A Just Gender Pay Gap?
Three Factorial Survey Studies on Justice Evaluations of Earnings for Male and Female Employees

Carsten Sauer
Carsten Sauer

A Just Gender Pay Gap? Three Factorial Survey Studies on Justice Evaluations of Earnings for Male and Female Employees

SFB 882 Working Paper Series, No. 29
DFG Research Center (SFB) 882 From Heterogeneities to Inequalities
Research Project A6
Bielefeld, August 2014

SFB 882 Working Paper Series
General Editors: Martin Diewald and Thomas Faist
ISSN 2193-9624

This publication has been funded by the German Research Foundation (DFG).

SFB 882 Working Papers are refereed scholarly papers. Submissions are reviewed by peers in a two-stage SFB 882 internal and external refereeing process before a final decision on publication is made.

The Working Paper Series is a forum for presenting works in progress. Readers should communicate comments on the manuscript directly to the author(s).

The papers can be downloaded from the SFB 882 website http://www.sfb882.uni-bielefeld.de/

SFB 882 “From Heterogeneities to Inequalities”
University of Bielefeld
Faculty of Sociology
PO Box 100131
D-33501 Bielefeld
Germany
Phone: +49-(0)521-106-4942 or +49-(0)521-106-4613
Email: office.sfb882@uni-bielefeld.de
Web: http://www.sfb882.uni-bielefeld.de/
DFG Research Center (SFB) “From Heterogeneities to Inequalities”

Whether fat or thin, male or female, young or old – people are different. Alongside their physical features, they also differ in terms of nationality and ethnicity; in their cultural preferences, lifestyles, attitudes, orientations, and philosophies; in their competencies, qualifications, and traits; and in their professions. But how do such heterogeneities lead to social inequalities? What are the social mechanisms that underlie this process? These are the questions pursued by the DFG Research Center (Sonderforschungsbereich (SFB)) “From Heterogeneities to Inequalities” at Bielefeld University, which was approved by the German Research Foundation (DFG) as “SFB 882” on May 25, 2011.

In the social sciences, research on inequality is dispersed across different research fields such as education, the labor market, equality, migration, health, or gender. One goal of the SFB is to integrate these fields, searching for common mechanisms in the emergence of inequality that can be compiled into a typology. More than fifty senior and junior researchers and the Bielefeld University Library are involved in the SFB. Along with sociologists, it brings together scholars from the Bielefeld University faculties of Business Administration and Economics, Educational Science, Health Science, and Law, as well as from the German Institute for Economic Research (DIW) in Berlin and the University of Erlangen-Nuremberg. In addition to carrying out research, the SFB is concerned to nurture new academic talent, and therefore provides doctoral training in its own integrated Research Training Group. A data infrastructure project has also been launched to archive, prepare, and disseminate the data gathered.
Research Project A6 “The Legitimation of Inequalities – Structural Conditions of Justice Attitudes over the Life-span”

This project investigates (a) the conditions under which inequalities are perceived as problems of justice and (b) how embedment in different social contexts influences the formation of attitudes to justice across the life course.

We assume that individuals evaluate inequalities in terms of whether they consider them just, and that they hold particular attitudes toward justice because, and as long as, these help them to attain their fundamental goals and to solve, especially, the problems that arise through cooperation with other people (cooperative relations). As a result, attitudes on justice are not viewed either as rigidly stable orientations across the life span or as “Sunday best beliefs” i.e. short-lived opinions that are adjusted continuously to fit situational interests. Instead, they are regarded as being shaped by the opportunities for learning and making comparisons in different phases of the life course and different social contexts.

The goal of the project is to use longitudinal survey data to explain why individuals have particular notions of justice. The key aspect is taken to be changes in the social context – particularly households, social networks, or workplaces – in which individuals are embedded across their life course. This is because social contexts offer opportunities to make social comparisons and engage in social learning, processes that are decisive in the formation of particular attitudes to justice. The project will test this empirically by setting up a special longitudinal panel in which the same individuals will be interviewed three times over an 11-year period.

The results of the project will permit conclusions to be drawn on the consequences of changes in a society's social and economic structure for its members' ideas about justice. The project therefore supplements the analysis of the mechanisms that produce inequality, which is the focus of SFB 882 as a whole, by looking at subjective evaluations, and it complements that focus by addressing the mechanisms of attitude formation.

Research goals

(1) Analysis of the conditions in which justice is used as a criterion for evaluating inequalities.
(2) Explanation of attitudes toward justice as the outcome of comparison and learning processes mediated by the social context.
(3) Longitudinal observation of the individual development of attitudes to justice over the life course.

Research design

(1) Continuation and expansion of the longitudinal survey of evaluations of justice conducted by the German Socio-Economic Panel Study (SOEP).
(2) Commencement of an independent longitudinal panel with ties to the process-generated individual data of the German Institute for Employment Research (IAB) and information on companies and households (the plan is to carry out three survey waves over an 11-year period).
The Author

Contact: carsten.sauer@uni-bielefeld.de
A Just Gender Pay Gap? Three Factorial Survey Studies on Justice Evaluations of Earnings for Male and Female Employees

Carsten Sauer
Collaborative Research Center 882, Bielefeld University

Abstract
The study investigates justice evaluations of earnings for male and female employees and links them to the actual inequality people are embedded in. Previous studies in non-reflexive sociological justice research found mixed results. Some studies report a just gender pay gap favoring men; others do not find this gap. This study provides an explanation for these mixed results by combining sociological justice research and status construction theory. Three factorial survey studies were carried out consisting of descriptions of employees with varying characteristics including gender. One study was conducted with social sciences students and two with population samples of German inhabitants. Results show that social sciences students revealed no gender gap in their ratings. In the population surveys, both men and women, showed a rating behavior favoring male employees. The findings indicate that actual inequalities between men and women influence the existence, sign, and size of a just gender pay gap.

Keywords: Justice evaluations, just gender pay gap, gender inequalities, status beliefs, factorial survey

Distributive justice judgments are driven by notions of entitlement. Entitlement, as an expectation with normative force (Singer, 1981), is connected to characteristics of recipients that can be ascribed, such as gender, or achieved, such as performance (Lerner, 1987; Major, 1994). Individuals as observers who hold expectations and evaluate outcomes in regard to justice are embedded in the social structure (Granovetter, 1985), meaning that their position within the social stratification influences their individual attitudes (Kiecolt, 1988) and shapes their status beliefs. Therefore, justice evaluations of men’s and women’s outcomes may reflect status beliefs and are influenced by the inequality structure individuals are embedded in.
It is important to know about how the recipient’s gender affects justice attitudes as these attitudes are, e.g., connected to behavior in negotiations or promotion procedures (Stuhlmacher & Walters, 1999; Kray, Galinsky, & Thompson, 2002; Kray & Thompson, 2004), and therefore directly reinforce actual inequalities between men and women.

Sociological justice research distinguishes between reflexive and non-reflexive justice evaluations (Jasso, 1999, 2007). In reflexive justice evaluations people evaluate their own outcomes (observer = recipient); in non-reflexive justice evaluations people evaluate outcomes of other people (observer ≠ recipient). Previous studies on reflexive justice and related research on satisfaction find a gap between the evaluations of men and women. Men expect higher wages than women (Major & Konar, 1984; Major, 1989, 1994; Sumner & Brown, 1996) and evaluate their higher earnings to a similar extent as being just or unjust (Liebig, Sauer, & Schupp, 2012). But, reflexive justice judgments are based on own outcomes and are therefore driven by two forces, justice deliberations and self-interest (Younts & Mueller, 2001). As only a small fraction of people evaluates themselves as being overpaid (Sauer & Valet, 2013) the gender gap in reflexive evaluations also occurs due to men and women both claiming higher earnings to a roughly similar amount. Thus, it is difficult to disentangle the impact of the justice deliberations and self-interest separately. Non-reflexive judgments are not affected by conflicts of justice perceptions and own interests, because people judge outcomes of which they are not affected. Non-reflexive judgments are, therefore, well suited to investigate justice attitudes in regard to gender.

Previous studies on non-reflexive justice evaluations have yielded mixed results in regard to gender differences. Jasso and Webster (1997) found a just gender pay gap—the difference between earnings evaluated as just for male and female recipients—in their factorial survey study conducted in 1974. Male and female observers (respondents) assigned higher just earnings to male recipients (vignette persons). In a later factorial survey conducted in 1995, using a student sample, they found only a marginal gap favoring women (Jasso & Webster, 1999). Jasso and Webster (1999) interpreted this result in comparison to their previous study (Jasso & Webster, 1997) as a possible consequence of changing gender norms over time.

An alternative explanation is that justice evaluations are related to status beliefs about gender that differ between sub-populations: If gender has a status value (Berger, Fisek, Norman, & Zelditch, 1977) it influences justice evaluations, otherwise not. People with higher status are assigned to receive higher just earnings and vice versa. The likelihood of gender being a status characteristic is influenced by the inequality structure. Status construction theory (Ridgeway, 1991, 2006, 2011) assumes that daily interactions between different categories of people generate and reinforce status hierarchies that are biased by resource inequalities. In a population, in which men earn more than women it is likely that gender becomes a status characteristic. This status characteristic gets then relevant in justice evaluation processes. If gender is a status characteristic, men and women evaluate higher earnings for the status-high
group as being just. Thus, the actual inequality determines the existence, sign, and size of a just gender pay gap produced in evaluation processes. Previous studies on non-reflexive justice evaluations did not link the justice perceptions of individuals to the macro context people are embedded in.

The contribution of this paper is to combine sociological justice research (Jasso, 1978, 1980; Jasso & Wegener, 1997) with status construction theory (Ridgeway, 1991, 2006, 2011) to derive hypotheses about the existence, sign, and size of a just gender pay gap in observers’ evaluations. Moreover, it provides empirical tests based on large respondent samples.

To detect differences in justice attitudes it is necessary to, first, use a method that allows to find gender gaps. Thus, the method should be less prone to social desirability bias than item-based non-reflexive justice evaluations are, as it is likely that people do not reveal their true preferences if they are asked directly whether or not they think that men should earn more than women. The data collection method used here is, therefore, a factorial survey design (Rossi & Anderson, 1982; Jasso, 2006; Wallander, 2009), in which respondents evaluated so called vignettes that described persons varying in multiple characteristics, including gender, and gross earnings. These vignette-based non-reflexive justice evaluations allow to measure the impact of the recipient’s gender and other characteristics on the justice evaluation of observers independently. In regard to this feature, factorial surveys have also an advantage compared to reflexive justice measures, as gender can be modeled as uncorrelated with other recipient’s characteristics, e.g., occupational status and gross earnings, which are confounded in the real world. Second, it is necessary to compare observers who are embedded in differing inequality structures. The empirical basis of this research, therefore, consists of three datasets: a sample of social sciences students and two population samples. The social sciences students are embedded in a structural context in which relevant resources are not correlated with gender and therewith, gender is not likely to become a status characteristic in their daily interactions. The respondents of the population samples were sampled in different regions in Germany with differing earnings inequality between men and women. Thus, it is possible to investigate justice evaluations of people embedded in differing gender inequality structures.

The findings reveal that the actual inequality people are embedded in plays a crucial role for gender preferences in their justice evaluations. Social sciences students showed no gender gap in their justice ratings. In both population samples female and male respondents produced a just gender pay gap in their ratings favoring male recipients. The size of the gap was bigger for those male and female respondents who lived in regions with a larger earnings inequality between men and women.

1The method is robust against social desirability bias due to the multifactorial design (Mutz, 2011; Wason, Polonsky, & Hyman, 2002). It has been used in studies with sensitive topics (e.g., Graeff, Sattler, Mehlkop, & Sauer, 2013; Sattler, Sauer, Mehlkop, & Graeff, 2013).
The Justice Evaluation Process

In justice evaluation processes people compare actual rewards to rewards perceived as just (Jasso, 1978, 1980, 1986). Jasso (1978) specifies this association with the following formula:

\[ J = \ln \frac{A}{C} = \ln A - \ln C. \]  

(1)

The justice evaluation \( J \) of an observer can be represented by the logarithmic ratio of the actual rewards \( A \) and the just rewards \( C \) of a recipient. The specification assumes a comparison process to be the central mechanism within justice evaluations. In case of justice evaluations of earnings the actual rewards (earnings) of a recipient are directly measurable. The just earning are amounts the observer considers as just for given recipients. This judgment is based on factors that the observer regards as important. But the specification leaves exogenous the substantive content of the just reward function (Jasso, 1980; Markovsky, 1985). Jasso and Wegener (1997) specify that the just reward depends on the reward-relevant factors \( x \), their weights and their combination. Thus,

\[ C = h(x_1, x_2, \ldots, x_n). \]  

(2)

To learn about the content of these factors theories are useful that provide substantive predictions. This study focuses on the relevance and weight of gender in justice evaluations, thus predictions for the relevance and weight of this factor are required. Reward relevant characteristics are those that entitle the trait carrier to receive a certain amount of rewards. These characteristics can be achieved, like performance, or ascribed, like gender (Berger et al., 1972). If these characteristics have a status value they can be defined as status characteristic (Berger et al., 1977). Status characteristics divide trait carriers in status-high and status-low individuals and entitle status-high individuals to receive higher rewards. The status value is not an intrinsic feature of a characteristic (in this case gender) but attached to the characteristic by generally shared beliefs. Reward expectations theory connects status characteristics to reward expectations and perceptions of justice and injustice (Berger, Fisek, Norman, & D. G. Wagner, 1985). Reward expectations are formed based on status characteristics and a referential structure. Berger et al. (1985) distinguish three types of referential structures: categorical referential structures are based on “who you are,” ability referential structures are based on “what you can do,” and performance-outcome referential structures are based on “what you have


\(^3\)The combination of these factors, deals with interactions between gender and other characteristics, e.g., experience or education. The theory of double standards focuses on these interactions between job related characteristics, like competence and performance, and gender (Foschi, Lai, & Sigerson, 1994; Foschi, 2000); empirical research by Jasso and Webster (1999) shows that these double standards exit in justice evaluations.
done.” Reward expectations theory implies that categorical, ability and performance-outcome characteristics may together determine reward expectations and therefore justice evaluations. Thus, status characteristics that refer to categorical differences, abilities or inputs are relevant for the observer to estimate the just earnings $C$ of a recipient. Assuming this evaluation process the justice evaluation stated in Equations 1 and 2 contains three types of characteristics: categorical variables, abilities and inputs. Gender is a categorical difference between recipients. If gender has a status value in the eyes of the observer, it will be relevant in the justice evaluation process. It is assumed, that the gender gap in just wages found in earlier studies (Jasso, 1994; Jasso & Webster, 1997; Jann, 2008) occurred because gender had a status value, dividing people in status-low and status-high groups. In other words, the existence, sign and size of a just gender wage gap is connected to the status value of this characteristic. If gender has no status value in the eyes of the observer, it is not a relevant factor for the justice evaluation. The observers produce in this case no just gender pay gap. To put this in a formal equation:

$$J = \beta_1 \times \text{gender} + \ldots + \beta_n \times \ln A.$$  

The term $C$ in Equation 1 is now replaced by characteristics that might be relevant for the justice evaluation including gender. $J$ is a function of the actual earnings ($A$) and the characteristics being evaluated as relevant for the assessment of the just reward.

The questions are, how do inequalities between men and women influence the existence ($\beta_1 \neq 0$), sign ($\beta_1 \leq 0$) and size of a just gender pay gap? To link the justice evaluation process to the structural context a theory is needed that predicts under which structural condition it becomes likely that gender becomes a status characteristic.

**Status Construction and Reward Expectations**

Status construction theory (SCT) provides an explanation for the evolution and diffusion of status beliefs regarding nominal groups in a society. SCT argues that, under some macro conditions and micro assumptions, it is likely that beliefs regarding status characteristics are commonly shared in a population (Ridgeway, 1991, 2006, 2011).

The macro assumptions focus on the distribution of characteristics and resources: first, a nominal characteristic $N$ divides a population into at least two categorical groups ($A$s and $B$s). Second, $R$ is a scarce resource that may take on at least two states (resource-rich and resource-poor). This resource is assumed to be unequally distributed between the groups. Finally a probability exists that members of the two groups interact in their daily lives. The macro assumptions may be stated as follows (Ridgeway, 1991):

1. Unequally distributed resources between $A$s and $B$s: $R_A \neq R_B$
2. Differences with regard to a nominal characteristic $N$ between $A$s and $B$s: $N_A \neq N_B$.

3. A correspondence between the resource and the categorical characteristic: $\rho(N, R) \neq 0$.

4. A probability of contact of $A$s and $B$s unequal to zero: $p_I(A, B) > 0$.

On the micro-level, SCT assumes that a status hierarchy is likely to emerge when people interact in goal-orientated tasks (Berger, D. G. Wagner, & Zelditch, 1985; D. G. Wagner & Berger, 2002). This hierarchy develops via individual behavior within interactions and results in expectations formed by the actors about each interactor’s contribution to the group task. In this formation process, it is not only relevant how people act in a specific situation but also what initial external resources they obtain, as these resources influence their confidence and non-verbal behavior and the expectations of others. Empirical studies clearly show that resource-rich people tend to behave more self-confidently and are less willing to compromise (Berger, Fisek, Norman, & D. G. Wagner, 1985; Stewart & Moore Jr, 1992). They also tend to gain a higher position in the immediate hierarchy-formation process (Ridgeway, Boyle, Kuipers, & Robinson, 1998; Ridgeway & Erickson, 2000). Resources are therefore a mediator in the emergence of status hierarchies in local contexts.

Given these assumptions, SCT predicts that it is possible for people to meet in doubly dissimilar encounters (Ridgeway, 1991, 2006). In doubly dissimilar encounters, actors differ in the nominal characteristic (e.g., gender) and in the resource endowment. In these situations, it is likely that the correspondence between the nominal characteristic and the resource endowment will be recognized by individuals and internalized in their behavior in other local contexts. The hierarchy formation process is biased by resource inequality. Diffusion processes spread these status beliefs over wide portions of a population and form commonly-shared status beliefs. As Ridgeway et al. (1998) showed, although resource inequality is a mediator, it is not causal for the emergence of status beliefs. Interactions in which members of the high-pay group act in a self-confident and dominant way and members of the low-pay group act in a self-effacing and deferring manner are highly relevant for status-belief formation. This process explains how a categorical characteristic like gender becomes a status characteristic.\footnote{A lot of papers tested the mechanisms (Ridgeway & Balkwell, 1997; Ridgeway et al., 1998; Ridgeway & Erickson, 2000; Ridgeway & Correll, 2006; Brashears, 2008) proposed by SCT and further developed the theory (e.g., Webster & Hysom, 1998; Walker, Webster, & Bianchi, 2011; Berger & Fisek, 2013).}

Hypotheses

Combining sociological justice theories and status construction theory it can be predicted under what macro conditions gender is likely to become a status chara-
teristic and therewith a relevant factor in the justice evaluation process formulated in Equation 3. Under the macro condition of resource equality it is likely that status hierarchies emerge in which gender is not important. Status hierarchies are in this case not correlated with gender and gender becomes no status characteristic. In a sub-population with resource equality the justice evaluation of the observer should not be affected by the gender of the recipient. The hypothesis refers to the question of the existence of a just gender pay gap.

**Hypothesis 1** In a sub-population with resource equality between men and women, it is likely that male and female observers will not attach a status value to the characteristic gender of the recipient. Observers do not produce a just gender pay gap with their ratings.

Under the macro condition of resource inequality in which gender is correlated with resource endowment, it is likely that status hierarchies emerge that are correlated with gender. If men are more likely to be resource-rich and women are more likely to be resource-poor, observers attach higher status to male recipients and assign higher earnings to the high-status group even though the recipients do not differ in other characteristics. This high-status group preference is shared by both the advantaged and disadvantaged groups and, accordingly, both male and female observers assign higher earnings to male recipients. The hypothesis refers to the question of the sign of a just gender pay gap.

**Hypothesis 2** In a sub-population in which men earn on the average more than women, it is likely that male and female observers produce a just gender pay gap favoring male recipients.

Moreover, the inequality structure does not only influence the relevance (relevant or not) but also the weight of the factor gender for the justice evaluation. More inequality between categorical groups leads to stronger correlations between the resources, in this case earnings, and the categorical characteristic gender. The more unequally earnings are distributed within a sub-population the more pertinent is gender as status characteristic. Therefore, the weight of this factor in the justice evaluation process depends on the skewness of the resource distribution. The hypothesis refers to the question of the size of a just gender pay gap.

**Hypothesis 3** The bigger the actual gender pay gap in the sub-population of the observer, the higher will be the just gender pay gap in the ratings.

**Methods**

**Respondents**

In order to test the hypotheses the analyses are based on factorial surveys conducted with one student sample and two general population samples. Both population
samples are independent of each other and have been conducted by different research institutions. The reason for having two population samples is replication, meaning that the hypotheses are tested with different data sources.

**Student Sample.** The student survey was conducted during the summer term in 2008. Students of social sciences from 27 universities all over Germany were interviewed via computer-assisted web interviews (CAWI) and computer-assisted self-interviews (CASI) in labs and in the presence of research assistants. The questionnaires consisted of the factorial survey module and additional questions on attitudes (after the factorial survey module) and questions on the socio-demographic background of the parents and students’ personal situation. The analysis sample consists of 1764 respondents. The student sample available is an example for equality in respect of two resources important for the emergence of status hierarchies: First, on the average the incomes of female and male students are equal (Isserstedt, Middendorff, Kandulla, Borchert, & Leszczensky, 2010). Second, in the study at hand study success was measured via self-assessment on an eleven-point rating scale (−5 to +5). There was no gender difference in these assessments (mean$_m$ = 1.26; mean$_f$ = 1.15; $T = 1.27; p = .20; n_m = 697; n_f = 998$) in the data used here. The resource endowment (income and performance) was uncorrelated with gender. Thus, the data are used to test Hypothesis 1.

**Population Sample 1.** The survey carried out in 2009 consists of randomly sampled respondents over 17 years of age, interviewed via computer-assisted personal interviews (CAPI) or self-administered interviews (paper and pencil [PAPI] or web interviews [CAWI]). The survey was conducted by a research institution with professional interviewers. The questionnaire consisted of the factorial survey module and additional questions on attitudes (after the factorial survey module) and questions on the socio-demographic background. As factorial survey studies go beyond the standard questionnaires the requirement in the CAPI version was to use experienced interviewers. Additionally, on two days training courses were provided by the researchers to show the interviewers how the respondents had to rate the vignette task and how the interviewers had to behave during the respondents rated the vignettes and how to react in the case of questions. The analysis sample consists of 1411 respondents.\(^5\)

**Population Sample 2.** The data from the survey conducted in 2008 were collected in a pretest of the German Socio-Economic Panel (SOEP; Schupp, 2009; G. Wagner, Frick, & Schupp, 2007) via computer-assisted personal interviews (CAPI). The program of the annual SOEP questionnaire for the following wave is pretested

\(^5\)In the CAPIs additional information were collected about the interviewer, e.g., work experience and training attendance, and the interview situations in interviewer questionnaires after each interview. With these pieces of information it was possible to find those interviews which did not fulfill the requirements for the analysis: interviews were excluded from the analysis sample if respondents did not do the task on their own but with the help of others. Moreover, three out of 81 interviewers had a tenure of one year or less and did not show up in one of the two training days. Their interviews were not used as they did not fulfill the requirements.
in each summer of the preceding year. The objective of these pretests is to test new modules and modifications of questions. Since 2002 the sample size is around 1,000 respondents and considered representative for the German resident population of 16 years and older (Siegel, Stocker, & Warnholz, 2009). There are two main differences between the pretest and the SOEP main survey. First, all interviews in the SOEP-Pretest are programmed as computer assisted personal interviews (CAPI), in contrast to paper and pencil questionnaires mostly used in the main survey. Second, whereas the main survey is a study of private households, the SOEP-Pretest is a sample of individuals. The pretest sample is not related to the main survey, meaning that these respondents are not part of the panel study. Due to the programming of the vignette module, it was not possible for the respondents to skip vignettes. The problem is described in more detail in Sauer, Auspurg, Hinz, Liebig, and Schupp (2009, 2014). Thus, interviews with less than 5 minutes processing time of the vignette module (less than 12 seconds processing time per vignette) were discarded from the analysis sample. The analysis is based on the judgments of 952 respondents.

The two population samples are useful to test (Hypothesis 2). In Germany, the gender pay gap is persistent in recent years and is with about 22 percent (Federal Statistical Office, 2010) higher than in most other countries.

Context Variables. To test Hypothesis 3 the average earnings of full-time employees and the actual gender pay gap (GPG) in different federal states in Germany were attached to the survey data. There exist large regional differences in gender pay gaps between federal states. The lowest pay gap in 2009 of full time employed people was measured in Saxony-Anhalt with 1 percent. The largest gap was measured in Baden-Württemberg with 28 percent. Table 8 provides the median earnings and gender pay gaps for each state separately. Therefore, this context variable is useful to compare how the gender of the recipient influences justice evaluations of observers living in different federal states.

Factorial Survey

The vignettes consisted of fictitious earners in full-time (40 hours per week) employment. Each vignette provided at least information on the gender, age, education, and occupation of the recipient described, among other dimensions in more complex settings, together with gross earnings.

The vignette samples were drawn via a quota design (D-efficient design) under exclusion of illogical or implausible cases (Dülmer, 2007). Illogical cases are, e.g.

---

6 The data of the average earnings of full-time employees are provided by the Federal Employment Agency (Frank & Grimm, 2010, p. 14). The gender pay gaps per federal state are provided by the Statistical Office of Rhineland-Palatinate (Schomaker, 2010b, 2010a).

7 D-efficient designs (Kuhfeld, Tobias, & Garratt, 1994; Kuhfeld, 2005) are built using a computer algorithm that specifies a sample characterized by a minimal intercorrelation between all or the most important dimensions (and interaction terms) and at the same time a maximal variance and balance of the frequency of the vignette levels. These designs ensure that the influence of vignette
medical doctors without university degree. The sampling technique ensured that the correlation of the ascribed characteristic gender and the other characteristics, e.g., occupation or gross earnings, was very low; therefore, no gender pay gap existed in the vignette samples.

In the student survey and the population sample 1, the number of dimensions (5, 8 and 12 dimensions) and the number of vignettes presented (10, 20 or 30 vignettes for each respondent) were varied in a between-subject design. Both studies used the same vignettes. In the population sample 2, a constant number of dimensions (10) and vignettes (24) was presented. This paper only focuses on five dimensions that were included in all studies (number of levels in brackets):

- Gender: [2] male, female
- Age: [4] 25, 35, 45, 55 years
- Education: [3] without vocational training, with vocational training, with university degree
- Occupation: [10] manufacturing laborer, doorman, locomotive engine driver, administrative associate professional, hairdresser, social work professional, computer programmer, electrical engineer, general manager, medical doctor
- Gross earnings per month (Euro): [10] 500, 950, 1200, 1500, 2500, 3800, 5400, 6800, 10000, 15000

The Tables 4, 5 and 6 in the Appendix provide information on the correlation structure of the vignette dimensions used for the analysis. The correlations between the vignette dimension gender and the other dimensions are in all three datasets low, which is a requirement to test just pay gaps without a priori inequality in the data. More details on the methodical setup of these studies can be found in Sauer et al. (2009, 2011, 2014).

dimensions and interaction terms are mutually uncorrelated. In addition, the design features lead to minimal standard errors in data analyses and, therefore, in comparison to other designs (like random samples), a higher statistical “power” and efficiency (in a statistical sense) to reveal the influence of single dimensions. The D-efficiency for all vignette samples was above 90 (ranging form 0 to 100).

8The variations were part of a method experiment that investigated effects of information load and fatigue during the interview. The results show only small effects of information load (number of dimensions) and fatigue (Sauer, Auspurg, Hinz, & Liebig, 2011) that do not affect the results presented here. The respondents were randomly assigned to experimental splits.

9Due to the design, the correlations between the dimension gender and the omitted dimensions are very low and the exclusion of other dimensions in the analysis does not affect the results.

10The categories are based on the percentiles of the income distribution of full-time employees 2007 in Germany (data source: Socio-economic Panel Study, 2007). The highest and lowest categories are added to have extreme cases.
Rating Task. The respondents’ justice judgments of gross earnings were obtained using two different rating procedures. In the student sample and the population sample 1 respondents were asked to submit their judgments of each vignette via an 11-point rating scale. The left extreme point (−5) was labeled “unjustly low,” the midpoint (0) was labeled “just” and the right extreme point (+5) was labeled “unjustly high.” The midpoint was coded as zero, the left segment as negative numbers, and the right segment as positive numbers. The population sample 2 had a three-stage rating task. First, respondents had to judge whether the income of a worker was just or unjust. If respondents rated the income as just, they were forwarded to the next vignette. If they rated the income as unjust, respondents judged in a second step whether the income was too high or too low. Third, the respondents stated the level of injustice on a 100-point scale. To achieve consistency with the two other samples—in which positive numbers indicate over-reward and negative numbers indicate under-reward—the ratings were transformed into a new scale in which perfect justice was coded as zero and the ratings that indicated under-reward were coded negatively. Thus, the new scale runs from −100 to 0 to +100. Table 7 in the Appendix provides descriptive statistics of the evaluations for the three datasets.

In all surveys, the respondents had the opportunity to change their judgments of earlier vignettes when they compared them to later vignettes and had to adjust the ratings. This possibility was introduced in the description of the vignette task immediately before the first vignette. Moreover, in all survey modes, including CAPI, the respondents evaluated the vignettes self-administered. In the CAPI versions the interviewers gave the laptop to the respondents and set opposite to them to not have the opportunity to look at the evaluations.

Analysis

The data structure of factorial surveys is hierarchical, as each respondent rates several vignettes. Because the assumption of uncorrelated error terms does not hold and standard Ordinary Least Square (OLS) regression models would therefore be biased (Hox, Kreft, & Hermkens, 1991; Cameron & Trivedi, 2005, 2009), the data were analyzed via Generalized Least Square (GLS) regression models.\(^\text{11}\)

The model in Equation 4 specifies that the justice evaluation \(J\) of vignette \(v\) of the \(i\)-th respondent is based on the given dimensions of each vignette. The outcome variable in the following regression models is the \(z\)-standardized justice evaluation per vignette.\(^\text{12}\) The independent variables are the five dimensions gender \((1 = \text{male})\), age, education (dummy coded as follows: ref = without vocational training; 1 = vocational training; 2 = university degree), occupation, and gross earnings. The occupation was

---

\(^\text{11}\)The models were estimated with the statistical software Stata 12.1 (StataCorp, 2011). The user-written program estout (Jann, 2005, 2007) was used to format the tables.

\(^\text{12}\)This is different to the approach of Jasso and Webster (1997, 1999) who estimate the just earnings based on the vignette judgment (see Jasso, 2012; Markovsky & Eriksson, 2012a, 2012b, for a critical debate).
transformed to a metric scale using the Standard International Occupational Prestige Scale (SIOPS; Ganzeboom & Treiman, 1996). Furthermore, according to the assumed evaluation process of Equation 3 the logarithmic representation of gross earnings was used. The regression equation (4) displays the models with an attached intercept \( (\beta_0) \), a respondent-specific residual \( (v_i) \) and an error term \( \epsilon_{iv} \). Equation 4 was used to estimate the results for the three models presented in Table 1.

\[
J_{iv} = \beta_0 + \beta_1 \times gender + \cdots + \beta_7 \times lnearnings + v_i + \epsilon_{iv}.
\] (4)

Furthermore, it is assumed that status beliefs differ between the population samples and the student sample. Additionally, both female and male respondents of the population samples are assumed to have similar status beliefs about the nominal characteristic gender. Equation 5 includes besides the gender of the vignette person \( (gender^{recipient}) \) and the other dimensions, the gender of the respondent \( (gender^{observer}) \), and a cross-level interaction term. Equation 5 was used to estimate the results presented in Table 2.

\[
J_{iv} = \beta_0 + \beta_1 \times gender^{recipient} + \beta_2 \times gender^{observer} + \\
\beta_3 \times gender^{recipient} \times gender^{observer} + \cdots + v_i + \epsilon_{iv}.
\] (5)

Additionally, it is assumed that there are differences between people living in federal states with high and low gender inequality. Thus, there should be an interaction effect between the vignette dimension gender and the actual gender pay gap in the federal state. Equation 6 includes the vignette dimensions, the structural context (the actual gender pay gap \( (GPG) \) and the average earnings per federal state), and the cross-level interaction between vignette persons gender \( (gender^{recipient}) \) and the gender pay gap in the federal state \( (GPG^{fed.state}) \). The analysis sample was restricted to those respondents who were full-time employed as gender pay gaps were available only for full-time employees, so they experience the difference in their daily interactions directly. The results are presented in Table 3. Additionally, the interaction effects were estimated for male and female respondents separately.

\[
J_{iv} = \beta_0 + \beta_1 \times gender^{recipient} + \beta_2 \times GPG^{fed.state} + \\
\beta_3 \times gender^{recipient} \times GPG^{fed.state} + \cdots + v_i + \epsilon_{iv}.
\] (6)

Results

Just Gender Pay Gap in Vignette Evaluations

The estimates of the GLS-models of the different respondent samples are presented in Table 1. First, the focus is on the effect of the vignette dimension gender for each dataset. In the student sample, the effect of gender on the justice evaluation is insignificant, which indicates that a minor importance is attached to this dimension. The second model in Table 1 provides the estimations of the population sample 1.
The effect of the dimension gender is—contrary to the findings in the first sample—highly significant. The negative coefficient indicates that male recipients were more often evaluated as under-rewarded than female recipients. In other words, respondents produce with their ratings a just gender pay gap as found by Jasso and Webster (1997). The third model in Table 1 provides the coefficients for the population sample 2. As in the previous model, the effect of the dimension gender is negative, indicating a rating behavior preferring male recipients. Moreover, the coefficient is similar to that in the population sample 1; thus, it is a robust result due to its occurrence in two independent population samples. The gender coefficients of the student sample and both population samples were tested against each other in an interaction model and are significantly different ($p < 0.001$).

The other coefficients and their interpretation are reported briefly, as follows: the effect of the dimension age is negative and highly significant, meaning that older vignette persons were evaluated more often as under-rewarded than younger vignette subjects. The effect of the SIOPS has a significantly negative value, meaning that those vignette persons described to work in occupations with higher prestige scores were evaluated as more under-rewarded than those with lower scores (occupation status reward). The effects of vocational training and university degree are also significantly negative. The reference category is the dimension level without vocational training. According to the respondents, the vignette persons who have a higher level of formal education should gain higher returns from their work (educational reward). Finally, the effect of gross earnings is positive because the more a vignette person earns, the more often respondents rated this person as over-rewarded holding other dimensions equal.

In sum, age, education, occupation, and the associated earnings provided information on the recipient that all respondents used in their justice evaluation. There seems to be a general agreement on the importance of these specific characteristics in justice evaluations of earnings; the coefficients including the constant are very similar. The only exception is the dimension gender, which was not important for students but crucial for the respondents in the two population samples. One must keep in mind that the vignettes of the student sample and the population sample 1 were designed equally, so differences can be attributed to rating behavior and not to design elements. On the other hand, the rating task differed between the population sample 1 and the population sample 2, thus, their similar evaluation patterns indicate reliable results and a robust design.

**Just Gender Pay Gap by Respondents’ Gender**

The results in Table 1 provided information on the overall difference between respondents of the student sample and the population samples. To gain insight into whether these rating patterns were similar for both male and female respondents, as it is suggested by Hypothesis 2, respondents’ gender was included in the regression. The models for the different samples are provided in Table 2. Model 1 and 2 report
Table 1

Multiple Linear Regression (GLS) of Justice Evaluations of Vignettes on Vignette Dimensions by Sample

<table>
<thead>
<tr>
<th></th>
<th>Student sample</th>
<th>Population sample 1</th>
<th>Population sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex [1 = male]</td>
<td>-.003</td>
<td>-.068***</td>
<td>-.074***</td>
</tr>
<tr>
<td></td>
<td>(.007)</td>
<td>(.007)</td>
<td>(.008)</td>
</tr>
<tr>
<td>Age</td>
<td>-.018***</td>
<td>-.024***</td>
<td>-.019***</td>
</tr>
<tr>
<td></td>
<td>(.003)</td>
<td>(.003)</td>
<td>(.003)</td>
</tr>
<tr>
<td>SIOPS</td>
<td>-.014***</td>
<td>-.014***</td>
<td>-.011***</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
<tr>
<td>Without vocational training</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>Vocational training</td>
<td>-.204***</td>
<td>-.127***</td>
<td>-.095***</td>
</tr>
<tr>
<td></td>
<td>(.008)</td>
<td>(.008)</td>
<td>(.010)</td>
</tr>
<tr>
<td>University degree</td>
<td>-.300***</td>
<td>-.198***</td>
<td>-.132***</td>
</tr>
<tr>
<td></td>
<td>(.009)</td>
<td>(.008)</td>
<td>(.010)</td>
</tr>
<tr>
<td>Earnings per month [ln]</td>
<td>.845***</td>
<td>.888***</td>
<td>.856***</td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td>(.004)</td>
<td>(.004)</td>
</tr>
<tr>
<td>Constant</td>
<td>-5.816***</td>
<td>-6.154***</td>
<td>-6.129***</td>
</tr>
<tr>
<td></td>
<td>(.031)</td>
<td>(.030)</td>
<td>(.035)</td>
</tr>
<tr>
<td>R²</td>
<td>.659</td>
<td>.744</td>
<td>.664</td>
</tr>
<tr>
<td>Vignettes</td>
<td>29121</td>
<td>23213</td>
<td>22848</td>
</tr>
<tr>
<td>Respondents</td>
<td>1734</td>
<td>1411</td>
<td>952</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p<.05, ** p<.01, *** p<.001 (two-tailed t-tests)

the coefficients for the student sample. Model 1 shows that the effect of respondents’ gender on the justice evaluations is significantly negative, meaning that male students evaluated on the average the vignettes as more unjustly low than female students. The interaction coefficient between the gender of the vignette person and the gender of the respondent in Model 2 indicates whether or not there were differences in the rating behavior between men and women. The interaction effect is significantly negative, meaning that the rating behavior of male and female students was different in regard to the dimension gender. Male students showed an insignificant tendency to favor male recipients (negative interaction effect), whereas female students showed an insignificant tendency to favor female recipients (positive main effect of the dimension gender). Thus, the rating patterns differ significantly but the main result for both groups is equal: they did not account for gender (statistically insignificant) in their justice evaluations. The Models 3 and 4 show the coefficients for the population sample 1. Model 3 indicates that male and female respondents evaluated the vignettes on the average to an equal extent as being just or unjust. The interaction effect in Model 4 is insignificant meaning that male and female respondents both produced to the same extent a just gender pay gap favoring male recipients with their evaluations. The Models 5 and 6 show the coefficients for the population sample 2. The results are
very similar to those of the first population sample. The results confirm Hypothesis 2 and show again similar coefficients indicating reliable results.

Table 2

Multiple Linear Regression (GLS) of Justice Evaluations of Vignettes on Vignette Dimensions by Sample

<table>
<thead>
<tr>
<th>Vignette level</th>
<th>Student sample</th>
<th>Population sample 1</th>
<th>Population sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Genderrecipient [1 = male]</td>
<td>-.003</td>
<td>.009</td>
<td>-0.068***</td>
</tr>
<tr>
<td></td>
<td>(.007)</td>
<td>(.009)</td>
<td>(.007)</td>
</tr>
<tr>
<td>Age</td>
<td>-.018***</td>
<td>-.018***</td>
<td>-.024***</td>
</tr>
<tr>
<td></td>
<td>(.003)</td>
<td>(.003)</td>
<td>(.003)</td>
</tr>
<tr>
<td>SIOPS</td>
<td>-.014***</td>
<td>-.014***</td>
<td>-.014***</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
<tr>
<td>Without vocational training</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>Vocational training</td>
<td>-.204***</td>
<td>-.203***</td>
<td>-.127***</td>
</tr>
<tr>
<td></td>
<td>(.008)</td>
<td>(.008)</td>
<td>(.008)</td>
</tr>
<tr>
<td>University degree</td>
<td>-.300***</td>
<td>-.300***</td>
<td>-.198***</td>
</tr>
<tr>
<td></td>
<td>(.009)</td>
<td>(.009)</td>
<td>(.008)</td>
</tr>
<tr>
<td>Earnings per month [ln]</td>
<td>.845***</td>
<td>.845***</td>
<td>.888***</td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td>(.004)</td>
<td>(.004)</td>
</tr>
</tbody>
</table>

| Respondent level                    |                |                     |                     |
| Genderobserver [1 = male]           | -.052***       | -.036*              | -.005               |
|                                     | (.013)         | (.015)              | (.013)              |
| Cross-level interaction             | -.031*         | .011                | .015                |
| Genderobserver × genderrecipient    |                |                     |                     |
| Constant                            | -5.795***      | -5.801***           | -6.152***           |
|                                     | (.031)         | (.031)              | (.030)              |

| R²                                  | .659           | .659                | .744                |
| Vignettes                           | 29121          | 29121               | 23213               |
| Respondents                         | 1734           | 1734                | 1411                |

Standard errors in parentheses
* p<.05, ** p<.01, *** p<.001 (two-tailed t-tests)

Just Gender Pay Gap and Structural Context

To investigate how structural differences shape justice perceptions and to test Hypothesis 3 the focus is now on the two population samples. The analysis was restricted to full time employed respondents as they were directly affected by the actual gender pay gaps in the different federal states. The results are presented in
Table 3. Models 1 to 3 show the coefficients for the population sample 1. Model 1 includes the structural variables average gross earnings and gender pay gaps per federal state. The effect of the average gross earnings is significantly negative, meaning that respondents living in federal states with high average earnings evaluated the gross earnings described on the vignettes more often as unjustly low compared to those respondents living in federal states with lower average earnings. This reflects the higher comparison standards observers from income high federal states have in mind during the evaluation process. The gender pay gap in the federal state did not affect the justice evaluation directly. The second model includes the interaction term between the vignette dimension gender and the gender pay gap in the federal state. The effect is significantly negative, meaning that the higher the gender pay gap in the federal state was the bigger was the just gender pay gap produced by respondents’ ratings. The main effect of the vignette dimension gender is insignificant indicating that there was no gender bias in the evaluations if the actual gender pay gap is zero. The third model shows the effects for male and female respondents separately. Again, the rating pattern was equal for male and female respondents (interaction coefficients are not statistically different). Models 4 to 6 show the coefficients for the population sample 2. The effects are very similar to those described above, again the interaction effects in Model 5 and Model 6 are significantly negative and the main effect of the dimension gender vanishes. This result confirms Hypothesis 3. Moreover, all coefficients are very similar in size and significance in both samples, even though the rating task was different, indicating stable results.

Discussion

The study investigated justice evaluations of earnings for male and female employees and linked them to the actual inequalities. The goal was to explain mixed results reported in previous studies on the just gender pay gap in non-reflexive justice evaluations (Jasso & Webster, 1997, 1999). Therefore, predictions of sociological justice theories (Jasso, 1978, 1980; Jasso & Webster, 1997; Berger et al., 1972; Berger, Fisek, Norman, & D. G. Wagner, 1985) and status construction theory (Ridgeway, 1991, 2006, 2011) were combined. It was assumed that resource inequalities between men and women lead to status differences. In the eyes of the observers, the status differences entitle male and female recipients to different just rewards. The existence, sign, and size of a just gender pay gap depends on actual inequalities. The empirical tests support the assumption of this mechanism.

The analysis was based on a factorial survey design (Rossi & Anderson, 1982; Jasso, 2006) conducted with one student sample and two population samples. The results show, that male and female students did not produce a just gender pay gap with their evaluations. Social sciences students are an example for a gender equal sub-population. In this population it is unlikely that gender becomes a status characteristic and therefore it is not a relevant characteristic within the justice evaluation process. One has to keep in mind that students are not only embedded in the struc-
Table 3
Multiple Linear Regression (GLS) of Justice Evaluations of Vignettes on Vignette Dimensions and Context Variables (full-time employees)

<table>
<thead>
<tr>
<th>Vignette level</th>
<th>Population sample 1</th>
<th>Population sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Genderrecipient [1 = male]</td>
<td>-.056***</td>
<td>-.023***</td>
</tr>
<tr>
<td></td>
<td>(.011)</td>
<td>(.005)</td>
</tr>
<tr>
<td>Age</td>
<td>-.023***</td>
<td>-.022***</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
<tr>
<td>SIOPS</td>
<td>-.015***</td>
<td>-.015***</td>
</tr>
<tr>
<td></td>
<td>(.000)</td>
<td>(.000)</td>
</tr>
<tr>
<td>Without vocational training</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>Vocational training</td>
<td>-.121***</td>
<td>-.122***</td>
</tr>
<tr>
<td></td>
<td>(.014)</td>
<td>(.014)</td>
</tr>
<tr>
<td>University degree</td>
<td>-.188***</td>
<td>-.188***</td>
</tr>
<tr>
<td></td>
<td>(.014)</td>
<td>(.014)</td>
</tr>
<tr>
<td>Earnings per month [ln]</td>
<td>.910***</td>
<td>.910***</td>
</tr>
<tr>
<td></td>
<td>(.006)</td>
<td>(.006)</td>
</tr>
<tr>
<td>Structural context</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average gross earnings of fed. state</td>
<td>-.230**</td>
<td>-.230**</td>
</tr>
<tr>
<td></td>
<td>(.084)</td>
<td>(.084)</td>
</tr>
<tr>
<td>Gender pay gap (GPG) in fed. state</td>
<td>.011</td>
<td>.030</td>
</tr>
<tr>
<td></td>
<td>(.039)</td>
<td>(.040)</td>
</tr>
<tr>
<td>Cross-level interaction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genderrecipient × GPG\text{fed. state}</td>
<td>-0.39*</td>
<td>-0.40*</td>
</tr>
<tr>
<td></td>
<td>(.017)</td>
<td>(.018)</td>
</tr>
<tr>
<td>Women: genderrec. × GPG\text{fed. state}</td>
<td>-0.036*</td>
<td>-0.045*</td>
</tr>
<tr>
<td></td>
<td>(.018)</td>
<td>(.020)</td>
</tr>
<tr>
<td>Men: genderrec. × GPG\text{fed. state}</td>
<td>-.041*</td>
<td>-.037*</td>
</tr>
<tr>
<td></td>
<td>(.017)</td>
<td>(.019)</td>
</tr>
<tr>
<td>Constant</td>
<td>-5.732***</td>
<td>-5.772***</td>
</tr>
<tr>
<td></td>
<td>(.158)</td>
<td>(.159)</td>
</tr>
<tr>
<td>R²</td>
<td>.755</td>
<td>.755</td>
</tr>
<tr>
<td>Vignettes</td>
<td>7788</td>
<td>7788</td>
</tr>
<tr>
<td>Respondents</td>
<td>483</td>
<td>483</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p<.05, ** p<.01, *** p<.001 (two-tailed t-tests); rec. = recipient
The respondents of both population samples produced a just gender pay gap favoring male recipients. This gap was equal for male and female observers as stated in Hypothesis 2. The reason is, that in a population with gender inequalities it is likely that gender becomes a status characteristic and therefore relevant in the justice evaluation process. Germany is a country in which a significant gender gap in earnings and income is still persistent and therefore the German population is an example for a macro context of inequality between men and women. Although, only a part of the observers participates in the labor market these status differences are shared beliefs in wide parts of the society as they have been spreading through out the population. The fact that male and female respondents showed equal evaluation patterns is in line with findings in previous factorial survey research using a population sample (Jasso & Webster, 1997). Other factorial survey studies also found a gender gap in ratings (Jasso & Rossi, 1977; Alves & Rossi, 1978; Shepelak & Alwin, 1986; Jann, 2008).

The analysis of full-time employees resembled the findings of the complete population sample. Full-time employees directly experience inequalities in their goal-oriented daily interactions at their workplaces. There exist regional differences regarding the gender pay gap. The results show that the gender pay gap observers experience influence their evaluations regarding recipient’s gender. Observers produced higher gaps in their ratings if they lived in federal states with a high actual gender pay gap. This evaluation behavior was measured for male and female full-time working observers in both population samples. The interaction effect between recipient’s gender and the actual gender pay gap in the federal state confirms Hypothesis 3. The experienced structural inequalities between men and women affect justice attitudes towards gender. As these findings were replicated with two independent surveys it is likely that these are reliable results.

A further note is that in all three datasets there were similar effects for the other dimensions indicating consensus regarding expected rewards for inputs and abilities such as education, occupation, and age. These findings are in line with results of earlier studies (e.g., Jasso & Rossi, 1977; Alves & Rossi, 1978; Jasso, 1994; Jasso & Meyersson Milgrom, 2008; Gatskova, 2013).

This research has some limitations. First, the mechanism assumed by sta-
tatus construction theory, daily interactions, was not tested directly. It was assumed that this mechanism is likely to occur given the conditional parameters. The sub-populations embedded in different inequality structures were analyzed and the results make this explanation plausible. But future research could test the effects of daily interactions at the workplace directly as they are important for the emergence and spread of status beliefs and for justice evaluation processes. Therefore, not only data on justice perceptions but also information on the interactions of men and women in the workplace and organizations, and on the inequality and power structures would be useful. Moreover, the comparisons between different sub-populations are based on cross-sectional data. The assumption is, that contexts shape justice attitudes, meaning that students and employees change their attitudes as they come into other contexts. To test this underlying assumption, longitudinal data would be useful to separate changes in justice attitudes in regard to gender from differences between observers.

Keeping the limitations of this study in mind, the findings are important for sociological justice research as they show how inequalities influence justice evaluations of people. The combination of status and justice theories is fruitful to predict justice attitudes regarding relevant factors and their weights. The combination provides a possibility to predict under which structural conditions it is likely a categorical characteristic to become a status characteristic, and a relevant factor for the justice evaluation. Moreover, the findings might be useful for inequality research as justice attitudes reinforce actual inequalities. In all Western countries levels of pay between men and women are only slowly becoming closer (Blau & Kahn, 2003, 2006). The differences in notions of entitlement could be one reason for the slow reduction in the actual pay gap.

References


Cameron, A. C. & Trivedi, P. K. (2009). Microeconometrics using stata. College Station, TX: Stata Press.


StataCorp. (2011). *Stata: release 12. statistical software*. College Station, TX: StataCorp LP.


# Appendix

**Table 4**  
*Correlation of Vignette Dimensions of the Student Sample*

<table>
<thead>
<tr>
<th></th>
<th>Sex</th>
<th>Age</th>
<th>SIOPS</th>
<th>Voc. training</th>
<th>Earnings (ln)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>-0.006</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SIOPS</strong></td>
<td>-0.022</td>
<td>0.040</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Voc. training</strong></td>
<td>0.001</td>
<td>-0.002</td>
<td>0.202</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td><strong>Earnings (ln)</strong></td>
<td>0.028</td>
<td>0.026</td>
<td>0.472</td>
<td>0.087</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**Table 5**  
*Correlation of Vignette Dimensions of the Population Sample 1*

<table>
<thead>
<tr>
<th></th>
<th>Sex</th>
<th>Age</th>
<th>SIOPS</th>
<th>Voc. training</th>
<th>Earnings (ln)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>-0.006</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SIOPS</strong></td>
<td>-0.035</td>
<td>0.035</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Voc. training</strong></td>
<td>-0.006</td>
<td>-0.001</td>
<td>0.205</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td><strong>Earnings (ln)</strong></td>
<td>0.022</td>
<td>0.021</td>
<td>0.476</td>
<td>0.086</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**Table 6**  
*Correlation of Vignette Dimensions of the Population Sample 2*

<table>
<thead>
<tr>
<th></th>
<th>Sex</th>
<th>Age</th>
<th>SIOPS</th>
<th>Voc. training</th>
<th>Earnings (ln)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>0.007</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SIOPS</strong></td>
<td>-0.006</td>
<td>0.036</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Voc. training</strong></td>
<td>0.007</td>
<td>-0.036</td>
<td>0.250</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td><strong>Earnings (ln)</strong></td>
<td>-0.009</td>
<td>0.018</td>
<td>0.538</td>
<td>0.144</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Table 7
*Descriptives of the Justice Evaluations by Dataset*

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student sample</td>
<td>-.278</td>
<td>2.689</td>
<td>-5</td>
<td>5</td>
<td>29121</td>
</tr>
<tr>
<td>Population sample 1</td>
<td>-.326</td>
<td>3.199</td>
<td>-5</td>
<td>5</td>
<td>23213</td>
</tr>
<tr>
<td>Population sample 2</td>
<td>-4.586</td>
<td>61.111</td>
<td>-100</td>
<td>100</td>
<td>22848</td>
</tr>
</tbody>
</table>

Table 8
*Median Earnings and Gender Pay Gaps by Federal State*

<table>
<thead>
<tr>
<th>Federal state</th>
<th>Median earnings (Euro)</th>
<th>Gender pay gap (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schleswig-Holstein</td>
<td>2502</td>
<td>18</td>
</tr>
<tr>
<td>Hamburg</td>
<td>3079</td>
<td>20</td>
</tr>
<tr>
<td>Lower Saxony</td>
<td>2598</td>
<td>24</td>
</tr>
<tr>
<td>Bremen</td>
<td>2921</td>
<td>25</td>
</tr>
<tr>
<td>North Rhine-Westphalia</td>
<td>2810</td>
<td>25</td>
</tr>
<tr>
<td>Hesse</td>
<td>2959</td>
<td>23</td>
</tr>
<tr>
<td>Rhineland-Palatinate</td>
<td>2688</td>
<td>22</td>
</tr>
<tr>
<td>Baden-Württemberg</td>
<td>2941</td>
<td>28</td>
</tr>
<tr>
<td>Bavaria</td>
<td>2779</td>
<td>25</td>
</tr>
<tr>
<td>Saarland</td>
<td>2748</td>
<td>26</td>
</tr>
<tr>
<td>Berlin</td>
<td>2510</td>
<td>18</td>
</tr>
<tr>
<td>Brandenburg</td>
<td>2004</td>
<td>8</td>
</tr>
<tr>
<td>Mecklenburg-Western Pomerania</td>
<td>1907</td>
<td>2</td>
</tr>
<tr>
<td>Saxony</td>
<td>1931</td>
<td>10</td>
</tr>
<tr>
<td>Saxony-Anhalt</td>
<td>1989</td>
<td>1</td>
</tr>
<tr>
<td>Thuringia</td>
<td>1914</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2648</strong></td>
<td><strong>21</strong></td>
</tr>
</tbody>
</table>

Sources: Federal Employment Agency and Statistical Office of Rhineland-Palatinate
Previously published SFB 882 Working Papers:


Fauser, Margit / Voigtländer, Sven / Tuncer, Hidayet / Liebau, Elisabeth / Faist, Thomas / Razum, Oliver (2012): Transnationality and Social Inequalities of Migrants in Germany, SFB 882 Working Paper Series, No. 11, DFG Research Center (SFB) 882 From Heterogeneities to Inequalities, Research Project C1, Bielefeld.


Cardona, Andrés (2013): Closing the Group or the Market? The Two Sides of Weber’s Concept of Closure and Their Relevance for the Study of Intergroup Inequality, SFB 882 Working Paper Series, No. 15, DFG Research Center (SFB) 882 From Heterogeneities to Inequalities, Research Project A1, Bielefeld.


Sauer, Carsten / Valet, Peter / Liebig, Stefan (2013): The Impact of Within and Between Occupational Inequalities on People’s Justice Perceptions Towards their Own Earnings, SFB 882 Working Paper Series, No. 21, DFG Research Center (SFB) 882 From Heterogeneities to Inequalities, Research Project A6, Bielefeld.


Faist, Thomas (2014): "We are all Transnationals now": The Relevance of Transnationality for Understanding Social Inequalities, SFB 882 Working Paper Series 25, DFG Research Center (SFB) 882 From Heterogeneities to Inequalities, Research Project C1, Bielefeld.

