The Impact of CV Recommender Systems on Procedural Justice in Recruiting: An Experiment in Candidate Selection

Verena Eitle  
*Technische Universität Darmstadt*, eitle@is.tu-darmstadt.de

Felix Peters  
*Technische Universität Darmstadt*, peters@is.tu-darmstadt.de

Andreas Welsch  
*Technische Universität Darmstadt*, welsch@is.tu-darmstadt.de

Peter Buxmann  
*Technische Universität Darmstadt*, buxmann@is.tu-darmstadt.de

Follow this and additional works at: [https://aisel.aisnet.org/ecis2021_rp](https://aisel.aisnet.org/ecis2021_rp)

Recommended Citation

[https://aisel.aisnet.org/ecis2021_rp/65](https://aisel.aisnet.org/ecis2021_rp/65)

This material is brought to you by the ECIS 2021 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2021 Research Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
THE IMPACT OF CV RECOMMENDER SYSTEMS ON PROCEDURAL JUSTICE IN RECRUITING: AN EXPERIMENT IN CANDIDATE SELECTION

Research Paper

Verena Eitle,
Technical University of Darmstadt, Germany, eitle@is.tu-darmstadt.de
Felix Peters,
Technical University of Darmstadt, Germany, felix.peters@tu-darmstadt.de
Andreas Welsch,
Technical University of Darmstadt, Germany, andreas.welsch@tu-darmstadt.de
Prof. Dr. Peter Buxmann,
Technical University of Darmstadt, Germany, peter.buxmann@tu-darmstadt.de

Abstract

Due to the increasing amount of digitally available applicant information recruiters have difficulties to manage applications through manual recruiting practices. Using CV recommender systems in the selection phase supports recruiters in identifying the most suitable candidates by computing the similarity between a candidate’s profile and job requirements. While recent research has mainly focused on technical improvements, we seek to gain more insights about human-algorithm interactions in recruiting. Our study aims to examine what impact the use of a CV recommender system has on procedural justice in the selection process. Through an experimental set-up with 74 recruiters from 22 multinational companies, our study shows that the incorporation of a CV recommender system helps recruiters to ensure the rule of consistency and bias suppression in the selection phase. Thus, our quantitative results indicate that CV recommender systems can have an impact on procedural justice in candidate selection.

Keywords: Candidate Selection, CV Recommender Systems, Procedural Justice.

1 Introduction

Advancements in information systems and social developments have significantly influenced the way of working in the field of human resource management (HRM). In recent years, organizations have shifted their priorities towards HRM as they perceive their workforce as one of their most important assets. The increasing demand for qualified talents might also result in a war for talents as the shortage of talents is considered one of the most worrying concerns among CIOs and IT executives in 2019 (Kappelman et al., 2020). To attract, select, and retain these talents, recruiting has become a strategic priority in organizations. Black and van Esch (2020) argue that digitization has made a major contribution to further developments in recruiting and emphasize the following eras of e-recruitment. Digital Recruiting 1.0 and 2.0 enable organizations to post job openings on digital job boards on the internet and social network platforms such as LinkedIn. Organizations have the opportunity to narrow down their target group of potential candidates and to contact them directly with concrete job postings (Black and van Esch, 2020). By searching through many digital job postings with a few simple clicks, potential candidates are able to submit multiple applications with less effort. As a result, the increase of incoming applications has made the manual recruiting process more difficult for organizations as
recruiters have to manually process digitally available applicant information. While coping with this large amount of applications, recruiters also need to ensure fairness in the selection process as their decision has a major impact on the applicants future (Arvey and Renz, 1992; Gilliland, 1993). However, procedural justice along the decision-making process in the selection phase is often impeded by recruiters’ previous work experiences, own beliefs or personal biases (Åslund and Skans, 2012; Eckhardt et al., 2014).

To cope with the increasing amount of data and to ensure fairness, different types of artificial intelligence (AI) technologies have been integrated into the recruiting process (Strohmeier and Piazza, 2015; van Esch et al., 2019) which Black and van Esch (2020) describe as Digital Recruiting 3.0. In particular, the development of Curriculum Vitae (CV) recommender systems is an essential research area in the selection phase of recruiting. These systems are typically applied in the selection phase of the recruiting process (Schneider, 1981) to estimate the person-job (P-J) fit (Caldwell and O Reilly, 1990; Edwards, 1991; Wilk and Sackett, 1996; Kristof-Brown, 2000). By computing the similarity between the details of a candidate’s profile and the given job requirements, CV recommender systems can support recruiters in identifying the most suitable candidates. While the performance level is constantly increasing due to technical improvements (e.g., Malinowski et al., 2006; Lu et al., 2013; Bansal et al., 2017), little is known about the socio-technical context of the interaction between human recruiters and CV recommender systems (Green and Chen, 2019a). Since the final decision in candidate selection still remains in the power of recruiters, further insights about the human-algorithm interactions are essential in order to investigate the effect on procedural justice. Therefore, our study aims to examine what impact the use of a CV recommender system has on procedural justice in the selection process. As a research design, we have chosen an experimental set-up in which 74 recruiters from 22 large multinational companies were given the instruction to create top-10 rankings of candidates for two fictional job postings. By randomly assigning the participants to either the control group which represents the non-CV recommender system supported settings or to the treatment group in which recruiters received a matching score generated by a CV recommender system, we were able to investigate our research question. Our study contributes to research and practice in the field of recruiting by providing quantitative findings that CV recommender systems tend to ensure procedural justice as recruiters are able to rank candidates in a more consistent manner and are more likely to assess a candidate’s knowledge, skills, and abilities when relying on the CV recommender system.

The rest of this paper is structured as follows: Section 2 outlines the theoretical background of recommender system in recruiting with a focus on CV recommender systems and elaborates on procedural justice in candidate selection. After describing the research design in the form of an experimental set-up in section 3, we present the results of the quantitative study in section 4. The discussion, the contributions to research and practice as well as the limitations and opportunities for future research are outlined in section 5, followed by the conclusion in section 6.

2 Theoretical Background

2.1 Overview of Recommender Systems

Over the last couple of years, the overload of information with which people need to cope on a daily basis has resulted in complex decision-making environments. The fact that humans have difficulties making decisions due to their limited cognitive resources and time constraints in evaluating and processing available information was coined by Simon (1955) as the phenomenon of bounded rationality. In order to help people deal with the overwhelming amount of data and to support them in the intelligence, design, choice, and implementation phase of complex decision-making processes (Simon, 1977), recommender systems have been developed. By generating personalized suggestions, recommender systems offer only a small number of selection options and eliminate irrelevant and excessive information (Burke, 2002; Adomavicius and Tuzhilin, 2005). To be more precise, the primary use of recommender systems is to predict elements that a user is likely to evaluate as positively according to his or her underlying preferences (Ricci et al., 2011). In general, recommender systems can be
classified into the following categories: Content-based, collaborative filtering, and knowledge-based recommender systems (Burke, 2002; Ricci et al., 2011; Aggarwal, 2016). Content-based recommender systems recommend items to users that are similar to those that they have historically favored or expressed interest in. In order to retrieve a user’s preferences, tastes, and desires, the recommender system uses long-term user profiles with user attributes that have been accumulated over time. By matching these user attributes to item attributes, new items will be recommended to the user (Adomavicius and Tuzhilin, 2005; Pazzani and Billus, 2007; Aggarwal, 2016). Since content-knowledge is mainly derived from unstructured or semi-structured data, item descriptions are composed of a set of textual features that can be acquired by various information retrieval or information extraction methods with the help of statistical, machine learning, or natural language processing techniques (Lops et al., 2011). In contrast, collaborative filtering recommender systems generate item recommendations based on the similarity towards other users’ preferences (Adomavicius and Tuzhilin, 2005; Schäfer et al., 2007; Aggarwal, 2016). This type of recommender system has to cope with a so-called cold-start issue as a new user has to first rate several items or a new item has to receive a couple of ratings before a user similarity can be determined (Ramezani et al., 2008; Bobadilla et al., 2013). Recommendations generated through knowledge-based recommender systems are derived from specific domain knowledge which have to be acquired through interviews or other knowledge discovery techniques (Aggarwal, 2016). A common form of knowledge representation are ontologies which display relations among attributes, objects, and item features. The main downside of this recommender system lies in the high efforts of knowledge acquisition (Ramezani et al., 2008).

2.2 Recommender Systems in Recruiting

According to the attraction-selection-attrition (ASA) framework by Schneider (1981), organizations tend to achieve a certain degree of homogeneity among their employees by identifying candidates during the recruiting phases of attraction, selection, and attrition who have similar characteristics and behaviors as the organization. The empirical study by Judge and Cable (1997) revealed that in the attraction phase, potential candidates search for suitable job postings and organizational cultures based on their own personality, preferences, and field of interest. Particularly in the attraction phase, there is a tendency of organizations to achieve a certain degree of homogeneity by seeking to recruit candidates with similar attributes and behaviors which is also described by the term "right types" (Schneider, 1981). In the selection phase, organizations seek to select candidates who possess specific competencies and skills required for the job position. By narrowing the applicant pool using pre-selection techniques and face-to-face interviews, companies are able to select a homogeneous group of candidates with specific skills (Schneider, 1981; Bretz et al., 1989). During the attrition phase, there is a tendency for employees who do not fit into the organization to eventually leave, while employees who embrace the organizational culture strive to retain their jobs and pursue their careers over time (Schneider, 1981; Chatman, 1991). By retaining the "right types" in the organization who share similar characteristics and behaviors, companies can increase the homogeneity among their workforce (Schneider, 1981). Since the increase in digital job and applicant data particularly impedes the screening and assessment activities of recruiters (Black and van Esch, 2020), the following sections mainly refer to the selection phase in recruiting.

The main task in the selection phase is the matching of potential candidates and job postings, which is an essential subject of the person-job (P-J) fit and person-organization (P-O) fit literature (Rynes and Gerhart, 1990; Adkins et al., 1994; Wilk and Sackett, 1996; Judge and Cable, 1997). The overarching research relates to the fit between a person and the environment, which has been a pervasive component in major research areas including personality theory, occupational psychology, personnel selection, and social psychology (Schneider, 2001). According to the person-environment fit concept (P-E), behavior is influenced by the congruence between personal and situational variables and not just by one of the elements alone. To be more precise, the compatibility between personal variables including abilities, needs, and values as well as environmental variables such as organizational culture, task demands, and job attributes leads to either positive or negative outcomes (Muchinsky and Monahan, 1987; Ostroff, 1993; Kristof, 1996; Schneider, 2001). Besides the P-J and P-O fit, the comprehensive P-E fit concept
comprises further sub-categories including the person-vocation (P-V) fit, the person-group (P-G) fit, and the person-supervisor (P-S) fit (Sekiguchi, 2004; Kristof-Brown et al., 2005).

In the recruiting literature, the concepts of P-J and P-O fit predominate the selection phase of Schneider's (1981) ASA framework since the primary objective is to match individuals and jobs. The operationalization of the P-J fit by Edwards (1991) refers to the demands-ability fit and the needs-supplies fit. To be more precise, the demands-ability fit determines the extent to which an employee's knowledge, skills, and abilities, the so-called KSA's, meet the requirements of a job. These KSA's comprise, for example, work experience, technical skills, problem-solving skills, academic experience, and leadership skills (Kristof-Brown, 2000). The needs-supplies fit, on the other hand, addresses whether needs, wishes, or preferences of an employee are satisfied by the jobs' characteristics and attributes (Edwards, 1991; Kristof, 1996; Sekiguchi, 2004). However, since candidates tend to select the vacant job positions according to their own needs and preferences (Judge and Cable, 1997), the primary task of recruiters is to identify candidates with the required KSA's. The study by Caldwell and O'Reilly (1990) showed that the match between the KSA's of a candidate and the job requirements positively influences an employee's job performance and ultimately job satisfaction. Furthermore, Wilk and Sackett (1996) reported that the match between an employee's skills and the complexity of the job even allows the employee to move up in the job hierarchy in the future. These empirical results indicate that the demands-ability fit is crucial for assessing the P-J fit (Kristof-Brown, 2000). With regard to the operationalization of the P-O fit, Chatman (1991) argues that the congruence between candidates' values as well as organizational norms and values can have a positive impact on the selection phase since this match increases the likelihood that a candidate identifies himself with the organizational culture. An experiment conducted by Kristof-Brown (2000) revealed that recruiters explicitly distinguish between the P-J fit and the P-O fit when selecting applicants. When assessing the first group of applicants, recruiters tend to follow the P-J fit as they primarily consider the KSA's as their main selection criteria. In the subsequent evaluation rounds of the recruiting process, the emphasis is on the P-O fit since the match between personal values and organizational values is given higher priority (Rynes and Gerhart, 1990; Kristof-Brown, 2000).

With the advancements of Digital Recruiting 1.0 and 2.0 (Black and van Esch, 2020), which allow organizations to post their job openings on digital job boards and professional and social networking platforms like LinkedIn, the amount of digital candidate data has increased significantly. Since the selection phase involves a high proportion of manual tasks, managing the large volume of digital applications can be time-consuming and costly for organizations (Eckhardt et al., 2014; Strohmeier and Piazza, 2015). While different types of artificial intelligence (AI) technologies can be integrated throughout the recruiting process (Strohmeier and Piazza, 2015; van Esch et al., 2019), the emergence of recommender systems have particularly simplified the manual tasks of recruiters in the selection phase. The study by Faerber et al. (2003) has compared the prediction performance of a content-based recommender system, a collaborative filtering recommender system, and a hybrid approach in the field of CV recommendations. According to their findings the content-based approach yields the best results in matching a candidate's profile and the job requirements. Based on the P-J fit, Malinowski et al. (2006) have developed a CV recommender system that follows the demands-ability fit approach (Edwards, 1991) by recommending candidates whose CVs most closely match the specific job requirements. In order to address the needs-supplies fit approach (Edwards, 1991) by matching a candidate's preference with the job attributes, the authors additionally developed a job recommender system. Based on a latent aspect model both recommender systems are able to compute the similarity between the candidate’s profile and the job requirements. Since the predictive quality of the two recommender systems was the main subject of the study, the computer-generated recommendations were compared with the original list of jobs selected by the study participants and the original list of top candidates. The results showed that the predictions of the CV and the job recommender systems largely corresponded to human choices, which indicate a high prediction quality and promising system performance. Moreover, the content-based recommender system proposed by Lu et al. (2013) is designed as a hybrid model that integrates a CV and a job recommender system in one system. The profile-based similarity of the candidate’s details and the job posting was computed by using latent semantic analysis (LSA) tools. In addition, the...
Procedural Justice in Candidate Selection

A recommender system is capable of not only including a candidate's profile and the job requirements, but also processing user interactions. In an experiment, the participants were able to indicate their preferences through the interaction features “posted”, “applied”, “favorited”, “liked”, and “visited”. The study by Almalis et al. (2016) extents the research of content-based CV recommender systems in a way that the match between human KSA’s and job attributes is based on different value ranges, such as specific values, a range with lower limit, a range with upper limit, and a range with both lower and upper limit. In other words, the proposed CV recommender system is able to differentiate between job requirements that refer to ranges of values such as “candidates must be at least 40 years old” or “between 18-40 years old”. The proposal of a further hybrid content-based recommender system by Bansal (2017) facilitates the matching of candidates profiles and job postings from the perspective of recruiters and job seekers in an integrated system. Instead of using words as textual features, the researcher focused on topic features by applying the topic modelling algorithm Latent Dirichlet Allocation (LDA). Since this unsupervised machine learning technique allows to detect latent topics that are hidden in the text corpus, low-frequency terms can become quite significant as they are linked to other high-frequency terms.

As shown, current research in the selection phase focuses mainly on enhancing the prediction performance of CV recommender systems by evolving algorithms and improving technical features (Faerber et al., 2003; Malinowski et al., 2006; Lu et al., 2013; Almalis et al., 2016; Bansal et al., 2017). However, instead of optimizing computational performance, Green and Chen (2019a) emphasize that attention in research should rather shift towards a socio-technical context to explore how human-algorithm interactions can be improved. According to their algorithm-in-the-loop framework, algorithmic aid can help to improve the decision-making process by incorporating algorithms which inform and advise humans in their decision-making while the final decision still remains with humans. Although the study of human-algorithm interaction is developing slowly in areas such as web journalism (Christin, 2017), forecasting (Dietvorst et al., 2018), and criminal justice (Green and Chen, 2019b; Grgic-Hlaca et al., 2019), research has not yet sufficiently taken into account the socio-technical context in the field of recruiting (Green and Chen, 2019a; Grgic-Hlaca et al., 2019).

2.3 Procedural Justice in Candidate Selection

Despite the fact that key performance indicators in the recruiting process are largely standardized, the selection process for candidates often differs among recruiters. Previous work experience, individual attitudes, and personal preferences lead to a variety of different behaviour patterns among recruiters which can significantly influence the selection of suitable candidates (Eckhardt et al., 2014). Furthermore, the existence of conscious or unconscious cognitive bias among recruiters might also contribute to the likelihood of inconsistent decision-making processes in the selection phase (Åslund and Skans, 2012; Black and van Esch, 2020). These diverse set of behaviour patterns among recruiters increase the risk of unfairness in the selection phase and can ultimately compromise a candidate’s chance of being selected.

In order to examine fairness in the decision-making process during the selection phase, the literature on organizational justice (Greenberg and Colquitt, 2005) must be taken into account which primarily addresses employees' reactions regarding unfairness and inequity in an organizational context and distinguishes between distributive and procedural justice. Distributive justice describes the degree to which an employee perceives the distribution of outcomes such as payments and rewards as fair in the sense of equity (Adams, 1965; Cohen, 1987) and equality (Deutsch, 1975). When considering the equity principles which determine the distribution of resources according to the contributions of employees, the foundation of distributive justice refers to Adam’s (1965) equity theory. In contrast, procedural justice refers to the perceived fairness in the actual decision-making process that ultimately determines the outcome (Greenberg and Colquitt, 2005). In order to ensure that the procedure can be assessed as fair, Leventhal (1980) defined the following six rules for procedural justice: consistency, unbiased suppression, accuracy, correctability, representativeness, and ethicality. Since fairness in the candidate selection process depends on procedural justice, Gilliland (1993) and Arvey and Renz (1992) have
defined specific procedural rules for the selection phase. In this context, the rule of consistency should be emphasized as Leventhal (1980) and Gilliland (1993) recommend a certain degree of uniformity in the selection procedure since all candidates should have the chance to receive the same decision-making process regardless of demographics, personality, or background. Arvey and Renz (1992) point out that consistency in candidate selection is only given when the content of the selection system, the scoring, and the interpretation of scores are standardized across all applicants. In addition, the rule of bias suppression by Leventhal (1980) is also crucial to ensure procedural justice in the selection phase as it determines that recruiters should not make decisions based on their own self-interest or be influenced by their own beliefs and opinions (Leventhal, 1980). To ensure objectivity rather than risking subjectivity, Arvey and Renz (1992) suggests to apply quantifiable methods which take certain criteria into account rather than relying on the recruiters’ instincts and experiences. The suppression of personal bias is also addressed in the propriety of questions as improper questioning and prejudicial statements impede the level of fairness in the selection phase (Gilliland, 1993).

By examining procedural justice in e-recruiting tools, the findings of Thielsch et al. (2012) show that applicants expect a higher level of objectivity when using an e-recruiting tool compared to traditional manual recruiting practices. In addition, the qualitative study by Ochmann and Laumer (2019) proposes that the implementation of AI-based instruments could contribute even more to increase the level of fairness by increasing objectivity during the selection phase. While traditional selection methods have been perceived as unfair due to the risk of personal bias on part of the recruiters, the qualitative findings suggest that AI technologies could assist in the decision-making process by focusing solely on the candidates’ skills and thus increasing objectivity. It should be noted, however, that the level of user reliance in a technology is also considered a critical factor in achieving procedural justice, as the final selection decision still remains in the power of recruiters. Reliance towards a technology depends primarily on user acceptance and the degree of influence that the user allows in their judgment (Arnold and Sutton, 1998; Madhavan and Wiegmann, 2007). Following the study by Ötting and Maier (2018) which empirically examined the impact of human and AI-based intelligent systems on procedural justice in a generic work-life situation, we aim to gain empirical insights into procedural justice in the selection process.

Based on the outlined literature on candidate selection (e.g. Schneider, 1987; Caldwell and O Reilly, 1990; Edwards, 1991; Wilk and Sackett, 1996), CV recommender systems (e.g. Malinowski et al., 2006; Almalis et al., 2016; Bansal et al., 2017), and procedural justice (Arvey and Renz, 1992; Gilliland, 1993; Greenberg and Colquitt, 2005), we believe that incorporating a CV recommender system in the selection phase could increase procedural justice by helping recruiters to ensure consistency and objectivity in their decision-making process. Under the premise that recruiters rely on a CV recommender system and take the generated suggestions into account, we anticipate that the top-10 rankings of recruiters who incorporate a CV recommender system into their decision-making process will be more consistent and similar than the top-10 rankings of those who rely solely on their own judgement without using a CV recommender system. Furthermore, we would like to gain further insights into the demands-ability approach in the context of the P-J fit (Edwards, 1991; Kristof-Brown, 2000) when incorporating a CV recommender system in the decision-making process of recruiters. As outlined above, CV recommender systems are based on the demands-ability fit as they compute the similarity between the applicant’s KSA’s and the respective job requirements (Faerber et al., 2003; Malinowski et al., 2006; Lu et al., 2013; Almalis et al., 2016; Bansal et al., 2017). Since the suggestions generated by CV recommender systems are based on the KSA’s of candidates and are not exposed to subjective discrimination or personal bias by recruiters, we anticipate that the CVs of the top-10 ranked candidates which were selected with the help of a CV recommender system possess stronger KSA’s than those ranked on the basis of the recruiters’ sole judgment.
3 Methodology

Since the aim of our study is to examine what impact the use of a CV recommender system has on procedural justice in the selection process, we conduct an true experimental research with a posttest-only control group that enables us to determine cause-effect relationships (Campbell and Stanley, 1963; Gay et al., 2012). The research design of the experiment is illustrated in Figure 1 and is described in detail in the following section.

Figure 1. Experimental research design

In order to establish a realistic experimental set-up for a decision-making process in the selection phase, we have involved professional recruiters rather than non-professional study participants. Through a cooperation with a national association for employer branding, talent marketing, and recruiting, we were able to randomly select recruiters who were willing to participate in our experiment. The random selection method is recommended primarily because it ensures external validity by increasing the degree to which the study results can be generalized to other groups (Campbell and Stanley, 1963; Kirk, 2013; Dean et al., 2017). By following a between-subject design, we have also applied the randomization method when assigning participants to either the control or the treatment group. Random assignment is particularly needed to ensure internal validity as it reduces systematic bias between the treatment and the control group by distributing participants equally among these groups (Campbell and Stanley, 1963; Kirk, 2013; Mikolov et al., 2013). The manipulation of the independent variable refers to a matching score generated by a CV recommender system and distinguishes the groups as follows: The control group represents the non-CV recommender system supported setting in which the participants have received CVs without any suggestions generated by a CV recommender system. The treatment group represents the CV recommender system supported setting in which the participants have received the same CVs but to which a matching score generated by the applied CV recommender system has been added in the upper right corner.

Following the current research on recommender systems in the field of candidate selection (Faerber et al., 2003; Malinowski et al., 2006; Lu et al., 2013; Almalis et al., 2016; Bansal et al., 2017), we have decided to also use a content-based CV recommender system that supports recruiters in the selection phase to identify suitable candidates. As we seek to examine the effect of CV recommender systems on procedural justice rather than improving the performance level through technical advancements, we decided to use an existing CV recommender system developed by a global enterprise software provider with sufficient training data. The underlying machine learning technique refers to the word2vec algorithm by Mikolov et al. (2013) which represents a neural network model with a single hidden layer. In the case of the applied CV recommender system, data cleansing activities such as functional removal,
lower case, and plural removal are performed on the input document in an initial step. After this prerequisite is fulfilled, the input document is tokenized into corresponding bigrams and trigrams. By using the word2vec algorithm, each token is assigned to a word embedding which ultimately represents a vector space. In order to remove irrelevant tokens and to generate interpretable token clusters, the tokens in the form of word embeddings are assigned to certain branches of a formerly created skill tree. As the final goal is to compare a CV and a job posting, the word embeddings are combined into document-level embeddings to compute the cosine similarity between the document vectors. The output of the selected content-based CV recommender system is a matching score which is expressed as a floating-point number between 0.0 and 1.0. The higher the value of the matching score, the closer the similarity between the CV and the job profile, and the higher the rank of the CV in a list of suitable candidates.

In regard to the experimental set-up, we created two fictional job posting based on examples from the participating companies: one for a Junior Full Stack Developer and one for a Junior Online Marketing Manager. We focused on junior positions as these positions oftentimes receive large numbers of applications and are thus more attractive for the implementation of a CV recommender system. As a second step, we collected a diverse set of 30 CVs for each job posting from computer science, information systems, business, and marketing students of higher education institutions. During the experiment, the participants received the task description through a survey tool in which the two job postings as well as the corresponding CVs were available in the form of PDF documents. Based on the random assignment to either the control group or the treatment group, the CVs either included a matching score generated by the CV recommender system or not. According to the task description, the participants of both experiment groups were instructed to first read the job postings and the corresponding CVs thoroughly. Based on a careful assessment of the candidates and the requirements of the first job posting, the recruiters were asked to create a ranking in the survey tool based on the suitability of the candidates under the assumption that the top-10 ranked candidates would be invited for a further interview. Subsequently, the participants were encouraged to proceed with the creation of the ranking list for the second job posting. Since the ranking represents the final outcome of the experiment and is considered in the further data analysis, our study is designed as posttest-only control group. This approach allows us to avoid testing effects that could have had an impact on the participants’ behaviour if they were exposed to any kind of information in advance (Gay et al., 2012).

4 Results

Regarding the participation rate, 89 professional recruiters voluntarily signed up for our experiment, out of which 74 completed the tasks (83% response rate). At the time of the experiment (January 2019), these 74 participants were employed in 22 large multinational companies. Among all participants, 74% were female, 83% were between 25 and 44 years old, and 78% had at least three years of experience in recruiting. To ensure objectivity within the quantitative data analysis, we manually extracted variables from all CVs in a two-stage procedure. First, variables were extracted independently by two of the authors. Subsequently, results were synchronized to reach consent and to apply consistent standards. In accordance with the P-J fit which determines the suitability between the KSA’s of candidates and the job requirements (Edwards, 1991; Kristof-Brown, 2000), we extracted the following variables from all CVs: study duration (in years) and relevant working experience (in years). By conducting statistical tests, we were able to relate these variables to the observed behaviour of the participating recruiters. The data was pre-processed using the Python programming language and subsequently analysed using SPSS. Given the nature of our study, we tested for significance at a 10% level to avoid discarding interesting relationships (Rosnow and Rosenthal, 1989; Schumm et al., 2013). In regard to the following section, we present the results as mean ± standard deviation, unless we state otherwise. While screening the experiment data, we detected three cases in which participants only completed the marketing job posting, and one case in which only the development task was finished. We decided to keep these partial completions in our dataset to account for the rather small sample size.
To examine whether the top-10 rankings of recruiters who were supported with the matching score generated by the CV recommender system are more consistent and similar than those who rely on their own judgement, we first calculated pairwise correlations between the rankings of participants separately for the CV recommender system supported group and the non-CV recommender system supported group (in the following referred to as inner group ranking correlation). Here, the ranked candidates received their respective position, while candidates outside the top-10 were being ranked as 11th, thus creating a lot of ties in our rankings. Consequently, we chose Kendall’s tau as correlation metric for this analysis, as this metric is more robust in the presence of ties in rankings (Kendall and Stuart, 1945). We then conducted two separate independent-samples t-tests (i.e., one for each job posting) to examine the effects of CV recommender system support on the inner group ranking correlation. The results of our quantitative analysis are summarized in Table 1.

Table 1. Effect of CV recommender system support on inner group ranking correlation

<table>
<thead>
<tr>
<th>Factor</th>
<th>Task</th>
<th>Levels</th>
<th>Inner group rank. corr.</th>
<th>df</th>
<th>t</th>
<th>Sig.</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV recommender system support</td>
<td>Development</td>
<td>Supported</td>
<td>.440 ± .233</td>
<td>1235</td>
<td>-21.989</td>
<td>.000</td>
<td>1.281</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unsupported</td>
<td>.144 ± .231</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marketing</td>
<td>Supported</td>
<td>.489 ± .274</td>
<td>1300</td>
<td>-18.958</td>
<td>.000</td>
<td>1.057</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unsupported</td>
<td>.192 ± .288</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Inner group ranking correlations are calculated as pairwise correlations between rankings from participants of the respective group, as measured by Kendall’s tau. Results are based on independent-samples t-tests.

Based on our quantitative analysis we found statistically significant differences between the CV recommender system supported and non-CV recommender system supported groups with regard to the inner group ranking correlation score. The results showed that the inner group ranking correlation was higher in the CV recommender system supported (development task: .440 ± .233; marketing task: .489 ± .274) than in the non-CV recommender system supported groups (development task: .144 ± .231; marketing task: .192 ± .288). In other words that means that the rankings from recruiters who received the matching score generated by the CV recommender system were more strongly correlated with each other than rankings from recruiters without the CV recommender system support. For both groups, the effects were statistically significant (development task: t = -21.989, p < .001; marketing task: t = -18.958, p < .001) and effect sizes were larger than one standard deviation, as measured by Cohen’s d (development task: 1.281; marketing task: 1.057).

According to our results, we can strongly support our anticipation that the top-10 rankings of recruiters within the CV recommender system supported group are more consistent and similar than those who did not received any matching score from the CV recommender system. We can further suspect that recruiters relied on the matching score generated by the CV recommender system. To further examine this finding, we also calculated the average correlation between the recruiters’ rankings and the ranking proposed by the CV recommender system. We found a strong correlation for both tasks (development task: .583 ± .253; marketing task: .599 ± .268), which further supports our assumption. For comparison, in the unsupported groups the observed correlations were much lower (development task: .062 ± .231; marketing task: .080 ± .310).

To examine our second anticipation that the CVs of the top-10 ranked candidates which were selected using a CV recommender system possess stronger KSA’s than those which were ranked without any CV recommender system support, we compared the ranked candidates of the control and the treatment group based on the extracted variables of study duration and relevant working experience. We calculated...
averages for ranked candidates on a per-recruiter basis and then compared between values from both groups using independent-samples t-tests. Once again, we considered rankings from development and marketing job postings separately. The quantitative results are summarized in Table 2.

<table>
<thead>
<tr>
<th>Task</th>
<th>Variable</th>
<th>Group</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>df</th>
<th>t</th>
<th>Sig.</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development</td>
<td>Study duration</td>
<td>Supported</td>
<td>4.909</td>
<td>.324</td>
<td>69</td>
<td>-1.532</td>
<td>.130</td>
<td>.365</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unsupported</td>
<td>4.792</td>
<td>.311</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Working experience</td>
<td>Supported</td>
<td>3.500</td>
<td>.505</td>
<td>69</td>
<td>-2.277</td>
<td>.026</td>
<td>.540</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unsupported</td>
<td>3.194</td>
<td>.622</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marketing</td>
<td>Study duration</td>
<td>Supported</td>
<td>5.063</td>
<td>.264</td>
<td>71</td>
<td>-2.537</td>
<td>.013</td>
<td>.604</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unsupported</td>
<td>4.924</td>
<td>.190</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Working experience</td>
<td>Supported</td>
<td>2.054</td>
<td>.308</td>
<td>71</td>
<td>-.670</td>
<td>.505</td>
<td>.154</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unsupported</td>
<td>1.993</td>
<td>.470</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Study duration and working experience are measured in years and were extracted from the submitted resumes. Results are based on independent-samples t-tests.*

Table 2. Effects of CV recommender system support on KSA levels of ranked candidates

By conducting our quantitative analysis we found that top-10 ranked candidates in CV recommender system supported settings displayed statistically significant stronger levels of KSA’s than top-10 ranked candidates in non-CV recommender system supported settings in two out of four observed cases (development – working experience: t = -2.277, p = .026; marketing – study duration: t = -2.537, p = .013). For both cases we observed medium effect sizes (larger than .5), as measured by Cohen’s d. In addition, we found a small effect size (larger than .2) for study duration in the development task, that was not statistically significant (t = -1.532, p = .130, d = .365). Considering these findings, we can partially support our anticipation that candidates of the top-10 rankings possess stronger KSA’s in the cases when recruiters have been supported by the CV recommender system compared to the cases were recruiters have not received a matching score generated by the CV recommender system.

5 Discussion

To cope with the increasing amount of digital applicant data (Strohmeier and Piazza, 2015; van Esch et al., 2019) and to ensure fairness in the candidate selection phase (Gilliland, 1993; Thielsch et al., 2012; Ochmann and Laumer, 2019), research has increasingly focused on the development of CV recommender systems. These systems serve to identify the most suitable candidates for a given job by calculating the similarities between candidate profiles and job requirements. Thus, CV recommender systems are typically applied in the selection phase of the recruiting process (Schneider, 1981) with the purpose of estimating the P-J fit (Rynes and Gerhart, 1990; Adkins et al., 1994; Wilk and Sackett, 1996; Judge and Cable, 1997). While prior research on CV recommender systems has mainly focused on improving the performance of CV recommender systems on a technical level (e.g., Malinowski et al., 2006; Lu et al., 2013; Bansal et al., 2017), our study addresses the socio-technical context by concentrating on the interaction between the human recruiter and the algorithm (Green and Chen, 2019a; Grgic-Hlaca et al., 2019). In detail, we examine what impact the use of a CV recommender system has on procedural justice in the selection process. Therefore, we conduct an experiment with 74 professional
recruiters from 22 multinational companies, where the task is to create top-10 rankings of candidates for two fictional job postings. According to our quantitative data analysis, we found statistically significant differences between the control and the treatment group with regard to the inner group ranking correlation score. We derive two main findings from our quantitative analysis. First, the analysis of our experiment indicates that the rankings correlated more strongly with each other when recruiters received the matching score generated by the CV recommender system than in the non-CV recommender system supported group. Since this stronger correlation is an indicator that the top-10 ranking list is more consistent and similar among the recruiters of the CV recommender system supported group, we can assume that the level of procedural justice increases through the assistance of the CV recommender system. Second, our quantitative results indicate that the CVs of the top-10 ranked candidates of the CV recommender system supported group contain stronger KSA’s in regard to working experience for the development job posting as well as in regard to study duration for the marketing job posting compared to the non-CV recommender system supported group. Due to the presence of these stronger KSA’s in the top-10 rankings, our results indicate that KSA’s are given more attention when creating the top-10 rankings with the support of a CV recommender system than if recruiters would make the decision on their own. In other words, CV recommender systems can help recruiters to base their decision-making on the pure set of KSA’s, rather than being influenced by their own judgment or personal biases. Thus, if candidates possess KSA’s required for a particular job posting and a CV recommender system is incorporated in the selection phase, the likelihood of these candidates being selected in the top-10 rankings tends to increase. To summarize, we show that incorporating CV recommender systems increases procedural justice in the selection phase as recruiters are more likely to adhere to the rule of consistency (Leventhal, 1980; Arvey and Renz, 1992; Gilliland, 1993) by ranking candidates in a more consistent and uniform manner. Moreover, we find that the candidates selected by CV recommender system supported recruiters typically possess stronger KSA’s than the candidates selected by non-CV recommender system supported recruiters which can be considered as an indicator of ensuring the procedural rule of bias suppression (Leventhal, 1980; Arvey and Renz, 1992).

Our study offers significant theoretical contributions regarding research in the area of human-algorithm interaction. To the best of our knowledge, we are among the first to study the effects of CV recommender system application on procedural justice in the selection phase of the recruiting process. Our study showed that the decision-making process of professional recruiters can be influenced by a CV recommender system by creating more consistent and uniform rankings in which the selected candidates possess stronger KSA’s. Thus, we provide quantitative evidence for findings of Thielisch et al. (2012) and Ochmann and Laumer (2019), i.e., that higher levels of objectivity and consistency can be achieved in the candidate selection phase when using an algorithmic aid instead of solely relying on human judgment. Consequently, we show that procedural justice in the selection phase of recruiting can be strengthened by deploying a CV recommender system. We propose that content-based CV recommender systems help recruiters to ensure the procedural rule of consistency (Leventhal, 1980; Arvey and Renz, 1992; Gilliland, 1993) by providing more consistent and uniform rankings. Moreover, relying on these types of systems might mitigate human biases in the recruiting process, such as subjective selection criteria by enabling accurate measurement of candidate’s KSA’s as proposed by the P-J fit (Caldwell and O Reilly, 1990; Edwards, 1991; Wilk and Sackett, 1996; Kristof-Brown, 2000).

Our findings also have significant implications for practitioners. We show that organizations should consider deploying CV recommender systems in the selection phase of the recruiting process. Here, the application of such systems might serve several purposes. First, using content-based CV recommender systems increases the likelihood that applicants with higher levels of KSA’s will be included in candidate rankings. This way, organizations can ensure that they more strongly consider candidates with a high P-J fit. As a result, more suitable candidates might be identified more efficiently, preventing costly hiring mistakes in the process. Moreover, content-based recommender systems could be used to partly automate the selection process, which would allow to direct further resources towards cognitively more challenging tasks, e.g., estimating the P-O fit via in-person interviews. Second, the deployment of CV recommender systems might reduce existing biases in the recruiting process by making sure that candidate rankings are more consistent across different recruiters.

*Twenty-Ninth European Conference on Information Systems (ECIS 2021), [Marrakesh, Morocco|A Virtual AIS Conference]*
While our study adds value for both research and practice, it is affected by some limitations that offer opportunities for further research. Despite the fact that we designed our experiment according to realistic recruiting standards and practices by involving professional recruiters, providing real CVs, and using a content-based CV recommender system, we are aware that the experimental set-up has some shortcomings regarding the recruiting process in practice. With regard to the selection phase, our study differs from procedures used in practice to evaluate applicants where CVs are usually reviewed as they are received rather than consecutively in batches. In addition, recruiters would usually receive more information on required skills and context from the hiring manager instead of just referring to the available requirements of the job posting. Furthermore, the choice of a junior job posting could have an impact on the matching score generated by the CV recommender system as the submitted CVs might contain fewer keywords and details than for a professional job posting. Lastly, as we used only one commercially available content-based CV recommender system, we are well aware that the results might differ for alternative solutions.

To increase realism, future research could improve the experimental set-up by providing a centralized CV upload that allows recruiters to review CVs at the time of upload. Therefore, the timeframe of the experiment should also be extended from one to three months in order to make the decision-making process of candidate selection more realistic. In addition, the between-subject design could also be varied by adding another treatment group of recruiters who receive additional information on the key features that influence the generated matching score. Thereby, the effect of increased transparency for recruiters compared to recruiting settings with less information could be studied further. By expanding the research scope with a focus on transparency, additional insights could be gained as to whether recruiters would integrate the suggestions of CV recommender systems even more strongly into their decision-making process as the key features become more transparent. This future research would contribute significantly to the study of human-algorithm interactions in the field of recruiting.

6 Conclusion

In recent years, recruiting qualified and skilled talents has gained considerable importance as organizations consider their workforce as strategic assets. While digitization has contributed to the emergence of job portals, recruiters face the challenge of dealing with large amounts of digital applications (Black and van Esch, 2020). In order to cope with this amount of data, CV recommender systems have been developed to support recruiters in the selection phase. By computing the similarity of the candidates KSA’s and the job requirements, CV recommender systems are able to identify the most suitable candidates as requested by the demands-ability approach of the P-J fit (Edwards, 1991; Kristof-Brown, 2000). However, relatively little is known about the socio-technical context in which such systems are deployed (Green and Chen, 2019a; Grgic-Hlaca et al., 2019). Our study aimed to examine the impact of a CV recommender system on procedural justice in the selection phase of the recruiting process. Therefore, we conducted an experiment with 74 recruiters from 22 large multinational companies. Using a between-subject design, we compare top-10 rankings of potential candidates for two fictional job postings between recruiters who are supported by a content-based CV recommender system and unsupported recruiters. Two main observations can be drawn from our quantitative analysis. First, candidate rankings from the CV recommender system supported group exhibit higher levels of similarity than rankings from the non-CV recommender system supported group. Second, candidates selected by recruiters who received the matching score generated by the CV recommender system contain stronger KSA’s than candidates selected by recruiters who relied solely on their own judgment. Thus, we find quantitative evidence that the deployment of CV recommender systems can increase consistency and reduce personal bias in the selection phase of the recruiting process, which might improve procedural justice in this process.
References


Dean, A., D. Voss and D. Draguljić. 2017. Design and analysis of experiments (2nd ed.). Cham: Springer International Publishing AG.


selection in personnel recruitment.” *Ninth Americas Conference on Information Systems.*


Kobsa, & W. Nejdl (Eds.), *The adaptive web* (pp. 325–341). Berlin: Springer-Verlag.