



Meet up with your manager more often

An experiment on the positive influence of
face-to-face contact on performance appraisal

Master's thesis
Erasmus Universiteit Rotterdam
Department of Business Economics
Section Accounting, Auditing and Control
Thesis supervisor: prof. dr. V.S. (Victor) Maas
2nd supervisor: prof. dr. E.A. De Groot

Author: Sousan Tam
Student number: 192561st
Rotterdam, April 16th, 2012

Table of contents

Introduction	2
Literature review	9
Hypotheses development	13
Method	17
Results	19
Conclusion	28
References	31

Introduction

This thesis looks into the possible impact of telecommuting on employees' performance appraisals. I describe the research questions of this thesis. Furthermore the current state of academic knowledge and the method I have used are described. Then the main empiric findings and overall contribution to the study are set out. And lastly the structure of this thesis is described.

This paragraph explains what telecommuting is and gives some statistics. I describe the circumstances which makes telecommuting blossom even more and how telecommuting has begun. I also describe some advantages and disadvantages to employers and as well to employees.

Telecommuting or telework are interchangeable terms that refer to a work arrangement in which employees enjoy flexibility in work location and hours. Telecommuters mostly communicate through electronic means for example by e-mail, intranet, company chat or by telephone (Kirk and Belovics, 2006). These employees generally have the freedom to avoid the daily commute to a central place of work and choose their own remote location. Sometimes they also have the opportunity to choose when they perform their work. The working mother, who picks up her children at 3 o'clock and continues her work in the evening when her children are asleep, is an example.

According to Statistical Indicators Benchmarking the Information Society (SIBIS) of the European Commission, telecommuting should be as high as 10.1% in the Netherlands. According to Van Tilburg (2007) the Netherlands has 1 million telecommuters and is the number one country in Europe with most people working at home. Based on SIBIS research (SIBIS, 2003) the number of teleworkers employed in EU member states was 13% on the average. In the USA the same index showed a level as high as 25%. The SIBIS research projects an annual 20-30% growth (Forgacs, 2010). In 2009, there were 2.9 million employees who primarily worked from home in the USA, an increase of 61 percent since 2005 (U.S. Census Bureau's annual American Community Survey, 2009). Most of that growth came from

teleworking federal and state workers. The number may vary according how the telecommuter is defined. But the number of telecommuters estimated over the world could be as high as 33.7 million if you include employees, contractors and business owners who work from home at least once a month (WorldatWork, 2009).

According to Van Meel (2010) there were several influential writers of management articles and academics such as Drucker, Toffler and Bell who proclaimed the rise of a different era in the 1970's where knowledge work and information technology would become much more important and prominent than factory production. Telecommuting is the proof that this development has become a reality and it has been made possible through different technological advancements like the first personal computers in the 1970s. Since the 1980s computers came within reach of a wider general public in developed countries and businesses have become more and more reliant on computers. By the end of the 1980s, having basic computer knowledge and skills became obligatory for many jobs. In the 1980s the first mobile phones were created, further establishing the digital era. In 1992 the World Wide Web was released to the public and many businesses started using it to place their ads and business websites. In 2000 half of the households in the United States used the internet on a regular basis on their personal computers. Also in that time cell phones became very common (Wikipedia Digital Revolution, 2012).

Mobile and flexible working practices are not very new and even started in the 1960s and 1970s. Nilles (1998) came up with the idea of telecommuting while he got stuck in a traffic jam. He was not the only one who got stuck in commuting and started to think about solutions for this problem which is very time-consuming and costly. He started a feasibility-study for an insurance company in Los Angeles for distance working. The firm was located in an overcrowded business district of Los Angeles. Also the firm could not find employees they wanted to have due to the ageing population and were forced to attract labor force who were living further away from the office. These circumstances made the firm willing to look into the possibilities of remote working locations. The study concluded that telecommuting would be feasible and beneficial to the company. It would reduce costs and limit environmental pollution while increasing productivity at the same time (Van Meel, 2010).

During the last few decades telecommuting has become more and more attractive for employers and employees for several reasons. For companies with knowledge workers their employees are generally very valuable staff and replacing them would take a lot of energy, time and costs. If they can support the lifestyle of the employees, for example having more flexible working hours for young mothers who have to pick up their children, employers would do that (Forgacs, 2010). Many HR departments would also state nowadays that telecommuting is a part of an attractive package of flexible working conditions. This helps companies to attract and retain the best and brightest employees. Even when specialists are not in the vicinity of the offices, telecommuting is providing opportunities to attract the right people for the job. The combination of fuel savings for commuting (some companies provide a company car or a compensation) and the savings for having less real estate are other benefits to companies (Meinert, 2011). Employers also stated that work efficiency and employee satisfaction have increased since the introduction of telecommuting (Forgacs, 2010). This results in less employee absenteeism and bigger turnover (Di Martino and Wirth, 1990). The government is giving a helping hand by providing cuts in employment taxes and contributions that would have a stimulating effect on the expansion of telework (Forgacs, 2010). Telecommuting is also supported by the Dutch government. According to Van Tilburg (2007) this is to be expected, since the population density is relatively high, the roads in particularly in the Randstad are hopelessly congested every day, a relatively large part of the labor force works part time and child care still needs much improvement. A newspaper article in the *Telegraaf* of November 23rd, 2011 announced that the federal government is setting up a fund which will help stimulate telecommuting by providing for costs for computer soft- and hardware. The goal of this 24 million euro fund is to reduce the traffic jams in the Netherlands. The government also believes the organizational costs for smaller companies to adapt to telecommuting are relatively large in the beginning, but cost reductions and profits gained from increased productivity will be collected in a later phase. Similarly telecommuting can have strong benefits for employees. Telecommuting allows employees to have a flexible time schedule. Employees find themselves more flexible in dividing their time between work and family, which improves the harmony of work and private life. Thus providing a work environment that is customized to the worker and allowing a more flexible choice of jobs. Telecommuting not only reduces transportation costs for the employer, but also

for the employee. All in all, telecommuting results in increased satisfaction and stress reduction for the employee (Knight and Westbrook, 1999).

Telecommuting also has its disadvantages. Spending more time out of the office while telecommuting means that an employee will be missing out on what is going on in the office. The office provides for a strong social function. Some employees cherish the opportunity to work at home once in a while but prefer to get out of the house and going to the office. The office provides for some people the sole place to meet others. This interaction with colleagues may give meaning to work and forms a basis to enter networks and friendships are formed. For newcomers the office also has the function of integration and where acculturation processes take place. Employees who are constantly out of the office are 'out of the loop'. They may not participate in processes anymore which build up social capital and are excluded from networks to exchange tacit knowledge (Davenport, 2005; Godfrey, 2008; Kurland and Cooper, 2002). Telecommuters feel isolated from organizational development and promotional opportunities. There are some challenges for the employer as well. According to Forgacs (2010) a company which introduces telecommuting must make sure they are ready to implement it. It means that issues on personnel and labor, law, tax and IT-questions must be tackled. For example handling data on laptops and mobile phones should be thought over well and policies should be enforced. Other issues are managing the performance of employees working out of the office, fragmentation of the social network of the workplace, disruption to teamwork, and competition or hostility between program participants and nonparticipants (Kurland and Bailey, 1999). The fact that opportunities for face-to-face contact are much less in organizations that allow employees to telecommute can strongly influence several formal corporate processes, such as IT-processes. The focus of this thesis is on the influence of face-to-face contact on employee performance appraisals.

First, let's take a closer look at performance appraisals. What are performance appraisals, what is the goal of using performance appraisals, how widely spread is the use of performance appraisal and what are the consequences when not using it well?

Performance appraisal, or also called performance evaluation or performance review is a human resource process. The performance rating process is one of the most

important human resource processes. The process often starts with setting personal development goals with the employee performing the job, while the superior who will be rating the employee, hereafter referred as the rater, collects information on the performance of the employee, evaluating by giving feedback in a conversation. The result of the appraisal of the rater comes either with a financial and/or a promotional reward or punishment, for example a salary cut or simply no financial reward. Looking at the next round of appraisals new goals and expectations for the future will be set (Kondrasuk, 2011). Information to perform the appraisal is often retrieved over time while observing the behavior of the employee. In general the employee can be measured on Key Performance Indicators (KPI), like the quality of one's work, how he or she is managing his or her work or the way of interacting with clients and colleagues. Depending on a specific position and tasks the KPIs which the employee is assessed on may vary. For example sales employees' performance is measured by looking at sold sales, managers for the result of the department and the way the department is lead.

Performance appraisals were and are still used to let the employees know what they are expected to do, aligning their expected performance to the goals of the company and appraise them accordingly (Kondrasuk, 2011). Performance appraisals serve two goals according to Kondrasuk (2011); the first goal is to assess the employee's performance and help to improve job performance, the second goal is for administrative purposes and serves as a basis to cut back or increase salary level, a promotion or determine what training or coaching is necessary. When an employee falls behind performance, the manager will try to improve that by coaching, training etc. If still the performance falls behind consistently, the employee may be laid off or transferred to other more suitable work (Johnson and Geal, 2010). It may for example have an influence on career paths on a longer term. And in a shorter term the wage of an employee may be influenced by it. Thus in short, the assessment of the performance is used by organizations for control, accountability and staff development (Wilson, 2002).

The use of performance appraisals began to be a common part of companies during the Industrial Revolution when more and more bureaucratic companies were founded (Fandray, 2001). Atchison et al. (2010) state that most organizations use some form

of performance appraisal. In some industries the use of performance appraisal is more pervasive. These industries are the so called white collar environments where you will find the clerical and managerial positions. In the so called blue collar environments more manual labor is done and the use of performance appraisal is less extended.

Grote (2002) states that the performance appraisal is the one process in the company which has the most influence on an employee's career and work lives. Roberts (1998) warns for the negative effects of the (mis)use of performance appraisal. Raters can be reluctant to give negative feedback for example. When an employee does not perform well enough and the rater is reluctant to give negative feedback, the company will be sure to lose a juridical case when there is no documentation of the performance under expectation of the company. If a performance management system is not well implemented and carried out, the costs to the organization is extremely high. 'Employees lose confidence in the organization's performance management abilities, and supervisors become even more reluctant to take time and effort to document performance problems' (p.309 Roberts, 1998). According to the research of Kavanagh et al. (2007) the way the performance appraisals are conducted are of influence on the perception of fairness to the employee. Perceived fairness in turn will probably relate to more positive organizational attitudes such as job satisfaction and organizational commitment in turn (Blau, 1999).

Performance appraisal processes are widely used and can have a large impact on both employer and employee. How will telecommuting (having less face-to-face contact) impact the performance appraisal process?

The following hypotheses were researched in this thesis:

1. The frequency of face-to-face contacts positively influences the performance evaluation.
2. The positive effect of more frequent face-to-face contact on performance evaluation is bigger for female than for male raters.

The focus in most research literature on the drivers of performance appraisals is on gaining an improved working relationship through the moments for contact. Face-to-

face contact offers opportunities to improve mutual understanding between the rater and the ratee by getting acquainted with each other and thus reducing the effects of stereotyping. There is yet little research on the sole effect of face-to-face contact on performance appraisal.

244 persons participated in the experiment. Participants were asked to imagine that they were the manager in a company who was asked to appraise a just above average employee. There were two different scenarios which were randomly assigned. In one scenario the participant had seen the employee once a week and in the other scenario the participant had seen the employee once every two months.

There is no support found for the hypothesis that increased face-to-face contact between manager and employee positively influences the performance appraisal. There is also no support found in this research for the hypothesis that this effect is larger for female raters. The data of the samples failed the normality check. Therefore the results are tested again with a non-parametric test and this confirmed the initial findings: the hypothesis is not confirmed by the results.

Although the hypotheses are not confirmed, this research has shown that there is a tendency to receive a higher performance appraisal when the rater has seen the ratee more often. Further research in a more realistic setting is needed to investigate the hypothesis. Future possible confirmation of these hypotheses would benefit employees in tips and tricks on how to improve their performance appraisals. For companies it could lead to insights in approach to improve the manager – employee relationships. Overall, more awareness and knowledge on the factors that impact performance appraisals are needed to improve the process.

The remainder of this thesis is structured as follows. Chapter 2 gives an overview of the current state of knowledge on this topic, introduces the theoretical perspectives which are used to formulate the hypotheses, describes the hypotheses and explains some terms which are used. In chapter 3 the research method is described. The results are discussed in chapter 4. Chapter 5 concludes this thesis.

Literature review

In this chapter I describe the current state of academic knowledge. I start by explaining how the focus on performance appraisal changed from the economic view to the psychological view.

‘The Landy and Farr (1980) article, along with another by Feldman (1981), shifted the focus of performance appraisal research from scales and rater training to understanding the rater as a decision-maker who processes social cues’(page 322 Ilgen et al, 1993). ‘From this perspective, research focused on the accuracy with which raters make such judgments and the application of knowledge about judgment processes to the development of appraisal systems’(page 322 Ilgen et al, 1993). According to Lefkowitz (2000) the cognitive revolution in the social and behavioral sciences has influenced the direction of the research on performance appraisals in the last 30 years. The performance rating is no longer a simple rating on the basis of mere performance, but since then it has been seen as a complex process of information-processing tasks influenced by psychological and sociological influences. It is important to understand which factors influence the performance appraisal process in order to come to a better understanding of the performance appraisal process and performance appraisal systems. Research on the processes, which the rater goes through, contributes in providing information that suggests ways to improve performance appraisal systems in organizations. This can be divided in the effects of characteristics of the ratee (demographic features for example gender, race, sex), rater attributes (for example cognitive style, knowledge of the job which is been rated) and characteristics of the setting (for example, has the rater been receiving training, purpose of the appraisal, Ilgen et al, 1993).

Telecommuting consists of physical distance and less face-to-face contacts. What are the effects of research on the working relationship in general and performance appraisal specifically? I am going to describe some research which state there is yet another variable in the game, for example an improved working relationship through more face-to-face contact.

Judge and Ferris (1993) discussed spatial distance (i.e., between supervisors and subordinates) as a potentially important factor in the performance evaluation process. Judge and Ferris (1993) found out that having a close working relationship (spatial distance), like having more frequent contact and the quality of the interaction, positively influences the probability that the supervisor like those subordinates more than others. Judge and Ferris (1993) researched this by letting the subordinate report how close he or she has worked together with the supervisor. The better the supervisor liked the subordinate, the more positive the performance appraisal was.

The hypothesis that physical distance is important in the interactions and appraisals between superior and subordinate is also confirmed by Mitchell; Wexley & Klimoski (1994). The underpinning reason is that the physical distance can affect the opportunity to observe behavior and performance.

Napier and Ferris (2001) compared different theoretical frameworks and have concluded that distance is a key variable in the different theoretical frameworks (Homans, 1951; Byrne, 1961; Triandis, 1959). Close distance in these models has a positive effect on performance evaluation, among communication effectiveness and subordinate satisfaction. Napier and Ferris (2001) have found a link between structural distance and a positive effect on performance evaluations. Structural distance is the physical distance between the rater and ratee and the amount of interaction. Napier and Ferris (2001) defined organizational structure, like the span of control, as a part of physical distance.

Napier and Ferris (2001) found in their literature review that greater physiological distance will lead to greater functional distance. 'Individuals who are closer in terms of experiencing less dissimilarity are hypothesized to experience a closer, better working relationship as well.'(p. 344 Napier and Ferris, 2001).

Subordinates who feel they have better access to their managers and who actually interact more frequently are developing a closer and better working relationship. Previous work in the area of proximity and interaction opportunity supports this, including research by Homans (1951).

Napier and Ferris (2001) also state that greater communication effectiveness may result from less formal distance in position between the supervisor and the subordinate. This in turn may lead to more accurate assumptions about the other person's behavior. The end result will be a better working relationship with a greater actual functional efficacy. Napier and Ferris (2001) think that a higher performance evaluation can be biased by a greater affect of the rater towards the ratee or established by the improved working relationship and actual improved performance.

In Jones' (2011) article he cites other authors who have researched that familiarity with each other will increase affect. The meaning of affect can be considered as liking (Lefkowitz, 2000). Jones explained that the interaction itself is an opportunity to get to know each other and an opportunity to come to a real understanding and facilitate liking. The sole effect of repeated exposure can start a whole variety of positive outcomes, like enhanced perceptions of similarity (Moreland & Beach, 1992), validity (e.g., Arkes et al., 1989), and positive affect (Claypool et al, 2007).

Seeing each other more often may lead to a certain relationship, according to Judge and Ferris (1993) this may lead to a better relationship and even entrance to the in-group if you were in the out-group first. The in-group is considered as the group which the individual can identify the most. The individual will be inclined to attribute positive characteristics to the member of the in-group. They are more familiar with each other and they consciously or unconsciously want to preserve the good relationship and the rater will therefore give the ratee a more positive rating.

Lefkowitz (2000) also stated in his literature overview that the general conclusions of 24 studies on the relationship between the rater's affect and the performance appraisal ratings are positively correlated. Meaning it resulted in higher ratings, a higher quality of relationship, less inclination to punish poor performance, and greater halo and less accuracy. Feldman (1981) suggested that affect can bias causal attributions in the performance appraisal process. DeNisi (1984) states that affect influences the way the information is being recalled at the moment the evaluation was made, resulting in a bias in the performance evaluation. The effect of interpersonal affect holds stronger for subjective performance measures versus objectives ones (Alexander and Wilkins, 1982).

The majority of the existing academic research on face-to-face contact and performance appraisal is on actually gaining a better working relationship.

Hypotheses development

The theories, which are described in the previous chapter, are focusing on variables which are mediating or moderating the relation and face-to-face. Mediating or moderating variables are a better working relationship and affect. I am actually interested in the mere effect of face-to-face contact and the appraisal performance. I continue with the theories which are the basis for my hypotheses and end form my hypotheses.

Karremans et al. (2009) sum up that research on literature concluded that social interaction may have an influence on cognitive processes. This depends on who the interaction partner is and which psychological processes are taking place. Dreisbach (2011) explains that cognitive control enables intelligent systems to select the relevant information through the mass of information out there. Selecting relevant information is part of the performance appraisal process. According to Dreisbach the social presence of another person will influence the cognitive processes. In particular a certain area of the brain (the orbitofrontal cortex) can read the emotional signs and social cues in a face-to-face situation. This area of the brain is the part where we can exhibit empathy and makes sure that we control our behaviors to not to upset other individuals (Beer, Shimamura, & Knight, 2004).

Affective and evaluative impressions can be formed almost instantaneously after initial exposure to the stimulus (Zajonc, 1980), and cognitive categories reflecting stronger affect tend to be associated with greater memory accessibility (Feldman & Lynch, 1988).

Varma et al. (1996) state that the level of affect can intervene in the cognitive processes. Several studies (Isen & Daubman, 1984; Wyer & Srull, 1986) found that different stages of the cognitive processes are influenced by affect. Stages consist of acquiring, storing and retrieving information.

Quickly and accurately perceiving others' facial expressions is very important for successful social interaction. Claypool (2007) did several experiments in his research

and let participants read the facial emotions of others. When the participant saw a familiar face, the emotion of the face would be rated happier. The results of Claypool suggest that familiarity is one of the signals humans use to interpret if we like the faces of the other individuals.

Kunst-Wilson & Zajonc (1980) conclude from their literature research that experience with a certain stimuli will increase the affective reaction to it. They describe the example with music, where the participants are being familiarized with a certain pattern. An explanation is that the participants can anticipate the music and this is what makes the music attractive.

It is researched that repeated exposure to stimuli increases liking (Zajonc, 1980). Jones (2011) has researched that if you see the other person more often, the result will be positively influenced. Jones has controlled his research for the factor liking, and only researched the effect of mere exposure. Effects are enhanced perceptions of similarity (Moreland & Beach, 1992), validity (e.g., Arkes, 1989), and positive affect (Claypool et al., 2007). Jones establishes that there is a firm relation between familiarity and positive attitude following it.

This thesis examines if having more face-to-face contact between the rater and the ratee will positively influence the outcome of an appraisal. More face-to-face contact is a trigger for social psychological influences. I am focusing on the mere effect of having more face-to-face contact.

Hypothesis 1

The frequency of face-to-face contacts positively influences the performance evaluation.

Furthermore, having less face-to-face contact may have a different magnitude of influence on performance appraisal for female raters versus male raters due to gender differences.

Kanter (1976) reported that the women at work have a so-called female orientation; meaning women are concerned with close, immediate relationship. Researchers as Bartol (1974) and O'Leary (1974) found that women differ in their socialization

experiences and are taught to value social relationships and affiliation, whereas men are taught to value achievement and competition. Research by Hofstede (1994) points out that men choose ego values, like the need for careers more than social values and cooperation in comparison to women. Different evolutionary psychologist (example given Buss and Kenrick, 1998) state that sex differences come from contrasting sexual strategies. Men compete with other men for sexual access to women, so they dispose an attitude which includes violence, competition and risk taking. Women established nurture and prefer long-term mates to support a family. According to Eagly & Wood (1999) more research psychologist are willing to acknowledge that some aspects of social behavior, personality, and abilities differ between women and men. Dobbins (1986) found that female raters gave higher ratings when appraisals are made for promotion or personnel decisions in comparison to experimental of feedback purposes. A possible explanation is that the consequences for a promotion or personnel decisions is larger than when an appraisal is only performed to serve as feedback for the ratee. This can be connected to the nature of nurture of women. Dobbins (1986) found no bias in the ratings of male raters.

According to Gunawardena and Zittle (1997) intimacy and immediacy are correlated both with physical distance. Immediacy is defined as the psychological distance which the person who communicates puts between him or her and the person who is communicated to. Therefore increased face-to-face contact should lead to more intimacy and immediacy. Since women are concerned with close and immediate relationship face-to-face contact will influence performance appraisal.

Also, the environment of performance evaluation is a pressure situation: earlier is mentioned that the consequences for the ratee can be substantial. A pressure situation like a performance evaluation will positively influence the outcome even more when the rater is female. Eagly & Carli (1981) concluded in their research that women, compared to men, were shown to agree more with other people in a larger sample of conformity and persuasion studies, which was in line with the research of Cooper (1979). Agreement, particularly in groups, is a way to show harmony. Whereas disagreement gives cues that one wants to remain independent and one attracts attention to itself (Eagly & Wood, 1991). In general, Eagly found that women are

also more caring for the relationship and are more prone to agree and are more conforming in pressure situations. Appraising someone can be categorized as a pressure situation, since the outcome can have severe consequences on someone's career. According to Scherer female raters give higher ratings than male raters. But there are scholars who do not agree with him. A possible explanation is that women are not as confident as men in their managerial position and try to compensate this insecurity by giving a higher rating.

The positive effect of more face-to-face contact will be larger for female raters than for male raters due to the fact that female raters are more concerned with social values like good relationships than men. More face-to-face contact will increase the desire to maintain a good relationship. In different research women, compared to men, do agree more than men in different persuasion and conformity studies. I expect the effect of having more face-to-face contact will hold stronger for women, resulting in higher performance appraisals – in relation to men.

Hypothesis 2

The positive effect of more frequent face-to-face contact on performance evaluation is greater for female than for male raters.

Method

The method I have used is described in this chapter. Characteristics on the participants are also given in this chapter.

An experiment has been conducted to answer the research questions. The participants in this experiment were colleagues from a large consultancy firm, the professional network and friends of the author. The participants were selected by their educational and professional background. Potential participants were invited by email, by Yammer, by LinkedIn and by Facebook in Dutch to take part in this research and to fill out an online questionnaire. Participation was voluntary and no reward or compensation was provided. Out of 900 persons who were approached to participate, 244 persons have filled in the questionnaire. Here are some characteristics of the sample. In total 210 out of 244 questionnaires are completely filled out. In 34 questionnaires at least one answer to the questions was not given. The mean, also known as the average, score for performance appraisal is 77,53. 36% of the participants were female en 48% of the participants were managers. The average participant was 34.9 years old, had 10.9 years of working experience and 4.3 years of experience in performing appraisals. See table I for an overview of descriptive statistics. The first and second columns are the lowest and the highest scores which are filled in by the participants. The third column is the average of all the scores filled in.

TABLE I
Descriptive statistics on sample

	Minimum	Maximum	Mean
Performance appraisal	6	99	77,53
Age	25	55	34,89
Gender (0=male; 1=female)	0	1	0,36
Years of working experience	1	30	10,90
Manager (0=yes; 1=no)	0	1	0,62
Years of appraisal experience	0	30	4,25

This was an experiment with two different scenarios. The frequency of face-to-face contact was manipulated by the following: in one story the participants have seen the employee once a week and in the other story the participants have seen the employee once every two months. All other information remained the same. The internet tool for conducting the experiment is provided by Thesistool.com and I used the randomization option for the two versions of the questionnaire.

Participants went to the questionnaire online in the first two weeks of February in 2012. Filling out the questionnaire took less than 5 minutes. The participants received written information in Dutch. The scenario stated that they are invited to imagine if they were a manager in a company. They had to appraise an employee on his performance over the past year. The employee's gender is not mentioned, in order to leave out the impact of gender of the ratee on the performance appraisal process. The employee is performing just above standard. The results and the standard results are provided. The scenario also described how often the manager has seen the employee. The participants are asked to rate the performance of the employee on a scale from 1 to 100. Then the participants were asked two questions to check if they had understood the scenario well enough. The questions were two manipulation checks on how often the participants have seen the ratee in the story. This could be indicated on 7-point Likert-scale, 1= strongly disagree and 7 = strongly agree. Lastly they were asked to provide some information on themselves: age, gender, number of years of working experience, whether they are in a managing position or not and the number of years of experience in rating employees.

A few questionnaires returned with missing data. Because the sample is large, there is no need to replace or delete the data for missing values.

Results

The results of the experiment are described in this chapter. I describe the effectiveness of the manipulation checks, characteristics of my sample, the results of my preliminary and hypotheses tests and lastly robustness checks.

The data reveals that the manipulation checks were successful. In order to check if the participants have read and understood well enough how often they have seen the employee, they were asked to indicate on a 7-point Likert-scale (1 = strongly disagree, 7 = strongly agree), to what extent they agree with the following statements: 'I have seen the employee often.' and 'I have seen the employee more often than once per month.' For the first item the group with the scenario of having seen the employee once a week has a mean of 5.0, while the group with the scenario of having seen the employee once every two months has a mean of 4.2, $t(220) = 4.6$, $p < 0.01$. Although the mean for the participants who have seen the employee once a week is higher than the mean of the participants who have seen the employee once every two months, the difference between the two values is small. For the second item, the group who has seen the employee once a week has a mean of 5.9 and the group which has seen the employee once every two months has a mean of 2.8, $t(218) = 4.5$, $p < 0.01$.

Table II shows the descriptive statistics on the dependent variable (frequency of face-to-face contact) and also shows the breakdown for gender. Overall, more men participated than women. The number of men and women are quite well-balanced for the two different scenarios (they were randomly assigned). The means for the men and women with the scenario with the face-to-face contact once a week are very close to each other. The means in the scenario with the once every two months are more different, but still the numbers are very close to each other. Looking at the standard deviations you can see that the variability of the scores for performance appraisal given by women who had seen the employee once every two months is the largest, even twice as large as in the other cells.

TABLE II

Descriptive statistics dependent variable per cell in a 2x2-design

Face-to-face contact	Gender	Mean	Std. Deviation	N
Once a week	Male	79,15	12,905	68
	Female	79,40	10,174	42
	Total	79,25	11,888	110
Once every 2 months	Male	77,13	12,931	68
	Female	75,53	23,949	32
	Total	76,62	17,127	100
Total	Male	78,14	12,910	136
	Female	77,73	17,477	74
	Total	78,00	14,642	210

Table III shows the correlations between the dependent variable (performance appraisal) and the independent variables, such as age and gender. There are some significant relationships with $p < 0.01$ between the dependent variable of gender and the control variables. I sum up a few. There are relatively younger women in my sample or relatively older men. The women in my sample had relatively less working experience than the men. They also had less years of experience in appraisals.

TABLE III

Correlations between performance appraisal (dependent) and independent variables

		Performance Appraisal	Age	Gender	Work Exp	Manager
Performance Appraisal ^a	Pearson Correlation	1				
	Sig. (2-tailed)					
Age	Pearson Correlation	-,054	1			
	Sig. (2-tailed)	,431				
Gender ^b	Pearson Correlation	-,011	-,168*	1		
	Sig. (2-tailed)	,877	,014			
WorkExp ^c	Pearson Correlation	-,052	,960**	-,155*	1	
	Sig. (2-tailed)	,452	,000	,024		
Manager ^d	Pearson Correlation	,068	-,324**	,111	-,357**	1
	Sig. (2-tailed)	,325	,000	,108	,000	
ApprExp ^e	Pearson Correlation	-,044	,710**	-,163*	,717**	-,459**
	Sig. (2-tailed)	,520	,000	,017	,000	,000

*.Correlation is significant at the 0.05 level (2-tailed).

**.Correlation is significant at the 0.01 level (2-tailed).

^aPerformance appraisal is measured on a scale from 1 to 100.

^b0 = Male, 1 = Female

^cYears of working experience on a scale from 1 to 30.

^d0 = I am a manager, 1 = I am not a manager

^eYears of appraisal experience on a scale from 0 to 30.

In order to continue with the data I first perform some preliminary analyses to check the underlying assumptions. There are two subgroups in de sample based on how often the rater has seen the ratee. According to figure I the performance appraisal score has an approximately normal distribution. Negative values for the skewness statistics indicate that scores tend to accumulate on the right of the distribution, but this does not affect the normal distribution. Positive kurtosis statistic values indicate that the distribution is pointy, but the number is not significant.

$Z_{\text{skewness}} \text{ subsample}_{\text{seen once a week}} = -2.188/.216 = -10.13$. This is significant. This means that relatively many scores of the respondents are on the right side of the distribution.

$Z_{\text{skewness}} \text{ subsample}_{\text{seen once every 2 months}} = -2.816/.223 = -12.63$ This is significant. The same applies to this subsample; the scores of the respondents are much more on the right side of the distribution.

$Z_{\text{kurtosis}} \text{ subsample}_{\text{seen once a week}} = \sqrt{9.425/.428} = 4.69$ This is not significant. A number of the Kurtosis would have said something about the flatness or the pointiness of the distribution.

$Z_{\text{kurtosis}} \text{ subsample}_{\text{seen once every 2 months}} = \sqrt{9.807/.442} = 4.71$ This is not significant.

Figure I

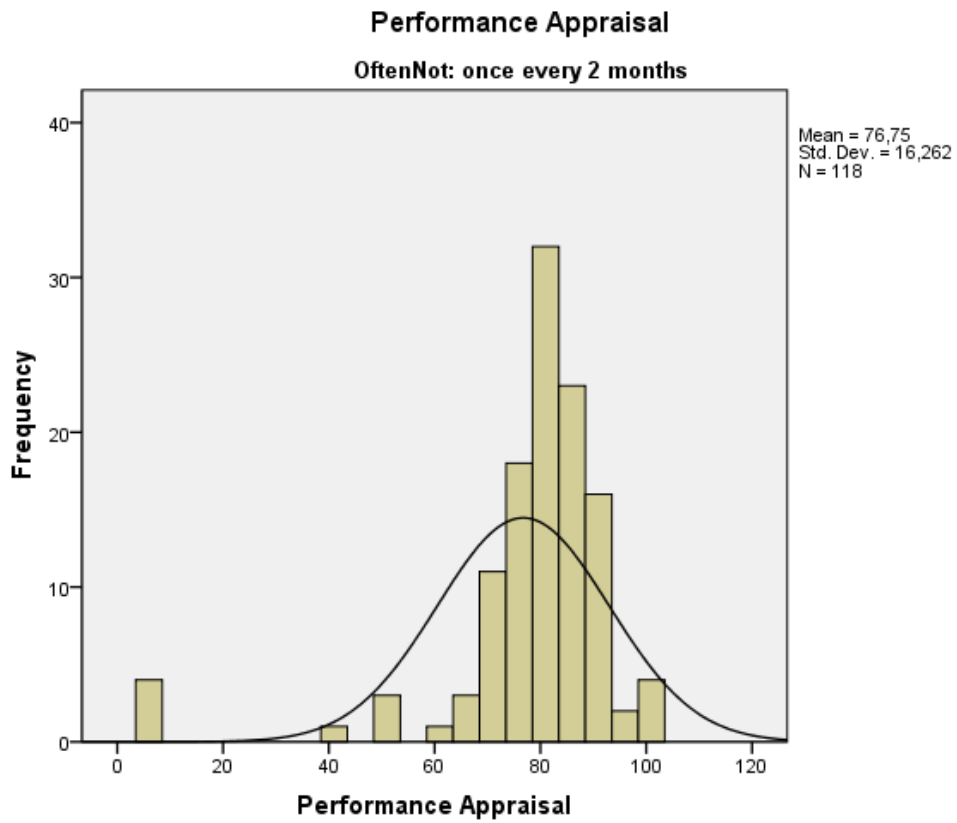
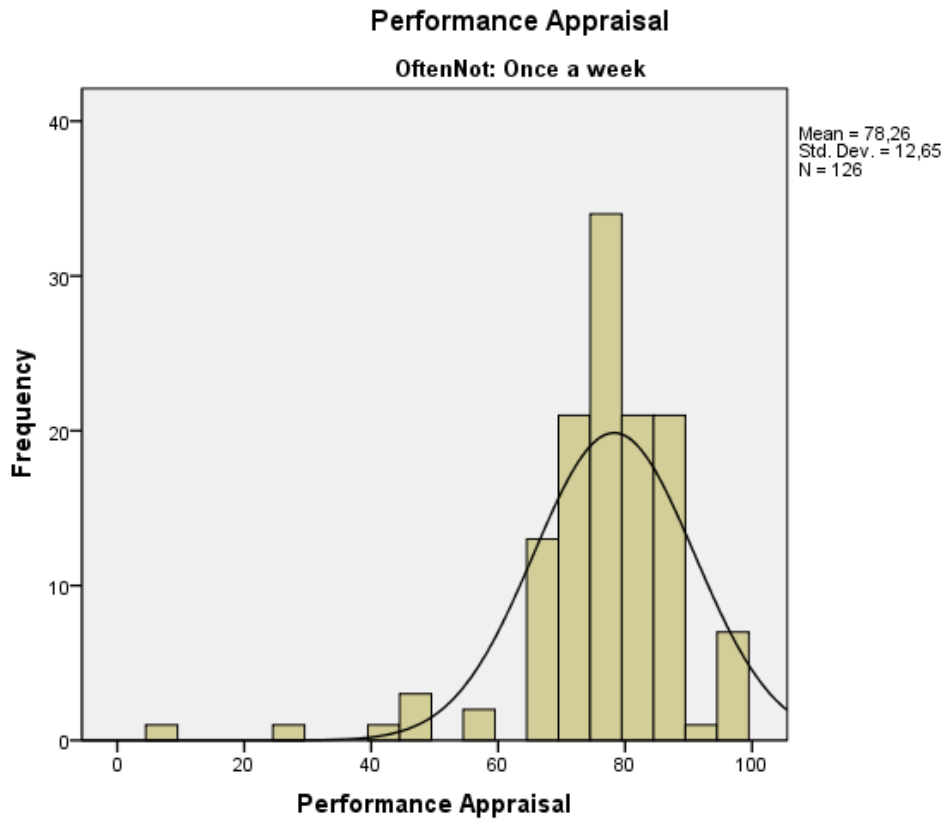


Table IV
Statistics normality

Performance Appraisal			
Once a week	N	Valid	126
		Missing	5
	Skewness		-2,188
	Std. Error of Skewness		,216
	Kurtosis		9,425
	Std. Error of Kurtosis		,428
	Once every 2 months	N	Valid
Missing			2
Skewness			-2,816
Std. Error of Skewness			,223
Kurtosis			9,807
Std. Error of Kurtosis			,442

I use another test to assess the normality of the distribution of a variable. This test is the Kolmogorov-Smirnov test. It is shown in table V that there is a significant score which indicates a deviation from normality. Still, taking a look back at figure I the shape is still bell-shaped and the distribution of the data is approximately normal. I conclude that my sample size is large enough and is approximately normal.

Table V
Test of normality

Frequency face-to-face contact		Kolmogorov-Smirnov ^a		
		Statistic	df	Sig.
Once a week	Perf. Appr.	,193	126	,000
Once every 2 months	Perf. Appr.	,238	118	,000

There are a few outliers, but very small in number compared to the sample. No manipulation on outliers is performed. A table with the 5 highest and lowest extreme values is shown. It can be assumed that 5 numbers in the next table are caused by misinterpretation of the given scale for 1 to 10 instead of 1 to 100.

Table VI
Extreme values

Frequency face-to-face contact				Value
Once a week	Performance Appraisal	Highest	1	99
			2	99
			3	99
			4	99
			5	99 ^a
	Lowest	1	7	
		2	28	
		3	42	
		4	49	
		5	49 ^b	
Once every 2 months	Performance Appraisal	Highest	1	99
			2	99
			3	99
			4	99
			5	94 ^c
	Lowest	1	6	
		2	7	
		3	7	
		4	8	
		5	39	

The Levene test is a test to assess the equality of variances in sample and one advantage is that it does not require normality of the underlying data. The Levene test, which is performed with the factorial ANOVA test in table VII hereafter, produced a significant F-value which means the variance of performance appraisal cannot be assumed to be equal in the two subsamples (one sample which has seen the ratee once a week and the other sample which has seen the ratee once every two months). Put in other words: the assumption that the variances in the two subsamples with different frequency of face-to-face contact cannot be met.

Now I continue with the testing of the hypotheses. First, a t-test is used to analyze if there is a significant difference between the answers regarding performance appraisal given by the two groups with the different scenarios in table VII. There is no significant effect found for the two subsamples of the group which has seen the employee once a week and the group which has seen the employee once every two months. The first hypothesis that having more face-to-face contact with an employee positively affects the performance appraisal is not confirmed by the results. Still, it is interesting to mention that the mean of the group which has seen the employee once a week ($M = 79.25$) is slightly higher than the group which has seen the employee once every 2 months ($M = 76.62$).

TABLE VII

Statistics subsamples

Frequency contact		N	Mean	Std. Deviation
Performance Appraisal	Once a week	110	79,25	11,888
	Once every 2 months	100	76,62	17,127

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
Perf. Appr.	Equal variances assumed	6,957	,010	,894	73	,374
	Equal variances not assumed			,822	41,327	,416

A factorial ANOVA is used to test the second hypothesis. The results are shown in table VIII (the first part of the table is repeated). There was no significant interaction effect found of gender and the frequency of face-to-face contact on appraisal performance. The second hypothesis is not confirmed by the results of this research. Nevertheless, it is interesting to mention that the mean of the female participants ($M = 79.40$) in the scenario of seeing the employee once a week is slightly higher than

the mean of the male participants ($M = 79.15$). In the scenario the participants have seen the employee once every two months the mean of the female participants ($M = 75.53$) is somewhat lower than the mean of the male participants ($M = 77.13$). The partial eta squared of the corrected model tells us that only 0.016% of the variance of the performance appraisal is explained by how often the participant has seen the employee or the gender of the participant. It comes down to that the model does not capture the cause and effect relationship (if any) well.

TABLE VIII (first part is repeated)
Descriptive Statistics

Dependent variable: performance appraisal

Frequency contact	Gender	Mean	Std. Deviation	N
Once a week	Male	79,15	12,905	68
	Female	79,40	10,174	42
	Total	79,25	11,888	110
Once every 2 months	Male	77,13	12,931	68
	Female	75,53	23,949	32
	Total	76,62	17,127	100
Total	Male	78,14	12,910	136
	Female	77,73	17,477	74
	Total	78,00	14,642	210

Tests of between-subjects effect

Dependent variable: performance appraisal

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	700,766a	7	100,109	,459	,864	,016
Intercept	12958,814	1	12958,814	59,352	,000	,227
ApprExp	7,113	1	7,113	,033	,857	,000
Manager	155,499	1	155,499	,712	,400	,004
Age	60,831	1	60,831	,279	,598	,001
WorkExp	30,117	1	30,117	,138	,711	,001
Frequency	405,338	1	405,338	1,856	,175	,009
Gender	48,973	1	48,973	,224	,636	,001
Frequency * Gender	41,061	1	41,061	,188	,665	,001
Error	44104,229	202	218,338			
Total	132289,000	210				

a. R Squared = ,016 (Adjusted R Squared = -,018)

Since the required assumption of normality is not met, a non-parametric test is also used to test the hypothesis. The two-tailed significance is not significant (shown in table IX), which confirms the results from the factorial ANOVA.

Table IX

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Performance Appraisal is the same across categories of OftenNot.	Independent-Samples Mann-Whitney U Test	.951	Retain the null hypothesis.
2	The distribution of Age is the same across categories of OftenNot.	Independent-Samples Mann-Whitney U Test	.631	Retain the null hypothesis.
3	The distribution of ApprExp is the same across categories of OftenNot.	Independent-Samples Mann-Whitney U Test	.383	Retain the null hypothesis.
4	The distribution of WorkExp is the same across categories of OftenNot.	Independent-Samples Mann-Whitney U Test	.308	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

I conclude that the underlying assumptions of normality and equal variances in the two subsamples are not met in different tests. As a consequence a non-parametric test is assessed in addition to the t-test and the factorial ANOVA. The hypotheses are not confirmed by the results of the t-test, the factorial ANOVA nor the non-parametric test.

Conclusion

This chapter gives a summary of the research and draw conclusions. It describes the ways in which this research has contributed to academic research and management. Next I consider the limitations of this research and make recommendations for future research on this topic.

The idea for this research started through the author's own experience of seeing her appraiser not very often when she started in the consulting business. Telecommuting is provided by the author's company and as a result the author and her appraiser would often work at different locations. The percentage of the employees telecommuting is still growing and it is important to understand the effects on the performance appraisal process. There are two hypotheses tested in this research.

The first hypothesis is more face-to-face contact between the rater and ratee will positively influence the performance appraisal. The second hypothesis this influence holds stronger for female raters than male raters. An experiment with two different scenarios is conducted among employees in a consultancy firm and the social network of the author. The scenarios are randomly assigned to the participants of whom 48% was a manager with an average experience of 4.3 years in appraising employees. The hypotheses are not supported by the results found in this research. No evidence has been found that having more face-to-face contact between a rater and ratee will positively influence performance appraisal.

Therefore the second hypothesis does not hold as well; no evidence has been found that this effect holds stronger for female raters than male raters. One possible reason that the hypothesis is not supported by the data is that the setting of the experiment was not realistic enough. Participants are asked to empathize with the written story in the experiment and pretend they are rating an employee, but a more realistic setting will be more suited. A more realistic setting can be actual appraisal situations where face-to-face contact before the appraisal is measured or laboratory experiments with interaction.

My contribution to academic literature has been to research the effect of face-to-face contact on performance appraisals, while controlling for other effects like actually gaining a better working relationship. There has not been research on this specific relationship yet. This research was only focused on the mere effect of having more face-to-face contact whereas previous research mostly focussed on gaining a better working relationship through increased contact, resulting in a higher performance appraisal. It is interesting to consider this research in light of research on for example the positive role of familiarity on liking of Claypool (2007) and the research on the positive effect of liking on performance appraisal (for example Varma et al, 1996). What is so unique about the performance appraisal process that this reasoning does not hold?

Management of companies can have a better understanding of the effect of less face-to-face contact on performance appraisal. They now know it does not influence the fairness or objectiveness of the performance appraisal outcome and hence point their effort to other possible influences. Employees know that by just showing their faces at the office will not positively influence their performance appraisal and hence work on other factors (e.g. a better working relationship) which may have according to other research.

The limitations of this research are an experiment using online questionnaires, which makes the experiment less realistic. This study suffers from invalid external validity. Also the participants are not all managers who are appraising employees in their actual jobs. Future research can focus on more realistic methods to conduct research and choose managers who already appraise employees in their jobs and hence obtain a more representative sample. Another limitation is the difference between the frequencies of face-to-face contact in the two scenarios. In the first scenario the rater has seen the ratee once every week and in the second scenario the rater has seen the ratee once every two months. Although it was significant, the difference is not large. Was it better to make the scenarios more different from each other and then again would it have been a realistic situation if for example the rater has seen the ratee only once a year? This is a challenge in scientific research, especially in experiments.

In the scenario the ratee has been performing just above the standard. Future research can use scenarios where the employee has underperformed. Research state (for

example DeNisi, 1992) that raters are hesitant to give bad news to the ratee if the performance is low. Will having more face-to-face contact positively influence leniency when the rater knows he or she has to give feedback to the ratee? Future research can also focus on different forms of real-time (face-to-face) contact, for example calling by phone (no face-to-face contacts, but there are social cues interchanged by the voice) and videoconferencing (face-to-face contact, but still physical distance).

References

- Alexander, E.R. & Wilkins, R.D. (1982). Performance rating validity: The relationship of objective and subjective measures of performance. *Group and Organizational Studies* 7, 485-496.
- Arkes, H.R., Hackett, C, & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making* 2, 81-94.
- Atchison, T.J., Belcher, D.W., & Thomsen, D.J. (2010). Wage and Salary Administration, retrieved 14/3/2012 on <http://www.eridlc.com/index.cfm?fuseaction=textbook.chpt14>
- Bartol, K. (1974). Sex differences in job orientation: A reexamination. *Proceedings of the Academy of Management*.
- Beer, J.S., Shimamura, A.P., & Knight, R.T. (2004). Frontal Lobe Contributions to Executive Control of Cognitive and Social Behavior. *Emotion en Social Neuroscience*, 1091-1104.
- Beutell, N.J., & Brenner, O.C. (1986). Sex differences in work values. *Journal of Vocational Behavior* 28, 29-41.
- Blau, G. (1999). Testing the longitudinal impact of work variables and performance appraisal satisfaction on subsequent overall job satisfaction. *Human Relations* 52, 1099-13.
- Buss, D.M., & Kenrick, D.T. (1998). Evolutionary social psychology.
- Byrne, D. (1961). Interpersonal and Attitude Similarity. *Journal of Social Psychology* 62, 713-715.
- Claypool, H.M., Hugenberg, K., Housley, M.K. & Mackie, D.M. (2007). Familiar eyes are smiling: On the role of familiarity in the perception of facial affect. *European Journal of Social Psychology* 37, 856-866.
- Cooper, H.M. (1979). Statistically combining independent studies: Meta-analysis of sex differences in conformity research. *Journal of Personality and Social Psychology* 37, 131-146.
- Davenport, T. (2005). Why office design matters.

- DeNisi, A.S., Cafferty, T.P., & Meglino, B.M. (1984). A cognitive view of the performance appraisal process: A model and research propositions. *Organizational Behavior and Human Performance* 33, 360-396.
- Di Martino, V., & Wirth, L. (1990). Telework: a new way of working and living. *International Labor Review* 129, 529-554.
- Dobbins, G.H. (1986). Effects of ratee sex and purpose of appraisal on the accuracy of performance evaluations. *Basic and applied social psychology* 7, 225-241.
- Dreisbach, G.H. & Botthcher, S. (2011). How the social-evaluative context modulates processes of cognitive control. *Psychological Research* 75, 143-151.
- Eagly, A.H. & Carli, L.L. (1981). Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: A meta-analysis of social influence studies. *Psychological Bulletin* 90, 1-20.
- Eagly, A.H. & Wood, W. (1999). The origins of sex differences in human behavior. *American Psychologist* 54, 408-423.
- Fandray, D. (2001). The New Thinking in Performance Appraisals. *Workforce* 80, 36-40.
- Feldman, J.M. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology* 66, 127-148.
- Feldman, J.M., & Lynch, J.G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology* 73, 421-435.
- Ferris, G.R., Munyon, T.P., Basik, K. & Buckley, M.R. (2008). The performance evaluation context: Social, emotional, cognitive, political, and relationships components. *Human Resource Management Review* 18, 146-163.
- Forgacs, T. (2010). Empirical research findings on telework: Management experiences and attitudes. *Business and Economics Horizons* 1, 6-13.
- Grote, D. (2002). Performance appraisal. *Executive Excellence* 19, 12-13.
- Gunawardena, C.N. & Zittle, F.J. (1997). Social Presence as a Predictor of Satisfaction within a Computer-Mediated Conferencing Environment. *The American Journal of Distance Education*, 11, 8-26.

- Hofstede, G. (1994). *Cultures and Organizations, Software Of The Mind*. Profile Books Ltd.
- Homans, G.C. (1951). The Human Group.
- Iacovelli, A. & Valenti, S. (2009). Internet addiction's effect on likeability and rapport. *Computers in Human Behavior* 25, 439-443.
- Ilgen, D.R., Barnes-Farrell, J.L. & McKellin, D.B. (1993). Performance Appraisal Process Research in the 1980s: What Has It Contributed to Appraisals in Use? *Organizational Behavior and Human Decision Processes* 54, 321-368.
- Isen, A.M., & Daubman, K.A. (1984). The influence of affect on categorization. *Journal of Personality and Social Psychology* 47, 1206-1217.
- Johnson, B. & Geal, M. (2010). Is performance appraisal valid? *Training Journal*, 49-53
- Jones, I.F., Young, S.G. & Claypool, H.M. (2011). Approaching the familiar: On the ability of mere exposure to direct approach and avoidance behavior. *Motiv Emot* 35, 383-392.
- Judge, T.A. & Ferris, G.R. (1993). Social context of performance evaluation decisions. *Academy of Management Journal* 36, 80-105.
- Kanter, R. (1976). The impact of hierarchical structures on the work behavior of men and women. *Social Problems* 33, 415-430.
- Karremans, J.C., Verwijmeren, T., Pronk, T.M. & Reitsma M. (2009). Interacting with women can impair men's cognitive functioning. *Journal of Experimental Social Psychology* 45, 1041-1044.
- Kavanagh, P., Benson, J. & Brown, M. (2007). Understanding performance appraisal fairness. *Asia Pacific Journal of Human Resources* 45, 132-150.
- Kirk, J. & Belovics, R. (2006). Making e-working work. *Journal of Employment Counseling* 43, 39-46.
- Knight, P.J., & Westbrook, J. (1999). Comparing employees in traditional job structures vs. telecommuting jobs using Herzberg's hygienes and motivators. *Engineering Management Journal* 11, 15-20.
- Kondrasuk, J.N. (2011). So What Would An Ideal Performance Appraisal Look Like? *Journal of Applied Business and Economics* 12, 57-71.

- Kunst-Wilson, W.R. & Zajonc R.B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science* 207, 557-558.
- Kurland, N.B., & Bailey, D.E. (1999). Telework: The advantages and challenges of working here, there, anywhere, and anytime. *Organizational Dynamics* 28, 53-67.
- Kurland, N.B., & Cooper, C.D. (2002). Manager control and employee isolation in telecommuting environments. *Journal of High Technology Management Research* 13, 107-126.
- Lefkowitz, J. (2000). The role of interpersonal affective regard in supervisory performance ratings: A literature review and proposed causal model. *Journal of Occupational and Organizational Psychology* 73, 67-85.
- Meinert, D. (2011). Make Telecommuting Pay Off. *HR Magazine of June*, 32-37.
- Moreland, R.L., & Beach, S.R. (1992). Exposure effects in the classroom: The development of affinity among students. *Journal of Experimental Social Psychology* 28, 225-276.
- Napier, B.J. & Ferris, G.R. (2001). Distance in organizations. *Human Resource Management Review* 3, 321-357.
- Nilles, J.M. (1998). *Managing Telework: Strategies for Managing the Virtual Workforce*. Wiley.
- O'Leary, V.E. (1974). Some attitudinal barriers to occupational aspirations in women. *Psychological Bulletin* 81, 809-826.
- Roberts, G.E. (1998). Perspectives on Enduring and Emerging Issues in Performance Appraisal. *Public Personnel Management* 27, 301-320
- SIBIS (2003). Statistical indicators benchmarking the information society. European Commission.
- Triandis, H.C. (1959). Cognitive Similarity and Interpersonal Communication in Industry. *Journal of Applied Psychology* 43, 321-326.
- Ullian, D. (1984). "Why girls are good": A constructivist view. *Sex roles* 11, 241-256.
- U.S. Census Bureau's annual American Community Survey, 2009
- Van Meel, J. (2010). The origins of new ways of working: Office concepts in the 1970s. *Facilities* 29, 357-367.

- Van Tilburg, L. (2007), Mondjesmaat thuiswerken. Binnenlands Bestuur. Retrieved on 15/3/2012 on <http://www.binnenlandsbestuur.nl/arbeidsmarkt-en-carriere/achtergrond/achtergrond/mondjesmaat-thuiswerken.99921.lynkx>
- Varma, A., Denisi A.S. & Peters L.H. (1996). Interpersonal affect and performance appraisal: A field study. *Personnel Psychology* 49, 341-360.
- Wexley, K.N., & Klimoski, R.J. (1984). Performance appraisal: An update. *Research in personnel and human resources management* 2, 35-79.
- Wikipedia (2012). Digital Revolution. Retrieved on 15/3/2012 on http://en.wikipedia.org/wiki/Digital_Revolution
- Wilson, F. (2002). Dilemmas of appraisal. *European Management Journal* 20, 620-629.
- WorldatWork (2009). Telework Trendline 2009.
- Wyer, R.S., & Srull, T.K. (1986). Human cognition in its social context. *Psychological Review* 93, 322-359.
- Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist* 35, 151-175.