TALLINN UNIVERSITY OF TECHNOLOGY
School of Business and Governance
Department of Business Administration

Tomoko Yamagata

COMPARATIVE ANALYSIS: THE PERCEPTIONS OF FAIRNESS IN PERFORMANCE REVIEWS WITH ALGORITHMS

Master's thesis
Programme International Business Administration, specialisation -

Supervisor: Merle Ojasoo, PhD
Co-supervisor: -

Tallinn 2021

I hereby declare that I have compiled the thesis independently
and all works, important standpoints and data by other authors
have been properly referenced and the same paper
has not been previously presented for grading.
The document length is 14,410 words from the introduction to the end of conclusion.

Tomoko Yamagata …Tomoko Yamagata, 11.05. 2021
(signature, date)
Student code: 194261TVTM
Student e-mail address: toyama@taltech.ee

Supervisor: Merle Ojasoo, PhD:
The paper conforms to requirements in force

……
(signature, date)

Co-supervisor: -:
The paper conforms to requirements in force
……
(signature, date)

Chairman of the Defence Committee:
Permitted to the defence
……
(name, signature, date)

TABLE OF CONTENTS

ABSTRACT

Fairness perceptions in performance reviews affect significantly the employees' motivations in the workplaces, therefore, it is important to achieve fairness in performance reviews for organisational performance. Though the new technologies utilising algorithms for the human resource field created an opportunity to increase fairness in performance reviews by reducing human evaluators' biases, the previous works showed that the participants in the experiments perceived algorithmic evaluations are less fair than human managers' evaluations because they felt algorithms do not take qualitatives aspects into account correctly. There is, however, a lack of information about the process of the employees from different demographic factors or backgrounds perceive fairness, express their opinions with the other workers, and the issues they care about when they imagine both humans and algorithms would join the performance reviews. This research focuses on qualitative analysis of how the participants with work experiences from different demographic factors and backgrounds perceive fairness when both human managers and algorithms join the performance reviews by conducting focus groups with qualitative experiments.

In conclusion, as well as the difficulty of discussing fairness, the several focus groups discussed that increasing the number of evaluators and evaluation criteria would improve fairness, while algorithms can contribute to improving fairness perceptions by joining as one of those additional evaluators and multiple axes of evaluation criteria. It was notably observed that the necessary factors for fairness were not the same as the important factors for agreeable performance reviews, while the employees needed consideration of several types of quality and contexts even though those were often not fair. In the future where humans need to work with algorithmic technologies, even challenging soft skills are required for humans such as strong listening, observing, understanding skills, considering different qualities and contexts in intercultural environments and building trusts with the employees, which algorithms would not be good at.

Keywords: Fairness Perception, Performance evaluation, Algorithms, People Analytics

INTRODUCTION

It is not new that the organisational literature provides strong evidence that the employees form fairness perceptions about the organisational procedures they experience, and that these perceptions are important for the organisational performance through employees' motivation, commitment and outcomes. (Taylor, *et al*. 1995; Konovsky, 2000; Colquitt, *et al*. 2001; Hartmann & Slapniˇcar, 2021). Algorithms created an opportunity to increase fairness perceptions in performance reviews since "algorithmic decision-makers will be perceived to have higher procedural fairness because algorithms follow the same procedures every time, are not influenced by emotional factors, and have no agency, and thus are perceived less biased than human decision-makers." (Lee, 2018, 4). Computer-based algorithms do not make decisions based on their personal preferences like human beings. However, the pilot research indicated that personnel decisions driven by algorithms are perceived to be less fair than those with human evaluators for two main reasons. Firstly, participants as the employees perceived that qualitative information is not taken into account in algorithmic evaluations, therefore, they could not be holistic evaluation. (Newman, *et al*. 2020, 160). Secondly, since data used to train algorithms can reflect the humans' biases, it has been proven that algorithmic decision-making systems have a risk to systematically reinforce racial or gender stereotypes, marginalise minorities or certain members of society (Veale & Binns, 2017; Žliobaitė, 2017; Starke, 2021). Yet, organisations could pair HR algorithms to assess quantitative factors and human evaluators who assess qualitative factors, while this combination might be perceived by the employees as the fairest. This approach might be particularly important in an intercultural and diverse environment because managers' biases about gender or nationality can influence a lot on the performance assessment of his or her subordinates. The research problem is that there is a lack of information about the process of the employees from different demographic factors or backgrounds perceive fairness, express their opinions with the other workers, and the issues they care about when they imagine both humans and algorithms would join the performance reviews. As per the systematic reviews about fairness perceptions of algorithmic decision-making by Starke, "the studies included in this review were almost exclusively conducted in Western democracies, predominantly the US." (Starke, *et al.* 2021, 29). This knowledge is important for those who take part in intercultural and diverse organisations because how they perceive fairness

and what factors they care about need to be taken into account when the organisations try to develop and introduce algorithmic technologies into human resource fields, and also to improve performance review processes themselves, where it would be indicated that which part of those processes human evaluators more value and focus on as qualitative factors.

The research aim is to investigate messaging from participants' feelings, mindsets and attitudes about fairness perceptions when they imagine algorithms join their performance reviews regarding consideration of demographic factors, such as nationality, gender, and backgrounds. The main research question is how the employees from different nationalities, gender or backgrounds would perceive fairness about performance evaluations when they imagine algorithms would join their performance reviews? Finding out the answers to this question would contribute to developing algorithmic assessment software in the future by suggesting what factors intercultural workers care about being assessed by algorithms to achieve fairness in performance evaluation. It would be also essential for improving the HR process and systems because the factors the employees valued and perceived that it is not appropriate to be done by algorithms are the very things that are required and expected to focus on with high-levelled skills for human evaluators. Applying computer-based algorithmic technologies into the human resource fields would be inevitable in the future, therefore, this research is conducted to provide information so that implementation of new technologies would be appropriate for both the organisations and employees in ethical ways.

In this research, focus groups are conducted as one of the qualitative research methods to collect data for complex and open-ended questions. Each participant is asked to read a scenario about a new policy that would be introduced to an organisation where employees' performance would be evaluated by both human managers and HR algorithms using their performance data and recorded videos. Data collected through the group interviews are analysed with a combination of Thematic Analysis (Saunders, *et al.* 2015, 579) and Content Analysis (*Ibid*., 608). In Thematic Analysis, data were coded to identify themes or patterns to analyse participants' perceptions about fairness and unfairness in performance reviews, feelings and comments about a scenario, considering gender and nationalities. A quantitative description of participants' messages would be analysed through Content Analysis. In this analysis, explicit messages such as participants' answers to the questions and interactions with the other participants during the interview are coded and transcribed to be analysed as "manifest content" (*Ibid.*, 609).

In the first chapter, the previous researches related to the topic were reviewed. The methodology of this research is discussed in the second chapter. Then the results of the fieldworks, findings and analysis of the results, the limitations of this research are discussed in the third chapter. At last, the main results of this research, assessments and proposals are mentioned as Conclusion.

1. THEORETICAL BACKGROUNDS

In this chapter, the previous researches related to the importance of individuals' fairness perceptions in performance evaluations for organisations, people's analytics (applying algorithmic technology to the human resource management field.) and applying algorithmic technology especially to the area of performance evaluations to increase perceptions of fairness are discussed.

1.1. Fairness perceptions in performance evaluations

Hereby, the importance of how fair the individuals perceive their performance evaluations are reviewed, as well as the basic definitions of fairness in the context of performance reviews. Fairness is defined as treating everyone equally or equitably based on people's performance or needs (Leventhal, 1980 referenced in Lee, 2018, 4). The pilot research by Robert *et al.* (2020) organised types of fairness based on reviewing the literature on organizational justice theory into three types. Distributive fairness refers to fairness with respect to the allocation of outcomes such as pay and other resources (Alexander & Ruderman, 1987, referenced in Robert *et al*. 2020, 6). This is fairness with respect to the allocation of outcomes such as pay and other resources. (Robert *et al*. 2020, 6). Procedural fairness is defined by the process employed to reach or decide the final outcome. (*Ibid.*, 7). At last, "Interactional fairness is defined by how workers are treated by their organization" (Luo, 2007, referenced in Robert *et al*. 2020,8). Perceived fairness in organizations refers to employees' global perception that decisions and procedures adhere to agreed-upon rules about equitable treatment. Importantly, different sets of rules influence perceptions of fairness (Colquitt & Zipay, 2015). In the other definition, Hartmann & Slapničar (2012), who approached the issues in perceptions of fairness and justice of performance evaluation by implementing uncertainty management theory, understood fairness perceptions in performance evaluations as a coping device, which reassures organizational members that they can expect favourable outcomes, and reduces anxiety about being exploited by the organization or its members (Diekmann *et al*. 2004, 240).

As per the previous researches, it could be said that employees' perceptions about fairness in performance evaluations play one of the important roles in organisations. "The organizational literature provides strong evidence that organizational participants form justice perceptions about the organizational procedures they experience, and that these perceptions explain important workplace outcomes, such as motivation, commitment and job performance" (Taylor, *et al*. 1995; Konovsky, 2000; Colquitt, *et al*. 2001, referenced in Hartmann & Slapniˇcar, 2012, 17). A recent study of the tech industry found that a perceived lack of fairness was the single largest driver of employee turnover, costing the industry $16 billion a year (Scott, Klein, & Onovakpuri, 2017). Indeed, scholars have long recognized the importance of understanding and improving employee perceptions of fairness, particularly the perceived fairness of decision-making procedures (Colquitt, *et al*. 2001).

There are many determinants of fairness in performance evaluations. The literature outlines several normative principles of procedural justice, such as 'consistency' and 'accuracy' (Folger and Bies, 1989; Taylor, *et al*. 1995), but these are not observable or designable characteristics of performance evaluation themselves, nor are they theoretically grounded. (Hartmann and Slapnicar, 2012, 17). Hartmann and Slapničar (2012) investigated (by implementing uncertainty management theory and using 178 managers from the banking industry) two elements related to the performance evaluation process, and two elements that relate to the metrics used in the process, since both are different and relatively independent design elements of performance evaluation processes. (*Ibid*., 20). Regarding the former, […]study the formality of the performance evaluation process (Moers, 2005), and subordinate voice in this process (Libby, 1999). Regarding the latter, […] analyse the use of outcome metrics (Hartmann and Slapnicar, 2012, 20) […] and the diversity of performance metrics relied upon by superiors (Ittner *et al*., 2003). As a result, they found that all four performance evaluation characteristics are related to justice perceptions, yet their effect depends on the level of task uncertainty and tolerance for ambiguity (Hartmann & Slapničar, 2012, 17). The only (subordinate) voice has a strong and universal positive effect on justice perceptions, yet its effect is stronger for managers with higher task uncertainty. (*Ibid*., 28). As the other example, […] (resource) allocations based on past performance and random draw rules lead to the highest fairness perceptions and the lowest expectations that the decisions made will lead to intragroup conflict. (Conlon, Porter, & MacLean Parks, 2004, 1).

This research concerns the procedural determinants of fairness, which require that procedures be consistent, free of bias, and based on accurate information. (Barrett-Howard & Tyler, 1986;

Brockner, 2002; Colquitt, 2001; Leventhal 1980 referenced in Newman, 2020,150 ; Thibaut & Walker 1975 referenced in Newman, 2020,150). In this research, to be free of influence from the personal bias of decision-makers is concerned as a factor that would be particularly significant in an intercultural and diverse work environment and has a possibility to be solved by applying technologies.

1.2. Fairness perceptions in performance evaluations in Japan

The importance of fairness perceptions is essential in both the international and domestic environment. For instance, researches support this idea about Japanese domestic companies as well, which are the organisations consist of employees with relatively homogeneous cultures in the world. According to Inoue (2016), human resource assessment about the performance of employees needs to increase employees' loyalty to the organisation and motivation to work and contribute to the sustainable development of the company by gaining employees' agreement with the result and process of evaluation. To gain their agreements, it is necessary to achieve fairness in the evaluation system, while it is hard to be fair since humans evaluate other humans. (Inoue, 2016, 55).

At the same time, there are particular reasons why it has been hard for Japanese domestic companies to realise fairness in human resource evaluations. Japanese domestic companies usually do not have clear evaluation criteria for employees' performance because they hire the employees at first, then allocate jobs to the employees. Since the scope of each job (a job description) is vague and ambiguous, the performance of an employee tends to evaluate "personality", "making efforts" "hard-working" and so on in Japanese large corporations. (Oguma,2019, No.1527/7209). Evaluation criteria under this system tend not to be clear. As historical background, labour unions in Japan were united for each company, not cross-industry after the second world war. While unions requested employers for long-term employment and salary based on age as an "equality among employees" instead of clarification of job descriptions or transparency in performance evaluations as a "fairness among the jobs". Through the negotiations between the unions and employers, opaqueness in promotions or recruiting were tolerated as a deal and a sacrifice in return for gaining long-term employment and salary based on age. (*Ibid.*, No.6757/7209).

If the transparency and openness (in criteria of hiring and performance evaluations) are achieved, cross-company labour market, the fairness of genders or related issues would be improved. (*Ibid.*, No. 6742/7209). In the past, there were several trials and movement which intended to change those rules and systems, while most of them failed (*Ibid.,* No. 6711/7209). This is because (the employers) tried to introduce only the result-driven evaluation system without improving transparency and openness in the criteria and process of performance evaluations, with the main intention to cut human resource costs. (*Ibid.*, No. 6749/7209).

Yoshida (2016) pointed out through her experiment that the satisfaction of employees with performance evaluations is highly influenced by the relationship between an employee and his supervisor under this system. In the survey JMA Management Center, 86.7% of respondents replied: "Strongly agree" or "Agree" to the question"the performance evaluations in HR is depends on a luck of in which department a person will be assigned, and which manager would be a person's supervisor." (Yoshida, 2016, 30). Yoshida also conducted a survey about supervisors' behaviours to find out that the more employees think their supervisors do not have a viewpoint of fairness, the more they untrust the supervisors, which lead to low satisfaction with their performance evaluations. (*Ibid.*, 32).

In light of the above historical backgrounds and the current situation, fairness perceptions of employees about performance evaluations in Japanese domestic corporations have a huge room to improve as well as the demands for it. Applying algorithms to the human resource management field can be an opportunity to change the rules and system in Japanese domestic corporations to create fairness perceptions in performance evaluations through improving transparency in evaluations, for the first time in our labour history.

1.3. Applying algorithms to human resource management

Algorithms are defined as "computational procedures drawing on some type of digital data that provide some kind of quantitative output through a software program" (Christin, 2017, 2). Another definition is that "a computational formula that autonomously makes decisions based on statistical models or decision rules without explicit human intervention." (Lee, 2018,3). The latter definition is used in this paper since algorithms are not used as independent tools but also the basis and

mechanisms which consists of Artificial Intelligence or structure of Machine Learnings recently emerged as new usages.

Big data analytics have transformed research in many fields, including the business areas of marketing, accounting and finance, and supply chain management (Hamilton, 2019). Human resource management (HRM) personnel have begun to incorporate specialized data analytics into their decision-making processes. To date, much of it has been focused on screening of candidates in the hiring process (Angrave, Charlwood, Kirkpatrick, Lawrence, & Stuart, 2016; McAbee, Landis, & Burke, 2017), using software to sort through employment applications and social media sources to reduce administrative expenses within the HR function. (Hamilton, 2019). For example, video interview evaluations, based on the emergence of companies like HireVue, which have developed algorithms to assess recordings of job candidates for clients including Goldman Sachs and JP Morgan (Corporate Finance Institute 2018, referenced in Newman *et al*. 2020). From the viewpoint of HR staff, this makes sense: Screening candidates is a major, difficult task for HR. Public databases provide a large volume of data to sort through for analyzing patterns that might predict appropriate hires, and accessing these databases could cost significantly less than manually processing applications while yielding a greater likelihood of interviewing appropriate candidates without having to process inappropriate ones. (Hamilton, 2019,2). Moreover, HR departments often justify their own relegation by overfocusing on transactional aspects of their job (e.g., completing required forms or distributing benefits; Armstrong, 2016) instead of strategic concerns. According to Marler and Boudreau (2017), HR managers with a clear focus on business issues are in short supply; they tend to concentrate on the administrative costs of processing applications or onboarding new employees rather than overall firm profitability (Hamilton & Davison, 2018; Marler & Boudreau, 2017).

But from an overall firms performance perspective, HR-related big data analytics should not be confined to screening candidates or to administrative effectiveness. Instead, HR's use of big data should be aimed at capturing the strategic linkage between human capital and profitability, determining how the HR function can enhance the skills and knowledge of employees to develop competitive advantage and improve overall firm performance (Jackson, Schuler, & Jiang, 2014).

1.4. Applying algorithms for the fairness of performance evaluation

As it was discussed in the previous subsection, the employees' fairness perceptions in performance reviews play an important role in the motivations of the employees and hence, the organisational performances. To view decisions as procedurally fair, one must perceive that they are free of bias (Leventhal, 1980). Empirical research has shown that the removal of bias is particularly important to procedural fairness in formal, business-like situations (Barrett-Howard & Tyler, 1986). Moreover, suppressing bias in decision making has been shown to increase employees' perceptions of fairness even when they receive lower performance evaluations (Taylor, *et al*. 1995), and Sheppard and Lewicki (1987) study of organizational executives found that bias suppression was one of the most frequently cited characteristics of procedural fairness. (Sheppard and Lewicki 1987, referenced in Newman, *et al*. 2020).

Increasing use of algorithms could be one of the potential solutions to increase fairness perceptions because algorithms make decisions based on persistent rules, being free from humans' personal likes or dislikes. Whereas human decision-makers are prone to judgment errors due to biases derived from intuition and other heuristics (e.g., Gilovich, Griffin, & Kahneman, 2002), algorithms can reduce, or even eliminate, such biases by relying on mathematical logic that converts various considerations (both quantitative and qualitative) into numerical factors. (Newman, *et al*. 2020). Thus, algorithms enable organizations to relieve decision making from subjectivity and other limitations, which has led to some organizations deploying algorithms to improve the fairness of personnel decision rules (O'Connor, 2016).

At the same time, however, it is needed to consider that algorithms can have biases as well. Algorithmic bias can be a result when the algorithm is poorly trained and inaccurately includes or excludes data (Hamilton, 2019, 89). Due to biased input data or faulty algorithms, unfair algorithmic decision-making systems have been proven to systematically reinforce racial or gender stereotypes, marginalize minorities, or flat-out denigrate certain members of society (Veale & Binns, 2017; Žliobaitė, 2017). For example, multiple implementations of facial recognition software have proven inaccurate with regard to identifying minorities when minorities were not included in the algorithmic training (Croasmun, 2018; Dickson, 2018). [...] algorithmic bias can also occur when the software replicates inherent human bias. (Hamilton, 2019, 89). In 2015, Amazon shut down an applicant screening system that inappropriately excluded qualified women

applicants because the data showed that the overwhelming majority of successful hires had previously been men (Reuters, 2018).

Based on the literature reviews above, algorithmic performance reviews could be unfair because its data can reflect biases of humans, while Newman, Fast and Harmon (2020) considered it is important to examine the employees' reactions to the use of algorithmic decision making in performance reviews since they have a different set of psychological needs. Newman, Fast and Harmon conducted four laboratory experiments and a large-scale randomised experiment in an organisational setting to confirm the relationship of algorithmic performance reviews, fairness perceptions, and organisational commitment. The result indicated that personnel decisions driven by algorithms are perceived to be less fair than identical decisions featuring more human involvement. (Newman, *et al*. 2020). They are subjectively perceived to violate procedural justice by reducing considerations to easily quantifiable performance data and failing to consider performance holistically. (*Ibid*., 2020). Their results also indicated several possibilities for future works: Firstly, it could be said the combination of human evaluators and algorithms are perceived as the fairest compared with reviews only either by humans or algorithms. For example, Nagtegaal (2021) distinguished between high- and low-complexity decisions, and algorithms were viewed as fairer for low-complexity tasks, because humans are subjective and biased while computers were perceived as more objective. Secondly, it would be interesting for future work to investigate explicit messaging and information regarding consideration of demographic factors such as race and gender. If, for instance, an algorithm is explicitly framed as ignoring race (or, alternatively, assessing race and assigning different weights to particular categories), this may moderate perceptions of the fairness of the algorithm. (*Ibid*., 2020). The pilot research which reviewed systematically literature in the fairness of algorithmic performance evaluations revealed that the studies […] were almost exclusively conducted in Western democracies, predominantly the US. (Starke, 2021, 29).

Through reviewing pilot researches, it is confirmed that employees' fairness perceptions in performance evaluations are important since it affects a lot on their motivations, commitments and outcomes which have impacts on organisational performance. While there are several determinants regarding fairness, such as the accuracy of the information, consistency in procedures, and being free from the personal bias of decision-makers, this research focuses on the issues of evaluators' biases which could be solved by computer-based algorithms and their technologies in the futures.

This research is to add the comparative analysis of participants from different demographic backgrounds to Newman, Fast and Harmon's quantitative research based on scenarios. While their research was quantitative researches to explain the relationship between variables, this research concerns the qualitative aspects of how participants from intercultural backgrounds perceive fairness when their performance are evaluated, what verbal or nonverbal messages are observed. This qualitative approach enables to describe the complex situation, participants' attitudes and values which could be lost through a quantitative approach. For example, the fairness perception from a female participant is not independent of her perception as her ethnic group factor, which is difficult to analyse as variables in quantitative approaches. It is assumed that gender issues are related to cultural issues as well. Besides, it can be expected that a group consists of multinational participants would be influenced by the nationality compositions and nationality context. This is because nationality context steered attitudes (social distances) and perceptions (conflict) of group members depending on how many nationalities or which nationalities were included in a group (Ayub and Jehn, 2018, 625). Since how honestly or comfortably each participant in the group can answer the questions depends on not only the relationship between participants and the author (the facilitator of the focus group) but also the relationship between participants, there could be observed influence of intercultural conflicts through the discussion in the group. This effect is also difficult to be taken into account and observed in a quantitative approach.

In light of the above, this research values how those messages and information collected in qualitative ways could be analysed to know the significant factors to improve fairness in performance evaluation at the workplace which includes different demographic workers by utilising algorithmic evaluation technologies in the future.

2. METHODOLOGY

This chapter gives an explanation of the methodology used in this research which includes research design, method of data collection, sampling procedure, data collection, and data analysis.

2.1. Research Design

The main research question of this research was how the employees from different nationalities, gender or backgrounds would perceive fairness about performance evaluations when they imagine algorithms would join their performance reviews? The objective of this research was to investigate messaging from participants' feelings, mindsets and attitudes about fairness perceptions when they imagine algorithms join their performance reviews, regarding consideration of demographic factors.

The nature of this research is a combined study of exploratory (Saunders, *et al*. 2015, 174) and evaluative research (*Ibid.*, 2015, 176) considering the characteristics above mentioned, to clarify the difference of fairness perceptions and find out how a way of human resource evaluations for different demographic factors can work well as organisations' issues. Besides, a mixed-method "qualitative experiment" (Robinson and Mendelson, 2012, 332-333) is used in this research to study participants' responses regarding fairness perceptions, rather than to examine the causal relationship between particular variables. Importantly, for example, the difference of perceptions in gender is not independent of the difference in ethnic groups. According to Robinson and Mendelson (2012), "qualitative experiment" is "a mixed-methods technique that uses qualitative strategies such as focus groups to capture the differences in the processing of meaning construction between groups in a single phase of experimental execution." (*Ibid.*, 332). Table 1 shows the general path a hybrid-technique (a mixed-method) design might take (*Ibid.*, 336). This was introduced in the works by Robinson and Mendelson (2012) as a sample of experimental qualitative technique, which consists of four stages. Phase I is a phase where a survey or open-ended questions are executed to measure participants' general preferences and attitudes. Phase II is described as a stage of experiments or observation to collect participants' data by either a

quantitative or a qualitative method. Phase III is a phase that enables a researcher to have interaction with data collected in Phase II, where additional qualitative data is collected through in-depth interviews or focus group to analyse deeper and richer meanings. The last phase as Phase IV is to have a survey or open-ended questions to collect information about whether participants reactions were changed through the process of the experiment, observation, In-Depth interviews or Focus groups.

Table 1. Sample experimental qualitative technique, hybrid-method design

	PhaseI	PhaseII	PhaseIII	PhaseIV (Optional)
Phase of Method Execution	Pre-Test	Application of Stimulus	Post-Test Interaction	Post-Test, Addendum
Sample Technique Per Phase	Survey, Open-ended questions	Experiment, Observation	In-Depth Interviewing, Focus Groups	Survey, Open-ended questions
Sample Question	How often do you visit the following news web sites?	When presented with unfamiliar news presentations, what actions do subjects take?	What are you thinking as you look at this? What in this news presentation do you consider most credible?	Have your attitudes changed?
Analysis Technique	Frequencies, Inferential Statistics	Discourse Analysis, Inferential Statistics	Discourse, Textual, or Content Analysis	Textual Analysis, Frequencies, or Inferential Statistics

Source: Robinson and Mendelson (2012, 337)

According to Robinson and Mendelson (2012), "This four-phased design should be modified according to the characteristics of the project as well as the time and other resources of the researchers. Phases might need to be expanded, curtailed, emphasized, integrated, or eliminated completely." (*Ibid.*, 336). The important point in using the technique of this mixed-method could be said to design the best way for the characteristics of the project, time and other resources which researchers could use, to analyse participants' data with a qualitative tool that enables participants to exchange their open opinions and attitudes to create a constructive process in the research.

As per the technique and conditions introduced by Robinson and Mendelson (2012), the design of this research combined a pre-survey to collect demographic and background information of participants and the focus groups to collect qualitative data from participants (Phase I and III in Table 1). The main method of collecting data is focus groups to observe and analyse both verbal and nonverbal response of participants from different demographic groups. Focus group interviews as a method that allows respondents to answer verbally in an extended, and often spontaneous, manner to open-ended questions so that researchers can compare differences among groups (Krueger & Casey 2000; Schutt, 1999 referenced in Robinson and Mendelson, 2012, 338).

2.2. Method of Data collection and Samples

In this research, focus groups are conducted to collect qualitative data, considering the characters of this research. This research is to add an analysis of different groups of participants from different demographic backgrounds including Japanese workers, who experienced one of the most unclear performance evaluations in their workplace focusing on qualitative aspects of employees feelings and attitudes in fairness perceptions. Considering the pandemic situation due to the COVID-19, all interviews are taken online through web-meeting tools, while this online method created fair conditions between participants who live in Estonia and Japan.

The participants of the focus group are assigned who fulfil the following requirements:

1) Those who have work experience of at least 6 months.
2) Those who have studied in universities or currently studying.
3) Those who can communicate in English or Japanese

The following participants are assigned to the focus groups based on the legal citizenships.

1) Estonian citizens (three groups)
2) Internationals (two groups)
3) Japanese citizens who have work experience in Japanese domestic companies (one group)

Participants in each group are assigned based on their legal citizenship, while it is predicted their nationalities as their identities are more diverse than their citizenships. The first, second and third groups consist of Estonian citizens who have worked in companies, firms or organisations. The

second and the third groups consist of international participants from different citizenships which can be compared with Estonian citizenship's participants. The common characteristic among international participants is that they also have work experiences of a minimum of 6 months, and they have studied abroad who have already experienced intercultural issues through academic places, daily lives and workplaces. The last group consists of Japanese citizens, who have worked in domestic Japanese companies for at least 6 months. In order to observe and analyse the influence of work experience in domestic Japanese companies' cultures which are largely different from those in companies in the other countries, this group does not include Japanese who have never worked in any domestic Japanese organisation for at least 6 months.

To avoid the bias of recruiting channels for participants, several different channels are used to recruit participants into the focus groups, which are through student organisations in the universities in Estonia, being introduced by local Estonian citizens, international communities in social networks (WhatsApp and Facebook) and private network of the author.

2.3. Procedures

Adapting the scenarios used in the surveys and experiment in Newman (2020)'s research, focus groups in this research also uses the measures where the participants were asked to read a scenario about an organisation where a new policy of performance evaluation assessed by computer-based algorithms is going to be introduced, then imagined they are the employees of the organisation to be evaluated, and answered questions regarding how they perceived fairness in this evaluation policy.

Since the methods of research and participants' backgrounds are different from those in the previous research, the measure is differentiated and customised in the following ways: Firstly, in the scenario, the performance evaluation policy of a department is described, where both a human manager and an algorithm join the performance reviews to evaluate the employees. Secondly, the scenario is shared with participants before the group interviews so that they can read the scenario in advance to have time to consider it. At last, questions to be asked to the participants are more qualitative about the process, the outcome of evaluations and what factors they think important about fairness in evaluations.

As per the principles above, participants were asked to do the following things:

1) Participants fill an online form about their demographic information (gender, citizenship, nationality), as well as academic backgrounds and work experience.
2) Participants read a scenario about performance reviews to imagine themselves as an employee going through the review process.
3) Participants were asked questions regarding the scenario or the topic in the online group interviews to answer their thoughts, feelings, or interactions with the other participants.

The scenario participants read is originally prepared by Norman's research (Newman, *et al*. 2020, 157,158) and edited by the author for this research as the following:

A SCENARIO OF A NEW PERFORMANCE EVALUATION POLICY
An algorithm joins the performance reviews

Please read the following scenario.
Imagine you are an employee of the organisation to be evaluated.
Think about fairness.
(The scenario might be lacking some information that is important for you.)

You are an employee of the Sales Department in an organisation.
Your organisation adopted a new performance review practice in which they evaluate individual performance data over the previous year as well as provide an opportunity for employees to articulate what they saw as their biggest challenges and contributions throughout the year. For the latter, employees record brief videos of themselves explaining what they achieved, what challenges they overcame, and what they learned during the previous year. This gives employees the opportunity to convey their thoughts and, if needed, put their performance data in context.

Employees' performance reviews are conducted by [their manager and an algorithm]. The [their manager and algorithm] evaluates the following data: the employee's total sales, customer experience survey results, duration of employment, amount of overtime worked, contributions to coworkers' sales, and potential for strong performance in the future. In addition, the [algorithm] is trained to evaluate the employee-recorded videos, paying specific attention to the employees' stated contributions, the lessons they say they have learned, and their nonverbal communication

(i.e., facial expressions, mood, positivity, persuasiveness, and honesty). After evaluating the data and the video recording, the [their manager and algorithm] come to a final decision on the performance review, which could affect the employee's salary, bonus, eligibility for promotion, and, in some cases, dismissal.

*Assessment algorithms are created and checked by the well-diverse team including engineers, HR professionals, managers, your direct boss, those who have deep understandings about the job in your positions, and check members from various nationalities and gender).

(The original scenario is Newman's research (Newman, *et al*. 2020, 157,158), the author edited for this research).

The focus groups are conducted as semi-structured interviews, where participants are asked some questions and the reason why they feel that way, as well as the interactions with the other participant's answers. The length of the focus group is about 50 minutes. The questions are not shared with the participants before the interviews. The prepared questions are about when participants felt "unfairness" regarding performance reviews in their experience, what unfairness can happen, fairness in performance reviews for the participants, which part of the performance review in this scenario is NOT fair compared with a human manager's performance review, and what kind of unfairness the participants think can be avoided or reduced if the algorithm joins the performance reviews. The order of the prepared questions and additional questions to be asked to the participants are changed depending on the participants' responses in the previous questions and interactions. In order to avoid misunderstandings by the influence of pronunciations, accents of the language, and the quality of the audio of the web meeting tool, the questions were written in the chat-box of the meeting tools and PowerPoint slides which were shared during the interviews.

The results of the interviews are recorded with permission by the participants, while the recordings of the interviews are only accessed by the author of this research for the limited purpose of analysing fairness perceptions for this research only.

2.4. Data Analysis and Ethics

The data collected through the focus groups are analysed with a combination of Thematic Analysis (Saunders, *et al*. 2015, 579) and Content Analysis (*Ibid*., 608). A quantitative description of participants' messages would be analysed through Content Analysis. In this analysis, explicit messages such as participants' answers to the questions and interactions with the other participants during the interview are coded and transcribed to be analysed as "manifest content" (*Ibid.*, 609). In Thematic Analysis, data were coded to identify themes or patterns to analyse participants' perceptions about fairness and unfairness in performance reviews, participants' feelings and attitudes about what kind of unfairness can be reduced or avoided with the performance reviews written in a scenario, and which parts participants fear about the performance reviews with algorithms. There could be similar themes in several groups, as well as different patterns in a certain group. This analysis method enabled analyse of free discussion that arose in the focus groups.

It is essential for the research to establishing ethics in collecting and analysing data. The research design should not subject those you are researching to the risk of embarrassment, pain, harm or any other material disadvantage (Saunders, *et al*. 2015, 201). There could be a risk of embarrassment for participants when they answer the questions, regarding their experience of being evaluated in unfair ways in the past. Therefore, permission from every participant is obtained to record the focus group, and the recording is accessed only by the author of this paper, for the limited purpose of analysing data only for this research. Besides, it is emphasised in each focus group that participants need not explain personal experience if they do not want to share during the group interview.

2.5. Limitation of the Methodology

This methodology has limitations. Firstly, the limitation as a qualitative methodology. The result of these interviews cannot be generalised as representatives for properties or features about a certain citizen. Especially, international participants who are studying or working in Estonia are not the majority or mainstream of the citizens in their home countries. Secondly, since the method of group interviews, participants would not always answer their honest thoughts or opinions, caring the responses and interactions with the other participants and the author as an interviewer. The responses are influenced to some extent by the human relationships between the participants

and the author as an interviewer. Thirdly, the limitation of scenario-reading based qualitative experiment. The AI-driven assessment technology began to be used mainly in assessing job interview videos to screen applicants and have not commercially used in performance evaluations for employees in organisations. As a character of this research based on scenarios, participants need to imagine the situation well by reading the scenario as a thought experiment. Thus, a participant's perception could depend on reading skills, thought skills for an abstract topic. At last, the limitation regarding the language used in the interviews. The interviews for Estonian citizen groups and the international groups were conducted in English, which was not the mother tongue of any of those participants. If similar focus groups are conducted in their mother tongues, more various participants with work experiences could join, and the participants could answer with more relaxed attitudes with free opinions during the interviews. In this viewpoint, the Japanese participants' group have a kind of advantage since they could listen, think, communicate, and discuss in their mother tongue which enabled them to share their opinions in more comfortable ways.

This research is not a holistic view of fairness perceptions, rather, the research supports the previous works by adding qualitative analysis of text and behavioural messages of participants from different demographic factors and backgrounds, regarding the feelings, mindsets and attitudes about their fairness perceptions. This research is executed to provide information for the process of developing algorithms for performance evaluations so that the implementation of new technologies would be appropriate for both the organisations and employees in ethical ways, in the future.

3. THE FIELDWORK RESULTS AND ANALYSIS

The aim of this work is to investigate messaging from participants' feelings, mindsets and attitudes about fairness perceptions when algorithms join their performance reviews regarding consideration of demographic factors, such as nationality, gender, and backgrounds. In this chapter, the result of the online interviews are organised, reflected, analysed and discussed.

3.1. Demographic factors and backgrounds of participants

As per the plan of the methodology, 6 groups which consist of different citizenship participants, a total of 31 participants were assigned for online focus groups. 14 out of 31 participants we male, 17 participants were female. International participants were mainly NON-EU nationals which reflect the recent compositions of international students in universities in Estonia (studyinestonia, 2020) to some extent. They are from Southern Asia, Western Asia, Eastern Asia, Western Africa, Northern Africa, Eastern Europe, and Northern America. The concrete name of the citizenship of international participants refrained from disclosing since it makes it easy for certain stakeholders to assume the individuals. All participants in the Japanese citizen group became male by chance, while this sample limitation is discussed in the limitation of this research. Most of the participants have more than one year of work experience, while 3 participants had work experience less than 1 year but more than 6 months.

In this research, "citizenship" means legal citizenship which an individual can confirm based on his passport, for example, while "nationality" means one's identity. The participants were assigned through different channels: The author's private networks, introductions by the author's friends, recruiting on SNS, and the student organisations of Tallinn University of Technology. The demographic factors, academic backgrounds work experience of the participants in each group are shown in Appendix 2.

Since some participants prefered having individual interviews to join a group interview to share their thoughts and experience more anonymously, the author took the individual interviews as well.

Though the data from the individual interviews should not be compared directly with focus groups because of the difference of conditions, they could be additional in-depth data to help to describe and analyse the topic of fairness perceptions in different cultures.

3.2. Results of the interviews

Table 2 to table 5 shows the core questions asked to participants and their responses about their thoughts, feelings and mindsets in each interview, using Thematic Analysis, where data were coded to identify themes or patterns to analyse participants' perceptions. Though the interviews were semi-structured and questions and the order of questions were different from one group to another depending on the participants' backgrounds, the relationships with the author, those 4 core topics were commonly mentioned and discussed in all groups and individual interviews. ● means that at least one of the participants in the group mentioned. ▲ means that at least one of the participants mentioned, while they told that they did not regard that point as an important factor of fairness or unfairness. F is not a group interview but individual interviews due to the participants' preferences and conveniences.

Table 2. The first question and participants' answers about their perceptions about the unfairness

Citizen category	Estonian			Internationals			Japanese
Interview groups	A	B	C	D	E	F*	G
Question: (In general) When did you feel "unfairness" in your work experience? *If you have never felt unfairness, what unfairness do you think can happen?							
1. Personal biases of the evaluator	●	●	●		●	●	▲
2. Disadvantages for less-experienced workers	●		●	●			
3. Unfair task allocations for the same position	●		●				
4. Disadvantages for minority groups			●	●			
5. Different salary for the same positions			●		●		

	6. Political involvement						●	
	7. Unfair in setting target goals							●
	8. Unfair between different projects							●

Source: Recordings of the focus groups and the author's notes in the focus groups (2021)

*F was not a group interview, conducted as three separate individual interviews.

In the question about fairness in performance reviews or the workplaces, it was observed that the participants in several groups told that they have never experienced unfairness in their work experiences. They, however, mentioned several unfair performance reviews in the discussion of the scenario. Most participants agreed that the personal biases of the evaluators would lead to unfair evaluations regardless they told that they actually experienced it in their workplaces or not. Though Group D did not mention directly using the word "personal bias" or "the evaluator's likes or dislikes", one participant shared the example of unfair treatment for a minor ethnic group, and this point could be largely related to personal biases of the evaluators. Group G also regarded it as unfair, they emphasised this unfairness is not something to reduced or avoided since building a good and advantageous relationship with the evaluators is an important skill to be evaluated in performance reviews. The unfairness of less experienced workers reflects the situation in many countries where it is hard for new graduates or young workers to find a job to build their career. The Japanese citizen group did not mention the issue of less-experienced workers, which indicates the recruiting system of domestic Japanese companies.

It was observed that the unfairness of political involvement was mentioned only in individual interviews. Besides, participants in the individual interviews did not mention the unfairness of political involvement in the first question at the beginning of the interview.

Table 3. The second question and participants' answers about their perceptions about the fairness

	Citizen category	Estonian	Internationals	Japanese

Interview groups	A	B	C	D	E	F*	G
Question: (In general) What do you think is fairness in performance reviews?							
1. Free from personal biases of the evaluator	●	●	●	●	●	●	▲
2. Clear job descriptions, agreeing with target goals at the beginning of the period	●	●	●	●			●
3. Comprehensive assessment (Quantity + Quality, regional context, processes+ results, .etc.)	●		●			●	
4. Apply the same evaluation rules		●	●		●		●
5. Opportunity to get feedback and to share opinions		●	●	●			●
6. The same salary for the same positions			●		●		

Source: Recordings of the focus groups and the author's notes in the focus groups (2021)

*F was not a group interview, conducted as separate three individual interviews.

Most groups mentioned being free from the personal biases of the evaluators is important for fairness in performance reviews, as well as the difficulty to achieve it. The participants in Group G did not point it out spontaneously, and when the author as an interviewer softly asked them "What do you think about unfairness related to personal biases of biases"? the participants admitted it as an unfair thing, while they mentioned that building up a good relationship with your evaluators is an important skill to be evaluated in the performance reviews rather than unfairness to be eliminated or reduced. It was also observed that several groups mentioned that it is an important factor for fairness to have clear a job description for each position, and the employees' agreement about their goals with the evaluators at the beginning of the review period. The importance of quality assessment is emphasised that a quantitative assessment way which counts how many tasks completed have the big issue of quality of services which influence a lot the company's reputation and trust by customers or consumers. Applying the same evaluation rules, the same salary for the same positions are mentioned in several groups, while participants also regard considering different conditions of regions, personal backgrounds are important for fair performance reviews. Opportunities to get feedback from the evaluators and share employees' opinions with the evaluators are observed in different groups. In their understanding, these opportunities give the employees to understand what the organisation require from a certain

position, the reason they got a certain evaluation, and justify their performance in the period including uncontrollable factors outside of the organisation. Those are the opportunities for the employers as well to understand the reasons in the performance and predict the organisational result in the period.

Table 4. The third question and participants' answers about their perceptions about unfair images of performance reviews with algorithms

Citizen category	Estonian			Internationals			Japanese
Interview groups	A	B	C	D	E	F*	G
Question: Which part of the performance review of a human manager + an algorithm is NOT fair compared with a human manager's performance review?							
1. Algorithms cannot evaluate qualitative factors (employees' soft skills, supportiveness, teamwork, process, environmental factors, etc.)	●	●	●	●	●	●	
2. Risks of employees' manipulation of the algorithm	●	●					●
3. Employees won't get feedback from, or discuss with the algorithm		●	●	●			
4. Some employees are not comfortable talking to the computers	●	●	●				●

Source: Recordings of the focus groups and the author's notes in the focus groups (2021)

*F was not a group interview, conducted as separate three individual interviews.

Most participants in different groups considered it would not be fair that algorithms won't evaluate employees' soft skills such as openness, motivated attitudes, supportiveness for other workers in the organisation. Some participants mentioned that if the algorithms cannot evaluate the process of how the employees' achieve the goals, it is not fair. Also, the environmental changes with the example of the COVID-19 crisis were pointed out as one of the things algorithms cannot evaluate fairly. Several groups considered it is unfair that the algorithms cannot give feedbacks to the employees, which is an important task for human evaluators. Some participants insisted it is necessary to have human interactions, where the evaluators or managers should have strong

interpersonal skills. Besides, Several groups discussed that some employees will manipulate the algorithms by pretending better persons than actually, they are. In their understanding, human evaluators are smart enough to detect those lies.

Table 5. The fourth question and participants' answers about their perceptions about fair images of performance reviews with algorithms

Citizen category	Estonian			Internationals			Japanese
Interview groups	A	B	C	D	E	F*	G
Question: What kind of unfairness do you think can be avoided or reduced if the algorithm joins the performance reviews?							
1. Free from personal biases of the evaluators	●	●	●		●	●	▲
2. Rationalise the evaluation process	●	●	●	●			●
3. Objective information from the algorithm for evaluations				●	●		

Source: Recordings of the focus groups and the author's notes in the focus groups (2021)

*F was not a group interview, conducted as separate three individual interviews.

Participants got similar answered regarding the unfairness which can be reduced if the algorithms join the performance reviews, which are being free from personal biases of the evaluators, Rationalising the evaluation process, and getting objective information from the algorithms. Most groups mentioned directly that the personal biases of the evaluators are reduced if the algorithms join the performance reviews. The participants in Group D did not directly mention it, while the participants mentioned the fairness of applying the same rules to everyone, which also indicate the reducing the personal biases. In this question, Grop D emphasised the negative impact of lacking humanity in evaluations, which is not good for the employees.

There were some different free discussions or comments in each group other than the core topics mentioned above. They mainly talked freely about what can improve fair performance reviews. Free discussions are summarised in Table 6, 7 and 8.

Table 6. Summary of free discussions which took place in Estonian citizen groups

Other discussions		
Estonian citizens		
	Group A	Team evaluations ・It is important to involve team members who know your everyday jobs well, not only the manager. ・Feedbacks from both your boss and your team members in the same levels (positions) are important.
	Group B	Training for managers ・Teach managers how to be open-minded, how to deal with people. You should not care about personal biases ・There is no way to eliminate unfairness ・Everyone needs different types of managers
	Group C	Improvement of interviews ・The evaluation in this scenario can be improved if the interviews became person-to-person interviews with support of professionals with psychological knowledge

Source: Recordings of the focus groups and the author's notes in the focus groups (2021)

Estonian participants had a different discussion about how to increase or improve fairness perceptions. Team evaluations are the common point that the international groups reached as well. Participants with relatively fewer experiences in Group B were often in the discussion of the algorithms or anything that can not eliminate unfairness, rather than the discussion to reduce it. Based on the unfair risks of the algorithmic evaluations, a participant in Group C suggested improving the interview process to involve the professionals with psychological knowledge so that they can support and encourages some employees to take it honestly.

Table 7. Summary of free discussions which took place in international groups

Other discussions		
Internationals		
	Group D	Do what you are expected to do by the organisation ・It is very important to get feedback about what they expect you to do by the organisation. Escalation to the higher authority, or the team for inclusive ・Escalate unfair issues to the higher authority About gender /racial discrimination ・If you are doing attacks according to what they expect you to do, there is no way your manager will keep on giving you a low grade
	Group E	Team Evaluations ・Bosses should be also evaluated by subordinates. Regulations by the government ・Government should also introduce some laws in order to regulate Protesting as Labour Union ・Strike to protest unfair managers

			Escalation to the higher authority, or the team for inclusive ・The company should have a department for employees to consult about unfairness. Risk of Escalation ・There is a risk that the consulted information will be told to the boss. Training for managers ・The managers should be trained so that they will just be aware of the procedures that go in their mind that they will be biased.

Source: Recordings of the focus groups and the author's notes in the focus groups (2021)

Group D emphasised that the employees can argue with the human evaluators but cannot with algorithms. Both Group D and Group E had opinions that the employees can escalate unfair issues to the higher authority, or the special department for inclusiveness, while Group E soon pointed out the risk of escalation that it is not safe for the employees to escalate because their opinions about unfairness would be shared with his boss or superiours, then he would get the negative impact on his careers in the organisation. A participant in Group D mentioned that if the employees do what is required from the company, the managers do not have reasons to evaluate him with low grades. Group E also discussed team evaluations are better and fairer than performance reviews by the only direct boss of the employees.

Table 8. Summary of free discussions which took place in the Japanese citizen group

Other discussions			
	Japanese citizens		
		Group G	Importance of understanding what is required, and do it · Some unfairness lies in the subordinates' misunderstandings. Those subordinates are not doing the right jobs in the right directions which their superiors and the company want them to do. Evaluator's personal biases are not unfairness · People assess the person whom he or she likes better than the others. In a company, this kind of political activities is important in some ways, after all. · It can be helpful if you let all employees know and realise that playing the game in office politics is the material you are evaluated based on. And the cosying up, bettering up, and sucking up to your boss is one of the evaluations indicators. Task allocation between humans and algorithms · Let algorithms do the part which can be affected by personal opinions, like or dislike. Then what we could agree as an organisational direction should be treated by human managers. · What the employees are obliged to do in the period can be assessed by the algorithms. The aspects related to the employees' motivation, potentials in the future should be evaluated by human managers. Team evaluations, multiple axes for evaluations · Increase the number of evaluators, other than your direct boss.

			· The employees in the same layer, the same position with you also should evaluate you to create fair evaluations ・ Have multiple axes for evaluations would be fairer Cons of Team evaluations · If you introduce this system into your office, the atmosphere of your workplace would become uncomfortable with superficial relationships. Everyone tries to butter up, suck up to each other in order to get a good assessment. Everyone makes great efforts to groundwork for internal matters. · This system won't solve the issue of personal likes or dislikes

Source: Recordings of the focus groups and the author's notes in the focus groups (2021)

The Japanese citizen group also reached the idea which is similar to team evaluations, where they thought increasing the number of evaluators and axes of evaluation criteria would be fairer performance reviews. This group, however, pointed out the negative aspects of team evaluations that every employee would spend their time more on internal evaluations buttering up to each other to get better performance review results.

They have some different perceptions about fair performance reviews. Firstly, regarding personal biases of the evaluator, emphasising it is significant for the employees to have skills of butter up to their superiours to build good relationships with them. Secondly, they mentioned that it is important for an employee to understand what the organisation is requiring from the employee, what job they should do, indicating that they need to understand without directly told by the superiors. "From the evaluator's perspectives, I don't want to assess good if the subordinates insist that they are making efforts in some jobs which the evaluator is not requiring. I have often felt 'why this subordinate is doing such a job?' He or she is doing unnecessary jobs, which is not required at all, and saying 'I'm busy'. They cannot be assessed well if they did not do what the superior want them to do. … What the superior want their subordinates to do are usually not understood well by the subordinates. (G3)"

While participants talked some common responses, some different discussions were observed in different groups. In the next section, those common points and differences are analysed.

3.3. Analysis of the results

Based on the results which were confirmed and organised in 3.2, what was found and analysed in this section. Some common key themes were mentioned during the discussions in several groups, while there were different opinions mainly depending on participants' experience and gender.

The difficulty of discussing fairness

In several group interviews, one or more participants began to talk that he or she had never felt unfair in the workplace in the past at the beginning of the interviews.

"I don't have any unfair experience in my works. (B2)" "In general, I have never felt that I have been evaluated unfairly in my career. I have worked in many companies and almost everywhere, if you do your job right, people do right by you." (D1). "I think that I was lucky, I don't recall in a situation like this" (E3). "So far, I have been working, I have not experienced unfairness." (E2). "I don't think I ever had those kinds of experience…. But what possibly could be? I'm not even sure." (F1).

It was notable that some participants shared the example of unfairness in their workplaces when they began to comment and discuss the scenario about how algorithms can help fairness in performance reviews or not.

"…people get hired because of various reasons. They get hired not because of their talent, they get hired because of some just random reasons, like maybe the manager is your uncle or something" (D1). "Sales Department persons will be judged based on the number of customers they bring in, while receptionists will be judged based on how they smile. Two different things. So if the receptionist or the person that sits in the office is close to the boss. Right? He is still going to get more advantage than the person going out, because it will only account for results. So your efforts will not be compensated. But somebody else that doesn't have to do with numbers, their efforts and their smiling will be compensated." (E2).

It was observed that the unfairness of political involvement was mentioned only in individual interviews, where they did not mention the unfairness of political involvement in the first question at the beginning of the interview, either.

"…well, let's say, experience in my home country is different. I know that some people get promoted just because they are good friends with the CEOs. Yeah, it's really bad. Or sometimes because of political questions. …Sometimes they win over the experience, over results, over everything. And this is what, well, oh, it's sad. I hope that it changes." (F1) "Okay, if you have somebody, […] an acquaintance with the person that evaluates your performance or decides about your salary, or even in another higher position than yours. Then, it can affect your salary to even your position, too. And evaluating process, and even your whole job career." (F3).

Those examples can be completely imaginary ones, or the ones which were talked about based on the participants' experiences or what they had seen in their workplaces regarding their colleagues. It could be said they shared the examples of unfairness in their workplaces through the reasons why and how they feel algorithmic evaluations can be fairer or not fairer compared with humans' evaluations. This indicates how difficult it was for the participants to share their personal perceptions about fairness or unfairness in front of others. Especially, the difficulties to talk about topics related to gender discriminations or racial discriminations were observed in all groups. This could be assumed with a common human interaction that no one wants to be misunderstood by the others that they were complaining about unfairness to make excuses about lacking abilities, efforts in the right directions and misunderstand what the company or the superiors are requiring. This was illustrated by some participants who have management experiences:

"The responsibility of the unfairness lies on those who feel unfair, too." (G4)

While this is a significant aspect for both employees and managers, it can be analysed it is also significant and healthy to have opportunities, place, rights to discuss fairness to create a better work environment since history shows labours could gain rights to work better only after they insisted, protested fought for their rights with the business owners.

The importance of fairness in performance reviews

While it was hard for participants to share their perceptions of unfairness, most participants who joined the interview mentioned the importance of fairness in performance reviews or agreed with those opinions. It could be said that fairness in performance reviews are perceived as important regarding motivation, not only for financial compensations, but also for different factors. The quote below was a participant's comment about her experience in a part-time job:

"…how come every time there was an evaluation, there was someone else is the employee of the month, where they don't even perform. … Although it's not a big thing, I just feel it would still affect employee performance. We would still be demotivated, although we don't want to make a career out of it." (D2)

Another quote from a full-time job worker:
"The evaluations will lead to the compensation for the employee, and the employees who go high evaluations will get an interesting job or project in his company. Though I might not have all information on what they will use the result of evaluations for, I think I will have a lot of advantages if I was garnered high acclaim. Then, if I cannot garner high acclaim by an unfair performance evaluation, I would feel hey what the hell this is." (G2)

The unfairnesses that can be reduced or avoided by algorithms

The unfairness which the participants pointed out as something that can be reduced or avoided when algorithms join performance evaluations were similar to each other. To summarise, to be free from biases of the evaluator, to rationalise the evaluation process, and to acquire objective information regarding performances were perceived as unfairness which will be reduced by using algorithmic evaluations.

Perceptions of personal biases of evaluators in the Japanese group

It was observed that only the Japanese group mentioned that personal biases are not something we should avoid, though they also regarded evaluators' personal biases as unfair.

"…humans and humans. People assess better the person whom he or she likes. In a company, this kind of political activities is important in some ways, after all… Playing the game in office politics is one skill." (G4)

This point reflects and indicates the unique features or characteristics of the hiring and HR system in domestic Japanese companies.

"Well, we don't have job descriptions, missions for a position in Japanese domestic companies." (G3). "The evaluation rules seem to differ from one company to another. The company which are closer to 'the village society' in Japan has that example, and once you are disliked by your boss or

other employees, your assessment results would be decreased dramatically… Domestic Japanese companies in general, are 'the village society'". (G3)

As it was reviewed in the literature chapter in this research, Japanese domestic companies have largely different HR systems and rules from any other country. They hire people first, then allocate jobs to people, thus, they do not have job descriptions. In this system, the scope of one employee needs to do is unclear and ambiguous, inevitably the criteria of evaluations are also unclear. Under this system, employers need to hire almighty employees who can work on any job in various fields by learning autonomy without instructions. Besides, it is extremely tough for the employees to change the company because the employees are transferred to one department to another in a completely different field such as Sales, IT, Accounting, Productions, once in 2 years or 3 years under the job rotation systems, where they cannot have cross-company specialisation. In these circumstances, it is essential for the employees to build good relationships with their bosses, colleagues, and those relationships matter largely their performance reviews, job allocations, work conditions.

While building good personal relationships with superiors is an essential skill and should be highly evaluated in performance reviews, there is a question of whether it is good if this skill overcomes all other skills or experiences. The fact that a skill of building good relationships with superiors is important is not the same with that this skill takes precedence over all other skills and experiences. Besides, the importance of buttering up to the evaluators for this Japanese group seems not to consider the discriminations against minority groups such as gender discriminations, racial discriminations. For instance, if a boss has a strong prejudice against female workers or East Asians, it is doubtful that an East Asian female subordinate can build good relationships with the boss to be promoted, no matter how hard she made efforts to butter up to the boss.

The differences between fairness factors and what participants need for evaluations

In each interview, participants discussed what is fair performance evaluations or unfair evaluations at the first half, then move on to the discussion regarding the scenario where they were asked "Which part of the performance review of a human manager + an algorithm is NOT fair compared with a human manager's performance review?" In the first half, most participants mainly mentioned being free from personal biases of the evaluator, clear job descriptions, agreeing with target goals at the beginning of the period, applying the same evaluation rules, opportunity to get

feedback and to share opinions, the same salary for the same positions as important factors in fair evaluations.

"…fairness can only be charged if you are being judged by your merits. What you can bring to your position, if you can complete projects, get along with people and be constructive in your work. And if you can charge this, not qualitatively, but quantitatively…" (A2). "…for every job, you need to have a clear description of what you need to do,… you have to have quantifiable metrics in order to judge somebody's performance." (D1).

When they were first asked what fair performance reviews were, only three participants in Group A and C began to mention that it is important for fairness in performance reviews to consider qualities and contexts, They were all female workers. The first type of opinion cared about the difference of experiences.

"Fairness is when you're judged not so equally to others in your company, but based on your personal criteria. For example, because the age might be the problem, experience might be the problem" (A5)

It was notable that one participant mentioned the unfairness in an intercultural environment where the regions in Europe shared similar historical backgrounds and values, while other regions outside of Europe have totally different conditions. This opinion was observed uniquely by this participant who had backgrounds in sociology.

"…it seems like it's a fair play because everyone has its own region and they have the same evaluation or performance review criteria. But as the region in their nature are bit different… In evaluation, for example, it's more easier to do business with Europe because we share the same culture… And the fairness is that, based on the same grounds, the specific details have been considered in evaluations. This is what I would say is fairness." (C5).

The third type described the issues in how companies can maintain or develop service-related jobs which were also supported by work experiences.

"…I think more emphasis should be on quality, even with numerical goals. For example, in customer service when you have to solve a certain amount of tickets, let's say, there is one person

who does all the required number, but answers in a very vague or offhand way. But there's someone who really puts their mind to it, and does, let's say a bit less tickets. But this is not taken into consideration, even if it contributes to the reputation of the whole company and customer satisfaction overall." (C4).

However, when they were asked "Which part of the performance review of a human manager + an algorithm is NOT fair compared with a human manager's performance review?", most participants began to talk about the importance of taking qualities and contexts into account in performance reviews, and algorithms, Artificial intelligence, or robots cannot consider or assess those qualities and contexts. The first type was about the uncertainty of algorithms' ability to assess qualitative aspects such as motivations, the process, conditions, and the context of the works and trust towards human managers' ability.

"…human manager could understand if the person is motivated or like lets' say emotion is stable" (A4). "…algorithms cannot always understand fully the background situation." (E3). "The algorithm cannot count the obstacles that workers face on the way to achieve that goal." (E5). "the employee put a lot of efforts and she is trying to work hard, maybe the result is not good, but from my point of view, we still get some encouragement to the employees." (E1). "algorithms cannot evaluate the effort of scientists or researchers, who kinda invest their time into R&D. So, because you know new inventions require time, they require ideas, discussions. Algorithms just cannot evaluate" (E5).

The second type of opinions was about cross-team or cross-organisational supportiveness.

"…those people that are very easy to speak to who are very supportive and work very well informed about how the business works and what's going on …They help with other people and other people's tasks by helping them figure out what to do, who to talk to, what kind of approach to use and so on. But this is not something that can be charged by this algorithm" (A2). "…they were valuable for the team in order that the whole team would work … this is something that I guess that algorithm couldn't take into account when evaluating the performance." (A3). "…now the tricky bit nowadays, a lot of corporations want to try if their people towards synergy and team effort. And more and more corporations are trying to bring that issue in, this is totally in the hands of the personal relationship between managers and employees." (C3).

The last type described the problem of an organisation where humans could not design appropriately the rules or systems rather than the ability of algorithms.

“I got into an accident and I was so much under pressure and in this company, let me remind you that this company was actually using software to do assess, give you bonuses, give you salary increases in these kinds of things. Ok, if I’m late, so they had this rule… HR system automatically deducts 500 rupi from your salary. SO I just didn’t’ want to have that,… then I still went to the office and worked at night… What I’m trying to emphasise on that if you have a robot making decisions for you then people will get into these situations.” (D1).

Those opinions, feelings indicated that those participants felt that those qualities and contexts were not necessary factors for fairness, but the important factors for a kind of better performance reviews. These feelings are explained by a Japanese participant in the following ways:

“I think a little bit that fair evaluations are relatively cruel evaluations, at the same time. Sometimes, pursuing fairness cannot be very good for organisations, and for motivations of employees. It is fair, but not agreeable for employees.” (G1).

The word “agreeable” was translated by the author from the original Japanese word “Nattoku-kan”, which might not the same meaning. “Nattoku-suru” is a verb that describes a person’s feelings where a person feels inside his mind that he can understand this is reasonable, rational. This “Nattoku” feeling could be said as a factor employees hope to have to create better performance reviews and better work environments. Besides, it can be said whether the employees can have this “Nattoku” feeling in the performance evaluations about the qualities and contexts which were described by the participants largely depend on whether those employees have trustworthy relationships with their human managers.

Team evaluations and fairness perceptions

Several groups mentioned the team evaluations which can improve fairness in performance reviews during free discussions.

“…you should add a little bit more people rather than just one boss. Because usually, boss is not participating in the everyday team’s performances…. you would rather put someone who is involved in the process of your team, like the team lead of your team.” (A5). “…the example from

one of the Estonian Unicorn companies…. they have to have the feedback from their teams. So people from working in the same position in their team and if the feedback is positive, then this, this kind of like, saying that my performance is good and I'm not the only one thinking it." (A2). "…if we want to get rid of unfairness within a workplace, we need a team evaluation of each other. Even boss needs to be evaluated." (E4). "It is important that everybody evaluates everybody. The bosses should be evaluated by the subordinates. It should go round like that so I think when it goes around in the chance of it becoming fairer." (E2). "If you increase the number of evaluators, fairer evaluations rules we can have than only one manager evaluate." (G3). "The several managers assess you is good. And I have an additional idea. The employees can be evaluated by the others in the same layer, the same positions…. The unfairness usually generated among the same layers." (G4). "Multiple evaluators or multiple axes for evaluations can be good, in a meaning, anyway." (G2)

On the other hand, the team evaluation method is not perfect, has also cons and risks. Japanese participants quickly pointed those cons and risks out soon after they mentioned the idea of team evaluations.

"It would be fairer, but the atmosphere of the office would be uncomfortable, what to say? superficial relationships where they always try to find out each other's real intentions." (G1). "To get a high assessment from everyone, everyone started trying to suck up and butter up to each other. Everyone makes great efforts to groundwork for internal matters." (G2).

While this was an important con as one of the demerits for democratic systems and principles, it was observed that increasing the number of evaluators and have multiple axes of evaluations increase fairness perceptions in performance reviews in different cultural groups.

3.4. Discussion, Limitation and Possibilities for futures works

The research problem of this work was that there is a lack of information about the process of the employees from different demographic factors or backgrounds perceive fairness, express their opinions with the other workers, and the issues they care about when they imagine both humans and algorithms would join the performance reviews. This knowledge is important for those who take part in intercultural and diverse organisations because how they perceive fairness and what

factors they care about need to be taken into account when the organisations try to develop and introduce algorithmic technologies into human resource fields, and also to improve performance review processes themselves, where it would be indicated that which part of those processes human evaluators more value and focus on. As it was confirmed in the results of the interviews and discussed in the analysis part, while participants got similar responses about unfair points and improvement of fairness when they imagined both algorithms and human managers join the performance reviews, how they perceive fairness and unfairness were different. The difference was observed depending on the difference of participants' experiences including length of work experiences, intercultural experiences, and management experiences, while certain citizens have similar experiences because of the system of their home countries. Their important factors for performance reviews are increased when they were asked the question about the scenario, from the direct questions of "What do you think is the fairness in performance reviews?". The Japanese citizen group had different perceptions about unfairness and how to create a fair work environment, reflecting the fact that HR systems in domestic Japanese companies are largely different from the other countries. They indicated that fair performance evaluations are important, while sometimes the employees need more than fairness to feel agreeable in the work environment.

This research has methodological limitations and limitations in the compositions of participants. As it was mentioned in the methodology chapter, this research has methodological limitations as qualitative research, focus groups, a scenario-reading-based qualitative experiment, and language used in the interviews. The results of interviews cannot be generalised as properties or features of certain citizens since those participants do not represent the population of their home countries. The results of the interviews were influenced by the relationships between the participants, the participants and the author as an interviewer, the limitation of scenario-based-experiment, as well as language the author could use in the interviews. It might be interesting as future work to conduct quantitative research by surveys with enough random samples to investigate if what were observed in this research can be justified as the tendency of a certain gender and regional cultures.

Since it was hard to reach out to participants from random samples in a population in the pandemic situation, most of the participants were from the authors' private networks, hence it was assumed that they avoided the answers which would make the author feel uncomfortable. Besides, the Japanese citizen groups became highly homogeneous, where all participants were males, mid-careers, generalists in domestic Japanese large corporations, had over 10 years of work experience as a system engineer. While male, generalists and having enough work experiences are good

quality samples to describe the main features of human resource systems, rules, and cultures in domestic Japanese companies, the voice of female generalists who had worked in Japanese domestic companies as a minority group could not be collected. Future works can have three groups of all-males, all-females, and mixed groups to observe the common points and differences in how the discussions would develop.

It was notable that some Estonian citizens were confused about what to do in these focus groups. This issue was not observed in international groups, Japanese citizen groups, or international candidates who talked with the author but could not join the group interviews because of the schedule. An Estonian citizen refused to join the interview for this reason. "…what is this all about? I can't really provide any answers based on experience, just reasoning and thinking out loud at best." (An Estonian citizen whom the author asked if he can join a focus group for this research). Though the author took time to explain the aim of this research and that what the participants were asked to do was just to read a scenario and comment on their feelings, he could not understand what to do and therefore, did not join. A participant who kindly joined the focus group shared her feeling about reading the scenario. "I actually don't know what to say…. this scenario seems like a puzzle to me at the moment. " (C1).

Since the sample numbers of this qualitative research are too small to analyse if this is a cultural factor or difference in the educational system in a country, or an issue of the individuals, it cannot be described as a general characteristic of Estonian citizens. If it was a cultural issue, it was one of the limitations of this fieldwork, which could possibly affect the results of focus groups of Estonian citizens. It was common in the environment where the author was brought up to imagine the situation by reading sentences to comment on what they felt. For example, in the reading classes in the author's primary school, the pupils were asked to read one scene from science fiction about the future (thus, no one actually experienced any similar situation with that in the sentences), and share their feelings, comments or opinions to discuss with other classmates. Therefore, the author had been taking for granted to have these kinds of things, and could not understand which part was confusing. It could be analysed that some people in some cultures have difficulties to imagine the situation which they have never experienced visibly, and this can be a new interesting topic in a different field which the future works can investigate.

CONCLUSION

This research aimed to investigate messaging from participants' feelings, mindsets and attitudes about fairness perceptions when they imagine algorithms join their performance reviews regarding consideration of demographic factors, such as nationality, gender, and backgrounds. The main research question was how the employees from different nationalities, gender or backgrounds would perceive fairness about performance evaluations when they imagine algorithms would join their performance reviews. The followings were confirmed or found as main results: through the participants' responses and flow of the discussion, it was observed that how difficult it was for the participants to discuss fairness or unfairness in workplaces and the topic of algorithm could help them to share their feelings to some extent. It would be healthy and meaningful to have opportunities and open mindsets to discuss fairness to create a better work environment, since history shows that the employees would never get better work conditions unless they are aware and require. As the previous researches showed, the results of this research also supported the importance of fairness perceptions for the employees' motivations, as well as the different factors for motivations. Though most groups mentioned personal biases of evaluators as common unfairnesses that can be reduced when algorithms join the performance reviews, the Japanese group emphasised that skills to build good personal relationships with their superiors are very important in office politics to be evaluated as part of performance reviews, indicating not trying to avoid the personal biases of the evaluators. This reflects the unique HR systems in Japanese domestic companies where the job descriptions do not exist, therefore the evaluation criteria are inevitably unclear and the evaluations are largely depending on relationships with superiors. Besides, it is extremely tough to change the company because the employees cannot have cross-company skills under the job rotation system where they are transferred to the different department once in 2 or 3 years, therefore, it is crucial for the employees in Japan to build good relationships with the evaluators to stay in the company. There is room for doubt, however, if it is rational for organisations that the skill to build close relationships with managers can be highly assessed to overcome every other performance factor. It was notable that while a lot of participants pointed out quantitative factors, clear job descriptions, setting clear goals at the beginning of the period as important factors for fair performance reviews, they pointed out the importance of evaluating

qualities and contexts of performances which are often not measurable when they were asked: "Which part of the performance review of a human manager + an algorithm is NOT fair compared with a human manager's performance review?". Quality evaluations included factors directly connected to the quality of services for customers and the employees' supportiveness or openness for teamwork in the organisations. Contexts are described in different levels of the scope such as the differences in regional conditions in an international business environment, and the employees' personal differences in experiences. This flow of responses indicated that those participants felt that those qualities and contexts were not necessary factors for fairness, but the important factors for a kind of better performance reviews. The participants valued those qualities and contexts with hopes to have agreeable "Nattoku" feeling (Japanese language to express feelings where a person feels inside his mind that he can understand this is reasonable and rational.) in their evaluations and work environments. In free discussions in the focus groups, several groups reached team evaluations as a method to increase fairness perceptions in performance reviews. It can be said that increasing the number of evaluators and have multiple axes of evaluations criteria would improve fairness perceptions to some extent than the evaluations relying on only the direct boss. As Japanese participants soon pointed them out, however, the negative aspects of team evaluations should be taken into account when the team evaluations are designed: the human relationships with other employees in the office would be even harder, where everyone spends more time buttering up to each other and building internal relationships in order to get good assessments, rather than making efforts to improve customers' satisfactions.

In conclusions, fairness perceptions in performance reviews increase when the number of evaluators or axis of evaluation criteria increases, compared with relying upon only one manager. The HR algorithms would be able to contribute to improving fairness perceptions by joining as one of those additional evaluators and multiple axes of evaluation criteria. Algorithmic evaluations would add values in giving objective information about measurable performance factors, rationalising evaluation processes with certain rules, as result, contribute to reducing human biases. Besides, evaluating qualities of work and employees' soft skills which are hard to measure by numerical values considering contexts are significant for the employees, although those factors were often not fair, not perceived by the employees as necessary factors for fairness. As well as fairness perceptions, it is important for the employees to have agreeable "Nattoku" feelings for a better employee experience, and to create a better work environment for the employees from different cultures and backgrounds. For better or worse, it would be unable to avoid the situation where the workers need to work together with algorithms and technologies such as machine

learning or artificial intelligence in the future, including human resource fields. It was indicated that skills which are needed for human evaluators in the future would be even more challenging soft skills to listen with open-mind, observe, understand, getting the employees' voice heard, and evaluate different qualities in each job, the employees' emotional conditions, motivations, openness, supportiveness, different contexts in intercultural environments, which algorithms would not be good at doing. Whether the employees feel agreeable is depending on if human evaluators can build trust with the employees, which would also remain a huge challenge.

LIST OF REFERENCES

Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016). HR and analytics: Why HR is set to fail the big data challenge. *Human Resource Management Journal*, 26(1), 1e11.

Armstrong, M. (2016). *Handbook of strategic human resource management (6th ed.)*. London, UK: Kogan Page.

Ayub, N., Jehn, K. A. (2018). Exploring diversity effects: nationality composition and nationality context in workgroups. *European Journal of Work and Organizational Psychology*, 27:5

Barrett-Howard, E., & Tyler, T. R. (1986). Procedural justice as a criterion in allocation decisions. *Journal of Personality and Social Psychology*, 50, 296–304.

Brockner, J. (2002). Making sense of procedural fairness: How high procedural fairness can reduce or heighten the influence of outcome favorability. *Academy of Management Review*, 27, 58–76.

Brockner, J. (2006). Why it's so hard to be fair. *Harvard Business Review*, 84, 122–129.

Christin, A. (2017). Algorithms in practice: Comparing web journalism and criminal justice. *Big Data & Society*, 4, 1–14.

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, 86, 425–445

Conlon, D. E., Porter, C. O. L. H., & MacLean Parks, J. (2004). The fairness of decision rules. *Journal of Management*, 30(3), 329–349. https://doi.org/10.1016/j.jm.2003.04.001

Croasmun, P. (2018, May 29). Facing up to bias in facial recognition. American Banker. https://www. americanbanker.com/news/facing-up-to-bias-in-facial- recognition

Diekmann, K.A., Barsness, Z.I., and Sondak, H., (2004). Uncertainty, fairness perceptions and job satisfaction: a field study. *Social Justice Research*, 17,237–255.

Eisenhardt, K.M. (1985). Control: organizational and economic approaches. *Management Science*, 31 (2), 134–149.

Folger, R., Bies, R., (1989). Managerial responsibilities and procedural justice. *Employee Responsibilities and Rights Journal* 2 (2), 79–90.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, UK: Cambridge University Press.

Hartmann, F., & Slapničar, S. (2012). The perceived fairness of performance evaluation: The role of uncertainty. *Management Accounting Research*, 23(1), 17–33. https://doi.org/10.1016/j.mar.2011.10.004

Hamilton, R. H., & Davison, H. K. (2018). The search for skills: Knowledge stars and innovation in the hiring process. *Business Horizons*, 61(3), 409-419.

Hamilton, R. H., & Sodeman, W. A. (2019). The questions we ask: Opportunities and challenges for using big data analytics to strategically manage human capital resources. *Business Horizons*, 63(1), 85–95. https://doi.org/10.1016/j.bushor.2019.10.001

Inoue, H. (2016). Focus on Analytical Study of "Company A" Without an Established Personnel Evaluation System. *Osaka Sangyo University journal of business administration*, 17(1/2), 53-73. （井上 仁志.人事評価制度の信頼性に関する一考察: 人事評価制度が確立していない A 社の分析を中心に. 大阪産業大学経営学部経営学科, 2016 年）.

Ittner, C. D., Larcker, D. F., & Randall, T. (2003). Performance implications of strategic performance measurement in financial services firms. *Accounting, Organizations and Society*, 28(7–8), 715–741. https://doi.org/10.1016/S0361-3682(03)00033-3

Jackson, S., Schuler, R., & Jiang, K. (2014). An aspirational framework for strategic human resource management. *The Academy of Management Annals*, 8(1), 1-56.

Köchling, A., Riazy, S., Wehner, M. C., & Simbeck, K. (2021). Highly Accurate, But Still Discriminatory: A Fairness Evaluation of Algorithmic Video Analysis in the Recruitment Context. *Business and Information Systems Engineering*, 63(1), 39–54. https://doi.org/10.1007/s12599-020-00673-w

Konovsky, M.A. (2000). Understanding procedural justice and its impact on business organizations. *Journal of Management* 26 (3), 489–511.

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data and Society*, 5(1), 1–16. https://doi.org/10.1177/2053951718756684

Leventhal, G. S. (1980). *What should be done with equity theory? New approaches to the study of fairness in social relationships*. In K. Gergen, M. Greenberg, & R. Willis (Vol. Eds.), Social exchange: Advances in theory and research: Vol. 1, (pp. 27–55). New York: Plenum Press.

Libby, T., (1999). The influence of voice and explanation on performance in a participative budgeting setting. *Accounting, Organisations, and Society*, 24, 125–137.

Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR analytics. *International Journal of Human Resource Management*, 28(1), 3-26.

McAbee, S. T., Landis, R. S., & Burke, M. I. (2017). Inductive reasoning: The promise of big data. *Human Resource Management Review*, 27(2), 277e290.

McIver, D., Lengnick-Hall, M. L., & Lengnick-Hall, C. A. (2018). A strategic approach to workforce analytics: Integrating science and agility. *Business Horizons*, 61(3), 397–407.

Moers, F., (2005). Discretion and bias in performance evaluation: the impact of diversity and subjectivity. *Accounting, Organizations and Society*,30, 67–80.

Nagtegaal, R. (2021). The impact of using algorithms for managerial decisions on public employees' procedural justice. *Government Information Quarterly*, 38(1), 1–10. https://doi.org/10.1016/j.giq.2020.101536

Newman, D. T., Fast, N. J., & Harmon, D. J. (2020). When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organizational Behavior and Human Decision Processes*, 160(April 2018), 149–167. https://doi.org/10.1016/j.obhdp.2020.03.008

O'Connor, S. (2016). When your boss is an algorithm. Financial Times. September 7. Retrieved from https://www.ft.com.

Oguma, E. (2019). *The structure of Japanese society -Historical sociology of hiring, educations, and welfare-*. Tokyo: Kodansha. (小熊英二. 日本社会のしくみ　雇用・教育・福祉の歴史社会学 (講談社現代新書). 講談社. Kindle 版. 2019 年.)

Participants in the focus group of Estonian citizens. Fairness perceptions in performance reviews. Author's interview. Voice recording and the author's notes. Online. 14. March 2021.

Participants in the focus group of Estonian citizens. Fairness perceptions in performance reviews. Author's interview. Voice recording and the author's notes. Online. 17. March 2021.

Participants in the focus group of Estonian citizens. Fairness perceptions in performance reviews. Author's interview. Voice recording and the author's notes. Online. 27. March 2021.

Participants in the focus group of Internationals. Fairness perceptions in performance reviews. Author's interview. Voice recording and the author's notes. Online. 22. March 2021.

Participants in the focus group of Internationals. Fairness perceptions in performance reviews. Author's interview. Voice recording and the author's notes. Online. 24. March 2021.

Participants in the focus group of Japanese citizens. Fairness perceptions in performance reviews. Author's interview. Voice recording and the author's notes. Online. 31. March 2021.

Participant in the individual interviews. Fairness perceptions in performance reviews. Author's interview. Voice recording and the author's notes. Online. 27. March 2021.

Participants in the individual interviews. Fairness perceptions in performance reviews. Author's interview. Voice recording and the author's notes. Online. 4. April 2021.

Reuters. (2018, October 10). Amazon ditched AI recruiting tools that favoredmen for technical jobs. The Guardian. https://www.theguardian.com/technology/2018/oct/10/ amazon-hiring-ai-gender-bias-recruiting-engine

Robinson, S., & Mendelson, A. L. (2012). A Qualitative Experiment: Research on Mediated Meaning Construction Using a Hybrid Approach. *Journal of Mixed Methods Research*, 6(4), 332–347. https://doi.org/10.1177/1558689812444789

Robert, L. P., Pierce, C., Morris, L., Kim, S., & Alahmad, R. (2020). Designing fair AI for managing employees in organizations: A review, critique, and design agenda. *Human–Computer Interaction*, Volume 35, 2020 - Issue 5-6, 545-575.

Saunders, M.N.K., Lewis, P., and Thornhill, A. (2015). *Research Methods for Business Students*. 7th edition. England, U.K. Pearson Education, Limited. https://ebookcentral.proquest.com/lib/tuee/reader.action?docID=5175

Scott, A., Klein, F. K., & Onovakpuri, U. (2017). *Tech leavers study: A first-of-its-kind analysis of why people voluntarily left jobs in tech*. Kapor Center for Social Impact.

Sheppard, B. H., & Lewicki, R. J. (1987). Toward general principles of managerial fairness. *Social Justice Research*, 1, 161–176.

Starke, C., Baleis, J., Keller, B., and Marcinkowski, F. (2021). Fairness Perceptions of Algorithmic Decision-Making: A Systematic Review of the Empirical Literature. *Human-Computer Interaction*. arXiv:2103.12016.

Studyinestonia. (2020). *International degree students in Estonia: facts and figures 2020*. Retrieved from https://studyinestonia.ee/sites/default/files/International%20Degree%20Students%20in%20Estonia%202020.pdf, 13. April. 2021.

Taylor, M.S., Tracy, K.B., Renard, M.K., Harrison, J.K., Carroll, S.J. (1995). Due process in performance appraisal: a quasi-experiment in procedural justice. *Administrative Science Quarterly* 40, 495–523.

Thibaut, J., & Walker, L. (1975). *Procedural justice: A psychological analysis*. Hillsdale, NJ: Erlbaum Press.

Tyler, T. R., & Blader, S. L. (2003). The group engagement model: Procedural justice, social identity, and cooperative behavior. *Personality and Social Psychology Review*, 7, 349–361.

United Nations Statistics Divisions (2021, April 3). Standard Country or Area Codes for Statistical Use. [Online] Retrieved from https://unstats.un.org/unsd/methodology/m49/

Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, 4(2), 1–17. https://doi.org/10.1177/2053951717743530

Yoshida, T. (2016). Interpersonal Relationships with Supervisors and Consent to Personnel Appraisal. *Japanese Association of Industrial/Organizational Psychology Journal*, 2016, Vol. 30, No. 1, 29-43. （吉田朋子（筑波大学大学院. 上司との対人関係が人事評価の納得感に及ぼす影響. 産業・組織心理学研究 2016 年, 第 30 巻, 第 1 号, 29−43）.

Žliobaitė, I. (2017). Measuring discrimination in algorithmic decision making. *Data Mining and Knowledge Discovery*, 31(4), 1060–1089. https://doi.org/10.1007/s10618-017-0506-1

APPENDICES

Appendix 1. Transcripts of the interviews

https://www.dropbox.com/sh/18eqkguxwz1plx4/AAAeb1a4MrrbI8hUGCseZCZma?dl=0

Appendix 2. Compositions of participants in the focus groups

The compositions of nationalities, genders, academic backgrounds and work experience of the participants are organised as the followings:

Estonian citizens						
Group A						
	Nationality	Gender	Academic Backgrounds	Work Experience	Work Years	Management Experience
A1	Estonian	Female	Computer Science	Full-stack developer	2	-
A2	Estonian	Male	Business Administration	Production engineering, construction	10	Yes
A3	Estonian	Female	Marketing and Management	HR, CRM	10	Yes
A4	Estonian	Female	Integrated Science	A waitress, an actress, etc.	2	-
A5	Russian-speaking Estonian	Female	Business Administration; marketing specialization	Barista, restaurant waiter, Business Development Associate	0.7	-
Group B						
B1	Estonian	Female	Linguistics	Language editor, sales project manager, recruiter, HR manager	17	Yes
B2	Russian-speaking Estonian	Female	Business Administration	Customer service, communication specialist, group leader, advisor in a bank, KYC/AML	5	-
B3	Estonian	Male	Robotics	Customer service, robot operator, manufacturing, sales	4	-
B4	Estonian	Female	Heating, ventilation & air conditioning engineering	Waiter, admission officer, secretary, HVAC technician, resource efficiency consultant, student consultant	3	-
Group C						
C1	Estonian	Female	Business Administration	Project manager	18	Yes
C2	Estonian	Female	Politics and Governance	Customer service, violin teacher, babysitter (au pair), sales manager, digitiser	4	-

C3	Estonian	Male	Civil Engineering Business Administration	Materials laboratory Engineer, Civil Engineer, Export and Logistics Manager, Member of Management Board	26	Yes
C4	Estonian	Female	English Language and Culture	Dean's assistant, sales & marketing assistant, translator	5	-
C5	Estonian	Female	Sociology	Customer service, team and process management, analyst, head of the board	23	Yes
C6	Russian-speaking Estonian	Male	Power engineering	Electrician, railway traffic engineer	4	-

Internationals						
Group D						
	Nationality	Gender	Academic Backgrounds	Work Experience	Work Years	Management Experience
D1	Southern Asia	Male	Business Administration	Project Coordinator, Network Administrator, Network Engineer, Senior Customer Sales	10	-
D2	Southern Asia	Female	Business Administration	Customer Service and Marketing	8	-
D3	Western Africa	Male	Business Administration	Financial Accounting	15	Yes
D4	Northern America	Male	English Literature	Insurance, entrepreneur, sales & marketing, operations, importer, product development, consultant	55	Yes
Group E						
E1	Eastern Asia	Female	Business Administration	Marketing	3	-
E2	Western Africa	Male	Business Administration	Banking operations	1	-
E3	Eastern Europe	Female	Psychology of Management	Customer Care Specialist, Project Manager	7	-

E4	Northern Africa	Female	International marketing	Junior marketing officer	0.6	-
E5	Eastern Europe	Male	Business Administration	Customer Support, Project Management, Marketing	4	-
Individual Interviews						
F1	Eastern Europe	Female	Business Administration	Finance, Marketing	1	-
F2	Eastern Europe	Male	Business Administration	Supply Chain, Fundraising, Marketing, Business development	2	-
F3	Western Asia	Female	International law and Human rights	Corporate Lawyer	0.9	-

Japanese citizens						
Group G						
	Nationality	Gender	Academic Backgrounds	Work Experience	Work Years	Management Experience
G1	Japanese	Male	Arts and Sciences	System Engineer IT consultant	19	Yes
G2	Japanese	Male	Engineering	System Engineer IT administrator	18	Yes
G3	Japanese	Male	Accounting	System Engineer IT consultant	16	Yes
G4	Japanese	Male	Business Engineering	System Engineer	20	Yes

Source: Collected through the participants' responses and author's note of focus groups. The regional categories followed the standard by United Nations Statistics Divisions (2021)

Appendix 3. Non-exclusive licence

A non-exclusive licence for reproduction and for granting public access to the graduation thesis[1]

I, Tomoko Yamagata (*author's name*)

1. Give Tallinn University of Technology a permission (non-exclusive licence) to use free of charge my creation

(The title of graduation thesis in English) Comparative Analysis: The Perceptions of Fairness in Performance Reviews with Algorithms

(The title of graduation thesis in Estonian) Õiglustaju algorütmipõhise tulemushindamise kasutamisel: võrdlev analüüs,

(*title of the graduation thesis*)

supervised by Merle Ojasoo, PhD, (*name of the supervisor*)

1.1. to reproduce with the purpose of keeping and publishing electronically, including for the purpose of supplementing the digital collection of TalTech library until the copyright expires;

1.2. to make available to the public through the web environment of Tallinn University of Technology, including through the digital collection of TalTech library until the copyright expires.

2. I am aware that the author will also retain the rights provided in Section 1.

3. I confirm that by granting the non-exclusive licence no infringement is committed to the third persons' intellectual property rights or to the rights arising from the personal data protection act and other legislation.

---

[1] *The non-exclusive licence is not valid during the access restriction period with the exception of the right of the university to reproduce the graduation thesis only for the purposes of preservation.*