

Reducing Ethnic Discrimination in Rental Applications: the Development of a Training Intervention

Abel Ghekiere, Fanny D'hondt, Eva Derous, Stijn Schelfhout & Pieter-Paul Verhaeghe

Abstract

While field experiments are valuable in their ability to objectively analyse patterns of discrimination, research that focuses on reducing discrimination is systematically lacking. The present study addresses this gap. Building on both cognitive and behavioural processes that steer discriminatory decision-making and insights from diversity literature, we develop and test the effectiveness of a training intervention which is tailored to the specific context of the housing market, with 113 students in real estate. We develop our training around three theoretical mechanisms of discrimination that are measured through a specified vignette experiment in pre- and post-tests. Our results show a reduction in statistical and customer taste-based discrimination after the training, but no change in agent taste-based discrimination. The driver for this reduction in discrimination is being more selective on majority candidates rather than increasing the chances of ethnic minority candidates. Additionally, we found that training has no effect on *taste* and that most of these effects remain present on the long term.

Key words: ethnic discrimination, housing market, training intervention, vignette experiment, intersectionality

1. Introduction

Recent meta-analyses of field experiments clearly documented the persistent existence of ethnic discrimination on the housing market (Auspurg et al., 2019; Flage, 2018; Quillian et al., 2020). Hence, an ethnic majority applicant is almost twice as likely to be chosen as an ethnic minority applicant in OECD countries (Flage, 2018). Although these field experiments are valuable in their ability to objectively analyse patterns of discrimination (Gaddis, 2019), their focus has predominantly been on measuring ethnic discrimination on the housing market (Gaddis, 2018; Zschirnt, 2019). However, studies that focus on reducing rental discrimination are systematically lacking (Cui et al., 2020; Oh & Yinger, 2015).

Previous research centred largely around diversity training to reduce ethnic discrimination in educational and organizational settings (Bezrukova et al 2016; Kulik Roberson 2008; Paluch et al 2021). Multiple meta-analyses support the success of training interventions in increasing sensitivity in cultural diverse contexts (Bezrukova et al., 2016; Kalinoski et al., 2013; Mendenhall et al., 2004), but their usefulness in hiring or housing context have only limited empirical support (Deros et al., 2020). In general, the lack of research on these interventions comes mostly from the difficulty to have a good understanding of why a training is effective or not (Kulik & Roberson, 2008). This results from the lack of good instruments to measure the effectiveness of a training (Deros et al., 2020), the difficulty to measure underlying mechanisms of discrimination (Sawert, 2020) and the dominant focus on attitude-based outcomes, contrasting the multiple overview studies that show that attitudes are rather resistant to change (Bezrukova et al., 2016; Kalinoski et al., 2013; Paluck et al., 2020). We introduce an intervention that is tailored to the housing market, in which we focus on specific skills to combat discrimination rather than a change in attitudes.

Our study aims to develop a training which is tailored to the specific context of the housing market and deals with the aforementioned critiques. The training focuses on ethnic discrimination during the first stage of the rental process: the initial screening of rental applicants and inviting them to view the rental dwelling. This phase is mostly studied by correspondence studies on rental discrimination (Auspurg et al., 2019) and is crucial in combating discrimination on the housing market. We use 113 students in real estate and related subjects to participate in the intervention, of which we measure the effectivity with a pre- and post- test using a factorial survey experiment. The training was developed around three theoretical mechanisms of discrimination (taste based, customer based and statistical discrimination). This allows us to be very specific on the effects of the intervention (e.g., does our training affect statistical or taste-based discrimination). In the training, we focus on discriminatory questions and selection versus discrimination. A training group of 77 students received a training, while a control group of 36 students did not. The latter students only took the experiment in pre- and post-test without attending the training. Rather than having a measure of effectivity like a self-assessment test (e.g., *'In general, how would you rate your ability to accurately articulate a client's problem who comes from a cultural group significantly different from your own?'* an item from the MAKSS skill subscale) we make use of a factorial survey experiment which simulates real housing market situations and will be used in this study to measure effectiveness of the intervention. This survey experiment has been introduced and validated before by Ghekiere, Verhaeghe, et al. (2022).

This article is structured as follows. First, we elaborate on the theoretical learning mechanisms used in this study. Subsequently, we present the development and practicalities of the training. Later, we discuss the design and analysis of the intervention. Then, we present the findings of our analyses. In the final section, we state conclusions and reflect on limitations and future research.

2. Background

2.1. Theoretical Learning Mechanisms

The review of diversity trainings by Kulik and Robinson (2008) provides a useful understanding of the effects and measures of training interventions. The review distinguishes three broad types of diversity training outcomes within different contexts, namely training regarding cognitively diversity knowledge, training regarding diversity attitudes and training regarding skills and behaviour. Even though attitude change after diversity training receives the most attention in training literature (Kulik, Robinson, 2008), multiple reviews show that attitudes are rather resistant to change and hardly affected by training (Dalege et al., 2016; Paluck et al., 2020), and diversity training has larger effects on cognitive-based and skill-based outcomes relative to attitude-based outcomes (Bezrukova et al., 2016; Kalinoski et al., 2013).

In our study, we therefore focus on skill-based outcomes, namely, how to react to and handle discriminatory clients or situations in the first stage of the rental process. A more recent study by Verstraete and Verhaeghe (2019) sheds new light on the strategies realtors use when interacting with clients that request discrimination towards ethnic minorities and with which justifications they comply or oppose discriminatory requests of owners. For instance, almost 35% of the realtors agree to a discriminatory request, thus confirming discriminatory customer requests as an important source of discrimination (Verstraete & Verhaeghe, 2019). Hence, empowering the realtors to shape the client interaction into a collaboration without discrimination is crucial and necessary. The intervention builds up towards a phased plan on how to divert a discriminatory question into a professional cooperation without discriminating, strategies to disrupt the clients' prejudices, etc.

2.2. The Development of the Training

The desired outcome is a change in behaviour rather than a change in personal attitudes. Therefore, we derive our design from the theoretical foundations of experiential learning (D. A. Kolb, 2014). In order to effectively teach individuals, a training must follow a sequence of different steps. First, trainees have a concrete experience. Second, this experience is the basis for reflective observations. Third, these observations are being used to distil knowledge through theory and concepts (abstract conceptualization). Fourth, the knowledge can be actively tested in new situations (active experimentation). The main goal of the model is to obtain *deep learning*, in which all four steps of the learning cycle – experience, reflecting, thinking and acting - are fully integrated (A. Y. Kolb & Kolb, 2009). The Kolb model is used, in our study, as a structure for our intervention, not particularly as a model that we want to validate in practice.

Experiential learning has proven to be efficient in numerous studies. Alkan (2016) found evidence for the effectiveness of the experiential learning model on academic achievement when a treatment group, who received education through the experiential learning model, was compared with the control group that was taught with a traditional teacher- centred approach (Alkan, 2016). Similar results by Baker and Robinson (2016) show that teaching in an “experiential learning way” yielded greater practical use of the knowledge that was transferred, when compared to their direct instruction counterparts. Also, concerning studies related to racism and discrimination, Winkler (2018) found support for the efficacy of beginning with students’ concrete experiences and working toward more abstract theoretical conceptualization and application when teaching about compound concepts such as racism or discrimination.

2.2.1. Concrete experience

While the examples of a concrete experience used in the experiential learning model are endless (Morris, 2019), it is essential to first consider an intervention that fits the nature of the task

(engaging with discriminating clients) and behavioural outcomes of interest (unbiased screening and professional interaction with discriminating clients) (Derous et al., 2020). The concrete experience in this study are multiple simulations of a screening of applicants by the respondents through different clients who want to rent out their dwelling (see appendix for example). This is the first exercise of the training and immediately presents the students with challenging situations that include discriminatory questions and legal selection criteria. We assume, in line with research by Paluck et al. (2020) that this kind of experiences stimulates perspective taking, which in turn reduces prejudice and discriminatory behaviour.

2.2.2. Reflective observation

The reflective observation will take the form of an analysis of the experiment, with critical reflection of the results. We present the aggregated results, live, with a focus on what selection criteria have been used. More concretely, we discuss the limits of selection and the start of discrimination. These results will subsequently be linked to practice, through knowledge on discrimination law and real estate deontology. Previous experiences are challenged by new conceptualizations and theory of the results.

2.2.3. Abstract conceptualization

The mechanisms tested in the factorial survey experiment will consequently be used to theorize the underlying patterns and motivation behind discrimination. The students subsequently engage in a more theoretical part in which prejudice, bias and attitudes are being discussed and linked to the specificities of the housing market. This takes the form of a theoretical part in which the mechanisms are elaborated on and subsequently discussed by the students.

2.2.4. Abstract experimentation

The training ends with a simulation in which the participants have to actively take the role of both the real estate agent and the discriminating client. A phased action plan is presented

through which the participants must operate in order to interact with a discriminating client in a professional manner (see appendix). Also, we focus on the objective assessment of rental candidates in the first phase of the rental process. This includes how to legally select rental candidates instead of discriminating against certain groups, what information can you take into account when selecting a rental candidate, etc. Finally, we focus on ethnic discrimination in invitation. We teach and practice equal treatment of every candidate, disregarding irrelevant personal characteristics.

3. Participants, Design and Analysis

The training experiment took place among educational institutions in Flanders – the Northern part of Belgium. Already several studies showed persistent levels of discrimination on the Belgian housing market against ethnic minorities (Ghekiere & Verhaeghe, 2022; Martiniello & Verhaeghe, 2022; Verhaeghe & De Coninck, 2022) and against people with social benefits and low incomes (Verstraete & Moris, 2019). Almost all these studies found that private landlords tend to discriminate more than real estate agents. However, realtors still discriminate in circa one fifth of the rental advertisements, a discrimination rate of 20 percent, which allows for much room for improvement. Several studies suggest that Belgian realtors discriminate because of a combination of both customer taste-based and statistical discrimination (Ghekiere, Lippens, et al., 2022; Verstraete & Verhaeghe, 2020). Also, these studies found more discrimination in predominantly white and rich areas and – in addition - general invitations rates tend to be lower in socioeconomically deprived neighborhoods (Martiniello & Verhaeghe 2022). In sum, research indicates that there is a complex interplay of ethnic origin and social class in the exclusion of ethnic minorities on the rental housing market. Hence, if real estate agents become aware of this type of exclusionary mechanisms and learn to no longer consider

these mechanisms but focus on the relevant criteria in the preselection of candidates, they become important agents in structural societal change.

The research population for our training consists of 113 students in real estate, management or insurance across different higher educational institutions in Flanders (the northern region of Belgium). In total, all six educational institutions, who offer a program for real estate agents, agreed to the request to include our training in the curriculum. The training took place in the beginning of the second semester of the academic year of the program. A training group of 77 students was trained and compared with a control group of 36 students who only took the vignette experiment in pre- and post-test without attending the training, these differences are discussed in the result section.

These students are schooled to work as real estate agents in the context of the Belgian housing market. It is important to note that discrimination is not directly included in the student's curriculum. This training will be the first training they will receive on unequal treatment and discriminatory clients. In general, the agents act as an intermediary between rental candidates, who seek a dwelling and the owner of the property, who seeks a renter. This job consists of finding the best candidate for the owners, who, as we simulate in our experiment, could pose their own needs and requests. A preselection is mostly done by the agent before the candidates are introduced to the owner. Discrimination through the request of an owner is equally illegal as discrimination because of an individual driver and will be equally judged by the law.

The two groups in our study were either assigned a control condition or an intervention condition. This makes our experiment a quasi-experimental design, as the students were not randomly allocated to a training or control group but through their study area. However, because the two groups are still in their early years of the program, they overlap for almost all

general courses and are similar in age (the average is 20 years old). However, we did find a larger proportion of female participants in the trained group (74%) compared with the control group (39%). Additionally, we found no significant ethnic diversity (7% with a migration background) in our sample, a result that mirrors the situation of professional real estate agents in the field. Throughout the course of six months, the participants were asked to complete the 2 x 2 ethnicity (Belgian name vs North African name) and income (high income vs low income) vignette experiment in a pre-, and repeated post-test, as part of, what we presented as a longitudinal study on the measurements of biases in rental processes by students in real estate. The vignette was administered at three time points: at the start of the academic year (T1), right after the students followed the training intervention (T2) and 3 months after the training intervention (T3). Both the test and the control group completed these vignettes with the possibility to compare the two in later analysis.

Knowing that the long-term effects of diversity decay over time (Bezrukova et al., 2016), we insert a post-measure directly after the training and a repeated-post measure three months after the training. Few studies have included repeated post measures to measure long-term training effects (for an overview: (Bezrukova et al., 2016). The studies that did include long term effects, vary substantially in the definition of “long-term”. For instance, Ehrke and colleagues (2014) found positive long-term effects of diversity training tested after one month while Derous and colleagues (2021) found declining long-term effects after three months. Hence, literature needs more studies that test the effects of a training in the long-term. A better understanding of how training effects evolve over time is needed to evaluate which training elements succeed in creating a change of behaviour in the long run.

3.1. The Test of Effectiveness

In the subsequent part, we will discuss the instrument that measures the effectiveness of the training. A vignette experiment, or a factorial survey experiment, is an experimental method that is widely used to measure attitudes, beliefs or behavioural intentions. Respondents are asked to answer to a set of questions in the context of a certain scenario (vignette), describing certain situations, hypothesis or objects (Auspurg & Hinz, 2014; Düval & Hinz, 2020).

The vignette experiment introduces a simulation of a rental application, in which the respondent is confronted with a specific rental property, a client and four rental candidates. The client asks the participant to look for the right candidate and either adds *I would not like to rent out to foreigners*, or *I would only want clean and quiet renters*. The variation of the two client requests follows a between subject design and is later addressed in the analysis as the variable *discriminatory question*. Thus, every respondent received either a discriminatory or a neutral request. Following, the candidates are presented in a 2 x 2 within subject design, with ethnicity (Belgian name vs North African name) and income (high income vs low income). Concretely, four candidates, of which two with a north African name with either a low or high income and two candidates with a Belgian name with either a low or a high income, were presented to the respondent. The order in which the four fictitious candidates were shown, was randomised. The income is set so that the candidates' income is always high enough to pay for the rental price. Some additional information on every candidate was provided, like the marital status, number of adults, children and pets. This information was quasi-identical for each candidate.

Subsequently, each candidate must be rated on different items such as *this candidate will be able to pay rent in time*, or *I will invite this candidate for a viewing*. The differences in responses on items and invitation rates between ethnic minority and majority candidates (i.e., ethnic inequalities) are being used as a measure of ethnic discrimination. Secondly, by focusing on the individual drivers of each mechanism (agent taste-based, customer taste-based and statistical discrimination) we can address which one is most malleable and undergoes the

biggest change after the training. The mechanisms we test for in our study are presented in the following two paragraphs. The items are based on the study by Baert and De Pauw (2014) and are also used in Ghekiere, Verhaeghe, et al. (2022).

First, taste-based discrimination, as posited by Becker (1971), refers to the unequal treatment of an ethnic minority candidate based on a personal “distaste“ for ethnic minority candidates (Becker, 1971). It is argued to be a conscious act, as the cost of discriminating (not letting the dwelling to a good minority candidate) is taken into account by the realtor. This mechanism could be rooted in a personal preference for ethnic majority candidates (agent taste-based discrimination) or could be result of an interaction with a discriminatory client (customer taste-based discrimination). The latter is defined as discrimination through a discriminatory request of the client, being the property owner (Verstraete & Verhaeghe, 2019; Yinger, 1986).¹

Secondly, statistical discrimination (Arrow, 1973; Phelps, 1972) occurs when imperfect information about an individual leads to (falsely) ascribing group characteristics to that individual. For example, the perceived financial strength of an ethnic group could be used to assess a candidate’s financial situation (Auspurg et al., 2017). This could result in inducing higher risk for ethnic minority candidates because of imperfect information. Discrimination is not only a social problem of exclusion, it could also be detrimental for the firm’s economic situation (Pager & Shepherd, 2008). According to Becker, taste-based discrimination leads to suboptimal recruiting decisions. Unequal access to the housing market leads to, also for

¹ However, some discussion is at stake in the housing discrimination literature as to what extent customer taste based discrimination could be limited to taste based discrimination. Some scholars consider this a form of taste based discrimination (e.g. Ahmed et al. 2010; Baldini and Federici 2011; Auspurg et al. 2017), others interpret it as a form of statistical discrimination because it is more economic rational (e.g. Hanson and Hawley 2014). However, the argument for customer statistical discrimination, made by Hanson and Hawley (2014), relates to the agent’s past experiences to formulate the expected payoff for each potential home-seeker and selects a lessee by profit maximization. However, no lack of information is at stake here and a direct, discriminatory, personal request is being made by the client.

statistical discrimination, a loss of capable candidates which entail a cost of discrimination for the agent. However, competitive markets should help eliminate this type of discrimination as prejudiced employers face higher production costs.

To measure these three mechanisms, we make use of the vignette experiment. The items of this experiment are presented in **Table 1**. All items were rated on a 7-point Likert scale. Ethnic inequalities in the first two items, ‘As a real estate agent, I will enjoy working with this candidate’ and ‘I can trust this candidate’, signal agent taste-based discrimination. These items measure the personal preference towards a rental candidate, not regarding the income and other characteristics. Next, ethnic inequalities in the items: ‘My client will be happy with the candidate’ and ‘My client matches this candidate’, signal the customer taste-based mechanism. As argued in the theoretical framework, this mechanism measures the willingness to use the client’s characteristics or question to be selective in working with a rental candidate. In other words, they grasp the perceived judgement of the client towards the applicant. However, it is important to note that this only holds when a discriminatory question is asked. When the request is “to have clean and quiet renters”, it is the agent own judgement that would link ‘being clean and calm’ to an ethnic origin, not the owner. Which would measure the agent taste based mechanism and not the customer based mechanism. Finally, ethnic inequalities in the next items signal statistical discrimination, in which the agent uses generalized group characteristics to judge the applicant. The items root for perceptions about income and communication skills of the candidate. The first item: ‘The candidate will have difficulty to pay rent in time’ signals the perceived financial and punctual strength of the candidate, using a combination of personal (income is available for the participant) and group characteristics. The second item: ‘Communication with this candidate could be difficult’ signals the perceived communicative strength of the applicant.

Eventually, the question is raised whether or not the participant would invite the applicant to a viewing: ‘I will invite this candidate for a first viewing’.

The order in which the 4 applicants were shown was randomised so that every repeated test was different from the one before, to reduce order effects within the experiment (Auspurg & Jäckle, 2017). We use the combination of the two items for each mechanism, the latent constructs. These constructs are subsequently tested in a confirmatory factor analysis, to measure the internal consistency and the fit of the data to our model².

< Table 1 about here >

In the first part of our analysis, we perform a confirmatory factor analysis to check if our model fits the data. Subsequently, ANCOVA and repeated ANCOVA tests are conducted to measure the effects of the training. We use the applicant ethnicity as the within-subjects variable and training intervention as the between-subjects variable.

In the second part, we perform multilevel regression analyses to study the effect of the key variables, ethnicity, income, etc. on the different mechanisms of discrimination. Our data is hierarchical, with the results of time 1 and time 2 nested within the individual respondent, which allowed us to perform multilevel analyses on the different mechanisms measured in the factorial survey experiment. An additional robustness check is added by including a difference in difference analysis to the regression. We experienced an expected decline in respondents at the repeated-post test, 3 months after the training, compared to Time 2 (Nicholson et al., 2017). Notwithstanding the financial incentives that we offered; we could only reach 25 percent of the

² For more detailed information on the vignette experiment, we refer to the study by Ghekiere, Verhaeghe, et al. (2022).

initial trained population at Time 3. This is probably due to the absence of the students on campus because of Covid measures at the time. Subsequently, we conduct our main analysis with T1 and T2 results and include the T3 results in an additional analysis.

4. Results

The results from the confirmatory factor analysis suggest a good fit between the data and the model, with a Tucker Lewis Index (TLI) of 0.98, a Comparative Fit Index (CFI) of 0.99, and a Root Mean Square Error of Approximation (RMSEA) of 0.04. We consider values above 0.95 to be suggestive of a good fit for the two relative indices, TLI and CFI (Schreiber, 2008). For the RMSEA, a value lower than 0.06 indicates an acceptable fit (Schreiber, 2008). The factor loadings range from 0.70 to 1.18, indicating that the items fit the latent constructs (mechanisms of discrimination) very well and that these mechanisms are distinct.

4.1 Effect on mechanisms of discrimination

First, **Table 2** shows that there is ethnic discrimination when indicating whether to invite a candidate for a viewing. The unequal treatment in invitation rates between the Belgian and North African applicant is significant at Time 1 for the control group $F(1,139) = 14.74, p < .001$ and for the training group $F(1,315) = 5.90, p = .016$.³ Interestingly, where the ethnic discriminatory behaviour in invitation remains with the control group $F(1,139) = 14.62, p < .001$ at Time 2, the training group show no significant discrimination towards the ethnic minority candidate after the training intervention $F(1,315) = 3.93, p = .480$. When comparing T2 with T1, we found that the control group at Time 2 showed no decrease in discrimination on invitations when compared to Time 1, $F(1,66) = 1.182, p = .280$. For the trained group,

³ Discrimination exists for both training and control group, which implies the important notion that these two groups are rather similar in their assessment of the rental candidates at Time 1, before the training.

however, we found a significant decrease in ethnic discrimination through invitation rates, $F(1,141) = 3.680, p = .045$. This suggests that the training intervention had a significant effect on ethnic discrimination, namely ‘invitation rate’, as measured with the factorial survey experiment. The calculated eta squared, $\eta^2 = .025$, shows a small but significant change (Cohen, 1973; Yigit & Mendes, 2018).

Interestingly, the driver for these results is that the respondents of the trained group behave in a more selective way, indicating to a lesser extent to invite the candidates for a viewing, and the control group the other way around. Although there is a decrease in ethnic discrimination, we must keep in mind that this goes hand in hand with being more selective to all the candidates after the training. In other words, less candidates are generally invited, but at the same time, the priming of the majority candidate decreases.

Second, we tested for changes in agent taste-based discrimination (the personal “*distaste*” towards a certain candidate). There is a significant difference in scores for the Belgian compared to the North African candidate at Time 1 for the control group $F(1,139) = 4.82, p = .030$, which indicates that the respondent would prefer to work together with the Belgian candidate over the North African candidate. This difference increases when looking at the scores of the control group at Time 2 $F(1, 139) = 6.35, p = .013$. However, we did not find any significant agent taste-based discrimination for the trained group at Time 1: $F(1,315) = 0.27, p = .601$, nor at Time 2: $F(1,315) = 0.12, p = .734$. This result indicates that the students in real estate are not acting out of personal taste when judging a rental candidate.

Additional analysis compared T2 and T1 and showed that that there is no significant decrease in agent taste based discrimination, in other words, the training has no effects on the personal preference for certain rental candidates, not for the control group $F(1,66) = 0.066, p = .799$, nor for the trained group $F(1,154) = 1.075, p = .301$.

Third, our results show that the unequal treatment of candidates is largest in the mechanism of customer taste-based discrimination. Note that, for this mechanism, we only use respondents that received a client with a discriminatory question. As half of the respondents received no discriminatory question, we use only the results from the other 57 respondents in this analysis. At Time 1, the difference in the customer taste based mechanism (the unequal treatment of candidates because of the prejudices of the client) was significant for the control group $F(1,87) = 210.1, p < .001$ but also present for the trained group $F(1,155) = 245.7, p < .001$. The unequal treatment towards the ethnic minority candidate remains at Time 2 for the control group $F(1,87) = 135.7, p < .001$ and for the trained group $F(1, 155) = 124.166, p < .001$. When comparing T2 with T1, there is a significant decrease in ethnic discrimination for the customer taste-based discrimination mechanism for the trained $F(1,209) = 132.5, p < .001$, and for the control group $F(1,97) = 29.3, p = .001$. These effects are considered large (Cohen, 1973), with an eta squared of .388.

Fourth, statistical discrimination was only detected for the control group at T1 $F(1,139) = 6.83, p < .010$, but not for the trained group $F(1,315) = 120.06, p = .501$. This result is in line with the results of the study that introduced this Vignette experiment on the housing market (Ghekiere, Verhaeghe, et al., 2022). Ghekiere and colleagues (2022) argue that statistical discrimination is especially hard to measure with students. This mechanism tests the perceived financial and communicative strength of the candidate, a mechanism that could lead to discrimination because group characteristics that are used to assess these characteristics could be based on previous interactions with certain groups.

< Table 2 about here >

4.2 Long term effects of the training

When looking at the results from our sample at T3 in **Table 3**, we found that the findings of T3 are similar to those of T2. Most importantly, the control group continues to disadvantage ethnic minority candidate in the invitation for a viewing $F(1,59) = 3.56, p = .048$, while the trained group does not $F(1,51) = 0.65, p = .425$. In line with the results at T2, we could not find agent taste based or statistical discrimination at T3. However, there remains an unequal treatment of North African applicants in the customer taste based mechanisms at T3 for the control group $F(1,19) = 7.124, p = .016$, but not for the trained group $F(1,7) = 3.52, p = .110$. A disclaimer for these results is that these were only found in a small subsample of our initial population.⁴ While these results could not be used to generalize to the broader population, they do give us some interesting insight in the trends of our training effects.

< **Table 3 about here** >

4.3 Results of the regression analysis

We perform a multilevel linear regression to study the effect of the key variables, ethnicity, income and discriminatory question on the different mechanisms of discrimination. Doing so, we are able to deconstruct the underlying patterns of each mechanism. Allowing us to have a clear view on the effects of the intervention. Additionally, we add a difference in difference model in our regression analysis, as a robustness check. We found that the test and control group are not quite similar in their assessment of the rental candidates. Especially in invitation rates, we found significant more discrimination with the control group than the test group.

⁴ We ran additional analysis where we compare the results at T1, T2 and T3 for the same 28 respondents but found no significant results. This can most likely be addressed to the small sample size of T3 respondents.

Therefore, we use a difference in difference method, which is a quasi-experimental approach that compares the changes in outcomes over time between a population enrolled in the training (the treatment group) and a population that is not (the control group) (Lechner, 2011).

4.3.1 Dependent variables

The dependent variables in our analysis are the invitation rates and the three mechanisms of discrimination presented in **Table 1**. These are all continuous variables ranging from 1 to 7.

4.3.2 Independent variables

Ethnicity of the candidate is a binary variable that indicates the perceived ethnicity of the rental candidate, signalled through the name of the candidate. Two of the four presented candidates have a north African name, the other two have a common Belgian name. The variable is therefore equally dispersed over our data.

Income of the candidate is a binary variable that signals either a high income or a low income. Again, in the vignette, two of the four presented candidates have a high income, two have a low income.

Discriminatory question is a binary variable that takes the value of 1 if the respondent was presented with a client that asked to discriminate in their search for a candidate, it takes the value 0 if this question was not raised.

4.3.2 Results

The results from the multilevel analysis on invitation rates are presented in **Table 4**. The table is divided in four models, representing the different mechanisms and the invitation rate. Firstly, Model 1 shows that ethnicity has a significant negative impact on the invitation rate. In other words, candidates with a North African sounding name are significantly less invited to a viewing. The income has, on the contrary, a positive impact on the invitation to a viewing.

Candidates with a higher income are significantly more invited to a viewing. Additionally, there is a significant decrease in invitation rates when the client expresses a discriminatory request. But most importantly, the difference in difference analysis shows a positive, significant effect. Which implies a reduction in ethnic discrimination for the tested group after the training.

Second, in model 2, we analyse the underlying patterns of the agent taste-based mechanism. As with the invitation rates, the ethnicity of the candidate has a significant negative effect on the respondent's personal attitude towards that candidate. Income, on the other hand, shows a positive effect on personal *taste* for a candidate. In contrast with the invitation rates, we did not find any effect of the discriminatory question on the agent taste based mechanism. Interestingly, whether or not the client expresses a discriminatory request, does not affect the respondent's personal judgement of the candidates. Also, no significant decrease in discrimination through the mechanism was found with the difference in difference method.

Third, model 3 presents the results for customer taste based mechanism. The results show similar patterns as model 2. Ethnicity has a big negative effect on the perceived fit between client and customer. On the other hand, a higher income of the client has a positive effect. As with the invitation rates, we found a negative effect from discriminatory question on the mechanism. The effect of the discriminatory question was largest with regard to this mechanism. But also, the difference in difference shows a reduction in ethnic discrimination for the tested group after the training.

Finally, regarding the statistical mechanism, model 4 shows that ethnicity has a significant effect on the statistical mechanism. This implies that our respondents use the candidate's ethnicity to make claims about perceived income or communicative skills. Income has a positive effect on the statistical mechanisms, which was expected as the mechanism checks for financial strength and communication skills of the candidates. Next, there exists no significant

effect of the client's request on the statistical mechanism. This result implies that the respondents made their judgements on the candidate's financial and communicative strength disregarding the client's discriminatory request. No decrease in discrimination through this mechanism was found.

< **Table 4 about here** >

To dig deeper in the, what appeared to be, important factor: income, we added a multilevel linear regression on invitation rates for the entire sample (n=113) and the trained group after the training (n=77) for both low- and high-income candidates in **Table 5**. The results of the entire sample show that ethnic discrimination is higher for low-income candidates than for high income candidates. Also, the negative effect of a discriminatory request on the invitation rate is larger for the low-income candidate.⁵ This shows that intersectionality with income is present for both ethnic origin as the client's request. When we look at the tested group after training, we find that ethnic discrimination still exists with low-income candidates but disappears with high income candidates. No effect of the discriminatory request was found on invitation rates in this sample. These results are visualised in **Figure 1** in appendix.

< **Table 5 about here** >

⁵ Additional analysis show that the tested group, after the training, favor high income candidates when a discriminatory question is asked. Moreover, a discriminatory request not only eases the way for discrimination based on ethnic origin but also based on income. This shows the strong intersectionality of ethnicity and financial strength

5. Discussion and conclusion

While many studies have analysed the patterns and mechanisms of ethnic discrimination on the rental housing market (Flage 2018; Auspurg et al. 2019; Quillian et al. 2020), evidence-based research about how to effectively reduce rental discrimination is very limited. At the same time, there is a burgeoning literature about the potential role of diversity trainings to tackle discrimination in hiring (Kalinowski et al., 2013), but its effectivity is still unclear and it has not been applied yet to the context of the housing market. This study combines insights from both literature strands. It examines to which extent rental discrimination against ethnic minorities can be reduced by means of a specially developed diversity training – based on the experiential learning cycle (A. Y. Kolb & Kolb, 2009) – which instructs real estate agents about how to deal with discriminatory requests of property owners. Moreover, we address the profound critique about the difficulty to measure the effectivity of diversity trainings (Kulik & Roberson, 2008; Paluck & Green, 2009) by working with a pre- and two post-tests based on multi-factorial vignette surveys. This study is, to the best of our knowledge, the first to tackle rental discrimination on the housing market with an experimental design that measures effectiveness in pre and post tests.

Our results have shown that our developed diversity training is partly effective to reduce discriminatory behaviour among real estate students. Ethnic inequalities in being invited for viewing the rental dwelling are significantly reduced after the training. However, it appears that the decrease in relative differences invitation rates is especially due to a lower invitation rate for the majority candidates and not to a higher invitation rate for the minority candidate. It appears that that real estate students show stricter, but also more fair selection processes after the training. In other words, the training conveys professional values, making the trained students more aware of the legal selection criteria, which results in a decrease in the priming of the majority candidate over the minority candidate.

A different angle on these results could be that the decrease in discrimination is fuelled by the decreased in-group favouritism. As argued by Greenwald and Pettigrew (2014), a skewed positive assessment of the in-group candidate could equally result in discrimination, as discrimination often occurs as differential favouring. The interesting question arises as to whether our intervention tempered the exceeding positive assessment of the in-group candidates, or rather increased the positive assessment of ethnic minority candidates. The results seem to suggest the first. Additionally, our results show that when the assessment is executed through a positive, personal expression like “I would like to work with this candidate”, the gap between the ethnic groups is larger than when there is a direct negative expression towards the minority group. In other words, the respondents easily withhold positive emotions from the minority group, whereas they find it more difficult in expressing negative judgement towards an individual ethnic minority candidate. Future trainings should continue to focus on equal opportunities, where the selection process is identical for the ethnic minority and majority candidate. Besides, the importance of intersectionality should be considered in these interventions. We found that ethnic discrimination after the training remained present only for the low-income candidates. Moreover, we found that the discriminatory request result in not only ethnic discrimination but also discrimination based on financial strength.

While this training did not aim to change agent taste-based discrimination, we did find some interesting results regarding the mechanism. The results of the multilevel analysis reveals that the personal judgement of the candidates was not affected by the client’s discriminatory request. This shows the independence of personal tastes within the selection process. Not only did we not find any training effects on the personal mechanisms, but we also found that this mechanism is not really context dependent. In line with the literature, our results show that, to get to the agent’s *tastes*, more is needed than a single training intervention, subsequently this result questions the thoroughness of training interventions in general.

However, the training did work especially well in reducing customer taste-based discrimination. Since previous research has shown that many realtors answer positively on discriminatory requests of clients (Verstraete & Verhaeghe, 2019), this is a promising finding. In addition, some of these effects seem to work on the long term too. Our results show that the discrimination rates based on invitation and the customer based mechanism remain present with the control group but are repeatedly not found within the trained group after the training. Nevertheless, we must acknowledge that a decrease in customer taste-based discrimination does not always lead to a decrease in actual discrimination in access to housing. In most cases, it is the property owner who makes the final decision in selecting a rental candidate. Although, a decrease in customer taste-based discrimination comes along with more equal treatment in the first phase of the rental process. But we cannot identify yet whether this change will also lead to an actual increase in access to rental dwellings in the later phases of the rental process.

These conclusions should be considered within the confines of the following limitations. First, our study was conducted among students and not real estate agents. There is no guarantee that in a real, professional situation, a real estate agent would react in the same way as our student sample did. Certainly when we consider the financial incentive to follow the clients' discriminatory requests, we could expect that the training probably has a positive impact, but it might be overstated here. However, we stress that, while the magnitude of the decline in discrimination might be an overstatement, it remains crucial to gain insight into the mechanisms that are affected by this type of training. Also, our results could be confounded because the condition was not randomly assigned to the participant group. This brings up the question of the external validity of the study. However, Falk et al. (2013) show that students results in rating job candidates are nearly identical to those of professionals and that the responses are given in a less socially desirable manner. Additionally, the meta-analysis of Kalinoski and colleagues (2013) showed that student samples may provide similar or even

more conservative estimates of training intervention effects than employees in an organizational setting.

Second, our measure of behaviour is one that is simulated. We use a vignette study and not real behaviour. Future research should explore the possibilities of measuring the effects of a diversity training on actual behaviour with the use of a large-scale field experiments like correspondence tests or mystery calls. Additionally, while the CFA model seems to fit the data well, we could question the construct validity of the items used in the experiment. Especially regarding the distinction between the taste and statistical mechanisms. While our factor analysis shows a clear distinction between the two, there remains a substantial uncertainty as to what these items measure exactly. For example, regarding the taste based mechanism, one could argue that the concept of ‘trust’ in the item “I can trust this candidate”, could refer to a personality trait as well as a group characteristic. Though our factor analysis shows that this item loads primarily on the taste based mechanism, future research could analyse the validity of these items in more detail.

Third, while the sample size of our pre and first post test were considered quite large for a diversity training study (Paluck et al., 2020), we did encounter sample attrition. This resulted in a small sample of students who participated in last of the three test moments. Hence, our results of the long-term effects should be read with nuance. Also, because we used a within subject design for the candidates’ characteristics, some social desirability could be at stake. While only few studies have experimentally compared *within-subjects* with *between-subjects designs* (see Auspurg et al., 2015; Walzenbach, 2019), a full between-subjects design would still be valuable for future research. For now, we argue that enlarging the knowledge on how to behave in specific situations can act as a precursor of self-efficacy. Schelfhout et al. (2022) showed that individuals with a broader cultural intelligence indeed showed higher levels of self-efficacy. We are cautiously optimistic that the effects found are thus an indication that the

participants have a more broad cultural intelligence on how to avoid discrimination in future behavior.

The results of this study might not only prompt scientific implications but could also induce societal repercussions. Training students, future realtors, on this topic is crucial to enhance change in long term. While effects may decline over time, the awareness of legalities and biases when selecting rental candidates may shape future practitioners towards a fairer process. This could counter high discrimination rates, which soar up to 25 percent in most Flemish cities (Ghekiere & Verhaeghe, 2022; Verhaeghe & Ghekiere, 2020). However, as previous research has shown (Kalinoski et al., 2013), a one-time training is seldom enough to produce change in systems. Practitioners should have the opportunity to get information when needed; as Deros et al. (2021) proposes, nudging and e-learning modules could be used to do so.

References

- Alkan, F. (2016). Experiential learning: Its effects on achievement and scientific process skills. *Journal of Turkish Science Education, 13*(2), 15-26.
- Arrow, K. (1973). The theory of discrimination. *Discrimination in labor markets, 3*(10), 3-33.
- Auspurg, K., & Hinz, T. (2014). *Factorial survey experiments* (Vol. 175): Sage Publications.
- Auspurg, K., Hinz, T., & Schmid, L. (2017). Contexts and conditions of ethnic discrimination: Evidence from a field experiment in a German housing market. *Journal of Housing Economics, 35*, 26-36.
- Auspurg, K., & Jäckle, A. (2017). First equals most important? Order effects in vignette-based measurement. *Sociological Methods & Research, 46*(3), 490-539.
- Auspurg, K., Schneck, A., & Hinz, T. (2019). Closed doors everywhere? A meta-analysis of field experiments on ethnic discrimination in rental housing markets. *Journal of Ethnic and Migration Studies, 45*(1), 95-114.
- Baert, S., & De Pauw, A.-S. (2014). Is ethnic discrimination due to distaste or statistics? *Economics Letters, 125*(2), 270-273.
- Baker, M. A., & Robinson, J. S. (2016). The Effects of Kolb's Experiential Learning Model on Successful Intelligence in Secondary Agriculture Students. *Journal of Agricultural Education, 57*(3), 129-144.
- Becker, G. S. (1971). *The Economics of Discrimination*. University of Chicago Press *Economics Books*.
- Bezrukova, K., Spell, C. S., Perry, J. L., & Jehn, K. A. (2016). A meta-analytical integration of over 40 years of research on diversity training evaluation. *Psychological bulletin, 142*(11), 1227.
- Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and psychological measurement, 33*(1), 107-112.

- Cui, R., Li, J., & Zhang, D. J. (2020). Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on Airbnb. *Management Science*, 66(3), 1071-1094.
- Dalege, J., Borsboom, D., Van Harreveld, F., Van den Berg, H., Conner, M., & Van der Maas, H. L. (2016). Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model. *Psychological review*, 123(1), 2.
- Derous, E., Nguyen, H.-H. D., & Ryan, A. M. (2020). Reducing ethnic discrimination in resume-screening: a test of two training interventions. *European Journal of Work and Organizational Psychology*, 1-15.
- Düval, S., & Hinz, T. (2020). Different order, different results? The effects of dimension order in factorial survey experiments. *Field Methods*, 32(1), 23-37.
- Falk, A., Meier, S., & Zehnder, C. (2013). Do lab experiments misrepresent social preferences? The case of self-selected student samples. *Journal of the European Economic Association*, 11(4), 839-852.
- Flage, A. (2018). Ethnic and gender discrimination in the rental housing market: Evidence from a meta-analysis of correspondence tests, 2006–2017. *Journal of Housing Economics*, 41, 251-273.
- Gaddis, S. M. (2018). An Introduction to Audit Studies in the Social Sciences. *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, 14, 1.
- Gaddis, S. M. (2019). Understanding the “how” and “why” aspects of racial-ethnic discrimination: A multimethod approach to audit studies. *Sociology of Race and Ethnicity*, 5(4), 443-455.
- Ghekiere, A., Lippens, L., Baert, S., & Verhaeghe, P.-P. (2022). Ethnic discrimination on paper: uncovering realtors’ willingness to discriminate with mystery mails. *Applied Economics Letters*, 1-4.

- Ghekiere, A., & Verhaeghe, P.-P. (2022). How does ethnic discrimination on the housing market differ across neighborhoods and real estate agencies? *Journal of Housing Economics*, 55, 101820.
- Ghekiere, A., Verhaeghe, P.-P., Baert, S., Deros, E., & Schelfhout, S. (2022). Introducing a vignette experiment to study mechanisms of ethnic discrimination on the housing market. *PloS one*, 17(10), e0276698.
- Greenwald, A. G., & Pettigrew, T. F. (2014). With malice toward none and charity for some: ingroup favoritism enables discrimination. *American Psychologist*, 69(7), 669.
- Kalinoski, Z. T., Steele-Johnson, D., Peyton, E. J., Leas, K. A., Steinke, J., & Bowling, N. A. (2013). A meta-analytic evaluation of diversity training outcomes. *Journal of Organizational Behavior*, 34(8), 1076-1104.
- Kolb, A. Y., & Kolb, D. A. (2009). Experiential learning theory: A dynamic, holistic approach to management learning, education and development. *The SAGE handbook of management learning, education and development*, 42, 68.
- Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*: FT press.
- Kulik, C. T., & Roberson, L. (2008). Common goals and golden opportunities: Evaluations of diversity education in academic and organizational settings. *Academy of Management Learning & Education*, 7(3), 309-331.
- Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends® in Econometrics*, 4(3), 165-224.
- Martiniello, B., & Verhaeghe, P.-P. (2022). Does the neighbourhood of the dwelling and the real estate agency matter? Geographical differences in ethnic discrimination on the rental housing market. *Urban Studies*, 00420980221086502.

- Morris, T. H. (2019). Experiential learning—a systematic review and revision of Kolb’s model. *Interactive Learning Environments*, 1-14.
- Nicholson, J. S., Deboeck, P. R., & Howard, W. (2017). Attrition in developmental psychology: A review of modern missing data reporting and practices. *International Journal of Behavioral Development*, 41(1), 143-153.
- Oh, S. J., & Yinger, J. (2015). What have we learned from paired testing in housing markets? *Cityscape*, 17(3), 15-60.
- Pager, D., & Shepherd, H. (2008). The sociology of discrimination: Racial discrimination in employment, housing, credit, and consumer markets. *Annu. Rev. Sociol.*, 34, 181-209.
- Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology*, 60, 339-367.
- Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2020). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, 72.
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review*, 62(4), 659-661.
- Quillian, L., Lee, J. J., & Honoré, B. (2020). Racial Discrimination in the US Housing and Mortgage Lending Markets: A Quantitative Review of Trends, 1976–2016. *Race and Social Problems*, 12(1), 13-28.
- Sawert, T. (2020). Understanding the mechanisms of ethnic discrimination: A field experiment on discrimination against Turks, Syrians and Americans in the Berlin shared housing market. *Journal of Ethnic and Migration Studies*, 46(19), 3937-3954.
- Schelfhout, S., Vandecasteele, R., De Maesschalck, S., D’hondt, F., Willems, S., & Derous, E. (2022). Intercultural Competence Predicts Intercultural Effectiveness: Test of an Integrative Framework. *International journal of environmental research and public health*, 19(8), 4490.

- Verhaeghe, P.-P., & De Coninck, D. (2022). Rental discrimination, perceived threat and public attitudes towards immigration and refugees. *Ethnic and Racial Studies*, 45(7), 1371-1393.
- Verhaeghe, P.-P., & Ghekiere, A. (2020). Is de etnische discriminatie op de Gentse woningmarkt structureel gedaald. *Vakgroep Sociologie, Vrije Universiteit Brussel*.
- Verstraete, J., & Moris, M. (2019). Action–reaction. Survival strategies of tenants and landlords in the private rental sector in Belgium. *Housing Studies*, 34(4), 588-608.
- Verstraete, J., & Verhaeghe, P.-P. (2019). Ethnic discrimination upon request? Real estate agents' strategies for discriminatory questions of clients. *Journal of Housing and the Built Environment*, 1-19.
- Verstraete, J., & Verhaeghe, P.-P. (2020). Ethnic discrimination upon request? Real estate agents' strategies for discriminatory questions of clients. *Journal of Housing and the Built Environment*, 35(3), 703-721.
- Winkler, E. N. (2018). Racism as a threshold concept: examining learning in a 'diversity requirement' course. *Race Ethnicity and Education*, 21(6), 808-826.
- Yigit, S., & Mendes, M. (2018). Which effect size measure is appropriate for one-way and two-way ANOVA models? A Monte Carlo simulation study. *Revstat Statistical Journal*, 16(3), 295-313.
- Yinger, J. (1986). Measuring racial discrimination with fair housing audits: Caught in the act. *The American Economic Review*, 881-893.
- Zschirnt, E. (2019). Research ethics in correspondence testing: an update. *Research ethics*, 15(2), 1-21.