of Behaviourally Anchored rating Scales for Auditors. *A Journal of Practice & Theory*, 9 (3), 134-149
- Harris, C. (1988). A Comparison of Employee Attitude Toward Two Performance Appraisal Systems. *Public Personnel Management*, 17 (4), 443-456
- Hedge, J.W. & Kavanagh, M.J. (1988). Improving the Accuracy of Performance Evaluations: Comparison of Three Methods of Performance Appraiser Training. *Journal of Applied Psychology*, 73 (1), 63-73
- Jacobs, R. et al. (1980). Expectations of Behaviourally Anchored Rating Scales. *Personnel Psychology*, 33, 595-640
- Jelley, R.B. & Goffin, R.D. (2001). Can Performance-Feedback Accuracy Be Improved? Effects of Rater Priming and Rating Scale Format on Rating Accuracy. *Journal of Applied Psychology*, 86 (1), 134-144
- Kane, J.S. & Bernardin, H.J. (1982). Behavioural Observation Scales and the Evaluation of performance Appraisal Effectiveness. *Personnel Psychology*, 35, 635-641
- Kinicki, A.J. et al. (1995). Effects of Category Prototypes on Performance-Rating Accuracy. *Journal of Applied Psychology*, 80 (3), 354-370
- Kondrasuk, J.N. et al. (2002). An Elusive Panacea: The Ideal Performance Appraisal. *Journal of Managerial Psychology*, 64 (2), 15-31
- Latham, G.P. & Wexley, K.N. (1977). Behavioural Observation Scales For Performance Appraisal Purposes. *Personnel Psychology*, 30, 255-268

- Lee, C. (1983). Increasing Performance Appraisal Effectiveness: Matching Task Types, Appraisal Process, and Rater Training. *Academy of Management Review*, 10 (2), 322-331
- Martell, R.F. & Borg, M.R. (1993). A Comparison of the Behavioural Rating Accuracy of Groups and Individuals. *Journal of Applied Psychology*, 78 (1), 43-50
- Martell, R.F. & Leavitt, K.N. (2002). Reducing the Performance – Cue Bias in Work behaviour Ratings: Can Groups Help?. *Journal of Applied Psychology*, 87 (6), 1032-1041
- Maurer, T.J. & Raju, N.S. (1998). Peer and Subordinate Performance Appraisal Measurement. *Journal of Applied Psychology*, 83 (5), 693-702
- Parrill, S. (1999). Revisiting Rating Format Research: Computer-Based Rating Formats and Components of Accuracy. Unpublished manuscript. Virginia Polytechnic and State Institute, Blacksburg, VA.
- Rarick, C.A. & Baxter, G. (1986). Behaviorally anchored rating scale (Bars): An effective performance appraisal approach. *Sam Advanced Management Journal*, winter, 36-39
- Reinke, S.J. (2003). Does The Form Really Matter? Leadership, trust, and acceptance of the Performance Appraisal Process. *Review of Public Personnel Administration*, Vol. 23, No. 1, 23-37
- Ridge, K. (2001). Rater Halo Error And Accuracy In A Mathematics Performance Assessment. National council on measurement in education, 11-13 April
- Roberts, G.E. (2003). Employee Performance Appraisal System Participation: a Technique that Works. *Public Personnel Management*, 30 (1), 89-98
- 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
- Rotundo, M. & Rotman, J.L. (2002). Defining and Measuring Individual Level Job Performance: A Review and Integration. *Journal of Applied Psychology*, 90 (5), 225-254
- Rudner, L.M. (). Reducing Errors Due to the Use of Judges. Practical Assessment, *Research & Evaluation*, 7 (26), 241-271
- Sanchez, J.I. & De La Torre, P. (1996). A Second Look at the Relationship between Rating and Behavioural Accuracy in Performance Appraisal. *Journal of Applied Psychology*, 81 (1), 3-10
- Smith, R. (2001).] Favouritism, Bias, and error in performance ratings of scientists and engineers: the effects of power, status, and numbers. *Sex Roles*, 45 (5-6), 337-358
- Solmon, L.C. & Podgursky, M. (2000). The Pros and Cons of Performance-based Compensation. Santa Monica, CA: The Milken Family Foundation
- Stark, E. et al. (1998). A Field Study of Accuracy, Agreement, and Outcomes of Multi-Source Feedback: Implications Across International Boundaries. Budapest, Hungary: International Conference of the Academy of Business and Administrative Sciences
- Swiercz, P.M. et al. (1999). Do Perceptions Of Performance Appraisal Fairness Predict Employee Attitudes And Performance?. *Academy of Management Proceedings*, 304
- Tziner, A. et al. (1992). Effects of Performance Appraisal Format on Perceived Goal Characteristics, Appraisal Process satisfaction, and Changes in Rated Job Performance: A Field Experiment. *The Journal of Psychology*, 127 (3), 281-291
- Tziner, A. & Kopelman, R.E. (2002). Is there a Preferred Performance Rating Format? A Non-psychometric Perspective. *Applied Psychology: An International Review*, 51 (3), 479-503

- Wise, P.G. (1998). Rating differences in Multi-Rater Feedback. Chicago, IL: International Personnel Management Association Assessment Council's Conference on Professional Personnel Assessment
- Woehr, D.J. (1997). Distributional ratings of performance: more evidence for a new rating format. *Journal of Management*. 78, 126-149
- Zigon, J. (1995). How to Measure the Results of Work Teams. Wallingford, PA: Zigon Performance Group

Appendix A: Performance appraisal form 2004, Germany



1.1.a Form 2004 Germany

Planning and Feedback

Name, Surname:
Personnel Number:
Overall Performance Assessment:

A. Planning and Feedback

1.1.a

Surname, first name	Personnel number	No. of organisational unit
<input type="text"/>	<input type="text"/>	<input type="text"/>
Date of joining department/organisational unit	Position	
<input type="text"/>	<input type="text"/>	

Surname and first name of superior	Personnel number	No. of organisational unit
<input type="text"/>	<input type="text"/>	<input type="text"/>
Surname and first name of next higher superior	Reason for feedback	
<input type="text"/>	<input type="text"/>	

To be completed at start of planning period	To be completed at end of planning period
<p>Task planning for the period from <input type="text"/> to <input type="text"/></p> <p>Key tasks (obligatory for all employees)</p> <p>Agreed business objectives (obligatory only for management staff and senior specialists/relationship or product managers)</p>	<p>1.1.b Evaluation of results for the period from <input type="text"/> to <input type="text"/></p>
Individual specifications and quality criteria	

Surname, first name	Personnel number	Organisational unit
<input type="text"/>	<input type="text"/>	<input type="text"/>
Date of joining department/organisational unit	Position	
<input type="text"/>	<input type="text"/>	

Surname and first name of appraiser	Personnel number	Organisational unit
<input type="text"/>	<input type="text"/>	<input type="text"/>
Surname and first name of appraiser's manager	Planning and review period	
<input type="text"/>	<input type="text"/>	

Planning and Evaluation

Achievement against objectives set

Achievement relative to objectives set, using the following scale

- (A) Did not achieve
- (B) Partially achieved
- (C) Achieved
- (D) Exceeded
- (E) Substantially exceeded

Skills Assessment

- B = Still need for improvement in individual situations
- C = Overall good performance within the parameters of requirements
- D = Very good performance, partly beyond expectations of the position

In relevant situations was observed to be...					
unsatisfactory A	in line with requirements			of outstanding performance E	not observable/ not relevant
	B	C	D		

Date, signature of superior ()

Date, signature of employee

A. Planning and Feedback

1.1.a

Surname, first name	Personnel number	No. of organisational unit
<input type="text"/>	<input type="text"/>	<input type="text"/>
Date of joining department/organisational unit	Position	
<input type="text"/>	<input type="text"/>	

Surname and first name of superior	Personnel number	No. of organisational unit
<input type="text"/>	<input type="text"/>	<input type="text"/>
Surname and first name of next higher superior	Reason for feedback	
<input type="text"/>	<input type="text"/>	

To be completed at start of planning period	To be completed at end of planning period
<p>Task planning for the period from <input type="text"/> to <input type="text"/></p> <p>Key tasks (obligatory for all employees)</p> <p>Agreed business objectives (obligatory only for management staff and senior specialists/relationship or product managers)</p>	<p>1.1.b Evaluation of results for the period from <input type="text"/> to <input type="text"/></p>
Individual specifications and quality criteria	

	... gezeigt					
1. Specialist and methodological	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	
2. Personal skills	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	
3. Interpersonal skills	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	
4. Entrepreneurial skills	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	
5. Management skills	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Summary

Overall performance assessment

Make an assessment of the appraisee's overall performance in the context of the achievement of objectives, the skills assessment and the performance standards required for the role, using the following scale:

- (A) Objective not achieved, unsatisfactory performance
- (B) Objective partially achieved, scope for improvement in individual situations
- (C) Objective achieved, good performance within the framework of the requirements
- (D) Objective exceeded, very good performance which partially exceeds requirements
- (E) Objective substantially exceeded, extraordinary performance which exceeds requirements by far
- (F)

Surname, first name	Personnel number	Organisational unit
<input type="text"/>	<input type="text"/>	<input type="text"/>
Date of joining department/organisational unit	Position	
<input type="text"/>	<input type="text"/>	

Surname and first name of appraiser	Personnel number	Organisational unit
<input type="text"/>	<input type="text"/>	<input type="text"/>
Surname and first name of appraiser's manager	Planning and review period	
<input type="text"/>	<input type="text"/>	

Date, signature of superior ()

Date, signature of employee

A. Planning and Feedback

1.1.a

Surname, first name <input type="text"/>	Personnel number <input type="text"/>	No. of organisational unit <input type="text"/>
Date of joining department/organisational unit <input type="text"/>	Position <input type="text"/>	

Surname and first name of superior <input type="text"/>	Personnel number <input type="text"/>	No. of organisational unit <input type="text"/>
Surname and first name of next higher superior <input type="text"/>	Reason for feedback <input type="text"/>	

To be completed at start of planning period		To be completed at end of planning period
<p>Task planning for the period from <input type="text"/> to <input type="text"/></p> <p>Key tasks (obligatory for all employees)</p> <p>Agreed business objectives (obligatory only for management staff and senior specialists/relationship or product managers)</p>	<p>Individual specifications and quality criteria</p>	<p>1.1.b Evaluation of results for the period from <input type="text"/> to <input type="text"/></p>

Money Laundering Regulations – Reliability

This page is not relevant for the staff and executives of the following areas:

- Board members and Group Development
- Group Communications
- Group Compliance/Money Laundering Prevention
- Economics
- Research
- Information Technology
- Group Organisation
- Internal Services
- Asset Liability Management
- Market Risk Management
- Operational Risk Management
- Group Risk
- Group Finance
- Group Controlling
- Human Resources
- Workers Council

Please indicate below whether the jobholder is involved in any way with financial transactions for clients of the WestLB Group
yes no

If yes, with regard to the German laws on the prevention of Money Laundering and the PUR guidance notes, the appraiser is confirming the following: they have no knowledge of the jobholder having been involved in any irregularities in relation to Money Laundering. In addition, the appraiser confirms that there are no obvious signs of the job holder becoming susceptible to involvement in Money Laundering (e.g. gambling addiction, extravagant spending, and involvement in the drug or terrorist scenes). They also confirm that to the best of their belief during the reporting period, the jobholder has acted in compliance with all the requirements of both the local national and German laws on money laundering prevention [, the procedures contained in manual 123] and the local manual for Money Laundering Prevention (e.g. Client identification requirements, duty to report suspicious activity, participation in training).

The above mentioned employee is reliable.
yes no

Date, signature of executive ()

Date, signature of superior ()

Date, signature of employee

A. Planning and Feedback

1.1.a

Surname, first name	Personnel number	No. of organisational unit
<input type="text"/>	<input type="text"/>	<input type="text"/>
Date of joining department/organisational unit	Position	
<input type="text"/>	<input type="text"/>	

Surname and first name of superior	Personnel number	No. of organisational unit
<input type="text"/>	<input type="text"/>	<input type="text"/>
Surname and first name of next higher superior	Reason for feedback	
<input type="text"/>	<input type="text"/>	

To be completed at start of planning period	To be completed at end of planning period
<p>Task planning for the period from <input type="text"/> to <input type="text"/></p> <p>Key tasks (obligatory for all employees)</p> <p>Agreed business objectives (obligatory only for management staff and senior specialists/relationship or product managers)</p>	<p>1.1.b Evaluation of results for the period from <input type="text"/> to <input type="text"/></p>
Individual specifications and quality criteria	
<input type="text"/>	

After planning and individual concretion of tasks both the employee and the superior receive a copy. The BU Human Resources only gets a copy when the appraisal is totally completed.

Date, signature of superior ()

Date, signature of employee

B. Summary Feedback

Surname, first name

Appraisal period
 to

In relevant situations was observed to be ...					not observ- able/ not relevant
(still) unsatis- factory A	in line with requirements B C D			out- standing performance E	
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

B = Still need for improvement in individual situations
 C = Overall good performance within the parameters of requirements
 D = Partly beyond expectations of the position

1. Specialist and methodological skills

The employee

1.1 uses well-founded and up-to-date knowledge specific to the task	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
1.2 plans and organises working procedures in such a way that tasks are completed efficiently and on time	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

2. Personal skills

The employee

2.1 is a self-starter and does not always need specific instructions	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
2.2 responds in a flexible way to new situations (ideas, tasks, persons)	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
2.3 develops new products more in line with needs or more effective solutions to problems	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
2.4 performs his/her duties even in stressful situations	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
2.5 questions his/her own working methods/results, recognises mistakes and learns from them	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
2.6 makes the necessary decisions even in difficult situations, and defends them	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
2.7 has a self-confident manner and argues his/her case convincingly	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

3. Interpersonal skills

The employee

3.1 is friendly, tolerant and fair towards others	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
3.2 communicates clearly and comprehensibly	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
3.3 passes on important information	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
3.4 establishes and maintains contacts of importance to the task	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
3.5 shows commitment beyond his/her own area of responsibility, and offers support to others	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
3.6 clearly shows his/her own viewpoint and discusses with others in order to achieve workable results	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
3.7 takes a constructive approach to criticism	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
3.8 is aware of conflict and seeks to achieve clarification	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

B. Summary Feedback

Surname, first name

Appraisal period
 to

In relevant situations was observed to be ...					
(still) unsatisfactory A	in line with requirements B C D			outstanding performance E	not observable/ not relevant
<input type="text"/>					

B = Still need for improvement in individual situations
 C = Overall good performance within the parameters of requirements
 D = Very good performance, partly beyond expectations of the position

4. Entrepreneurial skills

The employee

4.1 takes account of the effects of his/her own actions on others	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
4.2 recognises and exploits profit opportunities in accordance with bank interests	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
4.3 acts in a cost-conscious manner and uses resource cost-effectively	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
4.4 seeks to add value and deliver higher quality for internal and external customers	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
4.5 wins over and retains internal and external customers through attentiveness, interest and reliability	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
4.6 takes calculated risks in order to exploit opportunities	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

5. Management skills¹

As a manager, he/she

5.1 takes account of the strengths and weaknesses of employees in allocating tasks	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
5.2 delegates authority and responsibility to allow employees to work autonomously	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
5.3 agrees realistic objectives as well as clear benchmarks for the measurement of success, jointly with employees	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
5.4 stands by employees in critical situations	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
5.5 gives employees feedback if required in specific situations	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
5.6 promotes the employees' career development	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
5.7 promotes teamwork and creates a positive working atmosphere	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Notes on non-observable points/Supplementary comments

¹ If the employee undertook managerial duties only for a limited time, please explain the reasons for it and name the relevant period of time.

C. Personnel Development

Surname, first name

Appraisal period

 to

1. Measures within parameters of current position

Development of specialist and non-specialist skills and/or expansion of area of responsibility
(e.g. through working in other departments, project work, special assignments, seminars, etc.)

2. Conceivable medium-term change in area of duties

- Potential direction(s) of development
- Time span
- Supporting measures

The result-evaluation interview for the past planning period has taken place.

Yes No if not, why

The new task planning procedure/agreement of objectives has taken place for the next period.

Yes No if not, why

Date, signature of superior ()

I am in agreement with the feedback and the proposals on personnel development.

Yes No

if not in agreement, on what points?

Date, signature of employee

Signature of next higher superior ()

7) Raters share a common conceptualisation of the performance standards.

1 2 3 4 5
absolute disagreement neutral absolute agreement

2.2 Distributive justice

1) The performance appraisal process results in a clear and unbiased appraisal.

1 2 3 4 5
absolute disagreement neutral absolute agreement

2) The performance appraisal process results in better communication between myself and my supervisor.

1 2 3 4 5
absolute disagreement neutral absolute agreement

3) The best workers receive the highest evaluation scores.

1 2 3 4 5
absolute disagreement neutral absolute agreement

4) The performance appraisal is well designed and leads to better performance and work quality.

1 2 3 4 5
absolute disagreement neutral absolute agreement

2.3 Procedural justice

1) My supervisor possesses adequate knowledge and training to properly implement my performance evaluation.

1 2 3 4 5
absolute disagreement neutral absolute agreement

2) My supervisor utilizes the evaluation system to assess my performance objectively and without bias.

1 2 3 4 5
absolute disagreement neutral absolute agreement

3) If I have problems with my performance evaluation I can communicate my concerns openly to my supervisor.

1 2 3 4 5
absolute disagreement neutral absolute agreement

4) My supervisor will be ethical in how he/she scores my performance.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

5) I know the standards used to evaluate my performance.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

6) My supervisor takes the performance appraisal procedure seriously.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

7) I feel treated fairly during the performance appraisal process.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

2.4 Goal-setting

1) My supervisor clearly expresses goals and assignments.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

2) The goals developed for my performance period are meaningful measures.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

3) I accept the goals I have been assigned.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

4) I agree that the performance goals set up for me are reasonable.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

5) I am determined to achieve my performance goals.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

6) It is clear to me which course of action I need to take in order to accomplish my performance goals.

1 2 3 4 5
absolute disagreement neutral absolute agreement

7) My supervisor allows me to help choose the goals that I am to achieve.

1 2 3 4 5
absolute disagreement neutral absolute agreement

8) Goal-setting gives me a broader picture of the work unit and the organisations objectives.

1 2 3 4 5
absolute disagreement neutral absolute agreement

2.5 Performance feedback

1) The performance feedback I receive is helpful in improving my on-the –job performance and in attaining my goals.

1 2 3 4 5
absolute disagreement neutral absolute agreement

2) I receive regular and timely performance feedback beside the annual performance review.

1 2 3 4 5
absolute disagreement neutral absolute agreement

3) The information provided by my supervisor during my performance feedback is sufficiently lucid.

1 2 3 4 5
absolute disagreement neutral absolute agreement

4) The information provided by my supervisor during my performance feedback is sufficiently detailed.

1 2 3 4 5
absolute disagreement neutral absolute agreement

5) I am satisfied with my performance feedback.

1 2 3 4 5
absolute disagreement neutral absolute agreement

6) Throughout my performance feedback I feel that I have the possibility to discuss work-related issues with the supervisor.

1 2 3 4 5
absolute disagreement neutral absolute agreement

7) Level of involvement in my performance evaluation is adequate.

1 2 3 4 5
absolute disagreement neutral absolute agreement

2.6 Performance-based pay

1) There is a clear, direct and compelling linkage between performance and pay in the performance appraisal system.

1 2 3 4 5
absolute disagreement neutral absolute agreement

2) The forming of a global evaluation is clear and valid.

1 2 3 4 5
absolute disagreement neutral absolute agreement

3) There is a clear and reasonable process established for grieving both evaluation and performance-based pay results.

1 2 3 4 5
absolute disagreement neutral absolute agreement

4) I believe the amount of performance-based pay I can earn through high evaluation ratings will make a noticeable difference in my future performance.

1 2 3 4 5
absolute disagreement neutral absolute agreement

5) Performance-based pay based on performance ratings is the most effective method for motivating employees to improve/sustain performance.

1 2 3 4 5
absolute disagreement neutral absolute agreement

2.7 Employee participation

1) I would be willing to participate in developing a new performance appraisal system.

1 2 3 4 5
absolute disagreement neutral absolute agreement

- 2) **Participation of employees in the development of performance standards leads to a better performance appraisal instrument.**

1 2 3 4 5
absolute disagreement neutral absolute agreement

- 3) **I would prefer my performance to be evaluated by an instrument developed and designed with the help of employees.**

1 2 3 4 5
absolute disagreement neutral absolute agreement

2.8 “360-degree” appraisal

- 1) **I would prefer my performance to be evaluated additionally by my colleagues because they have relevant performance information and insight.**

1 2 3 4 5
absolute disagreement neutral absolute agreement

- 2) **I would like to give my supervisor feedback on his/her performance.**

1 2 3 4 5
absolute disagreement neutral absolute agreement

- 3) **Appraising my own performance would enhance my awareness of my performance.**

1 2 3 4 5
absolute disagreement neutral absolute agreement

- 4) **Being appraised by several sources (supervisor, peers, customers etc) would enhance the accuracy of performance appraisals.**

1 2 3 4 5
absolute disagreement neutral absolute agreement

- 5) **Being appraised by several sources (supervisor, peers, customers etc) would provide me with valuable information about different important aspects of my performance.**

1 2 3 4 5
absolute disagreement neutral absolute agreement

3. Part “Open question”:

- 1) **Do you have any ideas how to improve performance appraisals?**

4. Part: “*Statements raters*” (for raters only!)

4.1 Rating techniques

1) I feel comfortable with the scales used to evaluate performance.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

2) I feel that the scales allow an accurate assessment of different dimensions of performance.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

3) The existing form is too complex.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

4) The existing form is too long.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

5) The existing form is easy to use.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

4.2 Rating accuracy

1) I am motivated to correctly evaluate employees' behaviour.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

2) I regularly record incidents of good/poor behaviour relevant for the performance evaluation of employees.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

3) I have sufficient opportunity to observe the employees I evaluate.

1	2	3	4	5
absolute disagreement		neutral		absolute agreement

4) I feel I have enough information regarding performance standards to make accurate judgements about employees on each performance dimension.

1 2 3 4 5
absolute disagreement neutral absolute agreement

4.3 Training

1) I need more training in conducting performance appraisal interviews.

1 2 3 4 5
absolute disagreement neutral absolute agreement

2) I am sufficiently trained in all skills needed in appraising performance.

1 2 3 4 5
absolute disagreement neutral absolute agreement

3) I am able to give useful feedback.

1 2 3 4 5
absolute disagreement neutral absolute agreement

4) I am able to clearly set goals that are relevant for the employee's position.

1 2 3 4 5
absolute disagreement neutral absolute agreement

5) I am able to use the appraisal instrument as intended.

1 2 3 4 5
absolute disagreement neutral absolute agreement

6) My appraisal skills are regularly refreshed and updated through training.

1 2 3 4 5
absolute disagreement neutral absolute agreement