Automatically Activated Stereotypes and Differential Treatment Against the Obese in Hiring

Dan-Olof Rooth

October 2008

IZA

Automatically Activated Stereotypes and Differential Treatment Against the Obese in Hiring

Dan-Olof Rooth
Kalmar University
and IZA

ABSTRACT

Automatically Activated Stereotypes and
Differential Treatment Against the Obese in Hiring[*]

This study provides empirical support for automatically activated associations inducing unequal treatment against the obese among recruiters in a real-life hiring situation. A field experiment on differential treatment against obese job applicants in hiring is combined with a measure of employers' automatic/implicit performance stereotype toward obese relative to normal weight using the implicit association test. We find a strong and statistically significant obesity difference in the correlation between the automatic stereotype of obese as being less productive and the callback rate for an interview. This suggests that automatic processes may exert a significant impact on employers' hiring decisions, offering new insights into labor market discrimination.

Corresponding author:

Dan-Olof Rooth
Department of Economics
Kalmar University College
391 82 Kalmar
Sweden
E-mail: Dan-Olof.Rooth@hik.se

1. Introduction

A large body of literature has analyzed the correlation between body weight and labor market outcomes.[1] While there are several explanations as to why such a correlation could exist this study focuses exclusively on the "discrimination" link, asking *why* obese job applicants are treated differently in the hiring process by employers.[2] Although much attention has been directed at the unequal treatment of various groups in society, e.g., ethnic minorities, women and obese, little is still known about what causes differential treatment in the hiring situation. Through the lens of "taste-based" and "statistical" discrimination theories economists have assumed that this differential treatment can be traced back to people's consciously mediated perceptions of the "discriminated" target group (see Altonji and Blank, 1999).

However, considerable evidence now supports the view that a person's stereotypes (and attitudes) in fact often also operate in an automatic, less conscious, mode (see Stanley, Phelps, and Banaji, 2008).[3] According to dual-process-theory[4] people can process information using simple, low effort, readily available decision rules, referred to as *automatic*, *spontaneous*, *implicit* or *impulsive* processes, as well as through the ordinary type of processing involving the active, effortful scrutiny of relevant information, referred to as *controlled*, *deliberate*, *explicit* and *reflective* processes, see Ranganath, Smith, and Nosek (2006).[5]

While people's explicit stereotypes have for long been captured directly by self-reported reflections in questionnaires, the measurement of their automatic/implicit counterparts have a much shorter history. The latter are often achieved at by respondents taking the implicit association test (henceforth IAT), a computerized response latency

---

[1] See for instance Cawley (2004) and the references therein.

[2] Since discriminating between people on the grounds of weight is lawful in Sweden, we henceforth use the term *differential treatment* instead of the more value-laden *discrimination*. The fact that labor laws in Sweden make it very difficult to fire people (the so called "last in-first out" principle), the existence of "solidarity" wages and a narrow wage distribution make it especially likely to find gender and ethnic discrimination in hiring, rather than in wages. This is also expected to be the case for differential treatment due to weight.

[3] De Houwer, Thomas, and Baeyens (2001) show that people frequently form such associations without conscious intent.

[4] See Chaiken and Trope (1999) for an overview of dual process models.

[5] We interchangeably use the terms *automatic* and *implicit*, and *controlled* and *explicit*, throughout the article.

measure designed to tap individual differences in associations between concepts and attributes, see Greenwald, McGhee, and Schwartz (1998). The IAT thus indirectly measures an individuals' automatic/implicit stereotype by the assessment of his/hers relative differences in strengths of associations.

The IAT has been found to predict prejudiced behavior in various contexts (see Lane et al., 2007; Greenwald et al., in press).[6] For example, as regards automatic stereotyping of Blacks in the U.S., IAT has been found to predict the degree to which people make negative judgments of ambiguous behaviors performed by a Black target (Rudman and Lee, 2002). Furthermore, Green et al. (2007) reported that implicit anti-Black attitudes and stereotypes predicted physicians' medical decisions in that they were more reluctant to prescribe medications to African-American patients diagnosed with the same condition as White American patients.

The present study examines the extent to which recruiters possess more negative automatic associations of individual productivity toward obese individuals relative to normal weight individuals and whether it predicts differential treatment in the context of deciding which job candidates to call for a job interview.[7] The extent to which implicit discriminatory stereotypes (or attitudes) are in fact related to discriminatory behavior in the hiring situation *as such* has so far only been studied in laboratory settings focusing mainly on ethnic discrimination.[8] However, recently Rooth (2007), using an identical design to the present study, was the first study to provide empirical support for automatically activated stereotypes of Arab-Muslim men relative to Swedish men inducing discriminatory behavior among recruiters in a real-life hiring situation. He finds a strong and statistically significant ethnic difference in the correlation between the

---

[6] In fact, the IAT has been found to predict criterion measures ranging from, for instance, interracial friendliness and impression formation to anxious and shy behaviors, consumer choices, and voting, see e.g., Bargh, Chen, and Burrows (1996), Uhlman and Cohen (2007), and Phelps (2006).

[7] The design of the IAT is such that it can only assess relative difference in association. Thus, when we say that the obese are associated with low productivity, this is not in an absolute sense. Rather, it means that the obese are more easily associated with attributes reflecting low productivity *compared with* normal weight individuals. While the relative nature of the IAT can in some cases be a limitation, it is actually beneficial in the case of research on differential treatment in the labor market, since it is the relative difference in how the groups are perceived and treated that is of interest.

[8] For instance, Ziegert and Hanges (2005) found that an implicit racist attitude, which interacted with a climate for racial bias, predicts discriminatory behavior, while an explicit counterpart does not.

automatic association of Arab-Muslim men as being less productive and the callback rate for an interview.

The research design combine two research projects, one in economics and the other in social psychology. In Rooth (in press) a field experiment specifically designed to test for differential treatment in recruitment using a weight manipulation of facial photographs attached to job applications.[9] Correspondence testing in this particular situation implies that the researcher sends two equal applications to advertised job openings with the only difference being the photo attached to the application: one obese and the other of normal weight. The degree of differential treatment is quantified by the difference between the two groups in the number of callbacks for a job interview. Related to this is the critique in Heckman (1998) that audit studies do not measure the marginal impact of discrimination since the sectors, or jobs, being investigated are not the ones applied to by the *minority*. This critique is less likely to hold in the case of the obese and the sectors used in this study. First, overweight individuals today make up a large share of the Swedish population. The Swedish National Institute of Public Health (2006) reported that 44 percent of Swedish males and 28 percent of females aged 18-84 can be classified as overweight, and another 11 percent of both groups can be classified as obese. Second, register information from the Survey of Living Conditions (Statistics Sweden) show that these numbers correspond also to the sectors used in this study (see Rooth, in press).

In Agerström, Carlsson, and Rooth (2007) an IAT was developed that measures automatic work performance stereotypes toward obese relative to normal weight, which then was administered to a sample of the recruiters in Rooth (in press). These IAT results were then used to predict decisions made in the hiring process.

The results show that recruiters find it significantly easier to associate obese with incompetence, laziness, and inefficiency than with productivity, ambition, and efficiency. In fact, 90 percent have at least a slight, and 82 percent at least a moderate automatic performance stereotype. Also, 56 percent of the recruiters explicitly state that they would prefer hiring a normal weight job applicant over an obese one. It is found that employers'

---

[9] In this context the Swedish labor market is an ideal market to analyze since the use of photographs in job applications is quite common. Interviews with recruiters at some Swedish companies show that every fifth application they receive have a personal photo attached to them.

job interview callback rates for obese job applicants are negatively correlated in a statistically significant way with their automatic performance stereotype, but not with their explicit hiring preference or performance stereotype. Hence, the result shows that an automatically activated stereotype is a significant determinant in the hiring process, while explicitly stated attitudes and stereotypes seem to be controlled and not acted upon.

Other than being the second study that shows the predictive power of automatically activated stereotypes in a real-life hiring situation this study has several advantages compared to the study by Rooth (2007). First, the results apply for a group for which no protective discrimination legislation exists. Hence, employers probably to a lesser extent than the case of ethnic minorities feel a need to hide their true explicit stereotypes, which makes a comparison of the predictive power of explicit and implicit measures of stereotypes more trustworthy. Second, both studies have a low participation rate. In this study we make an attempt to analyse if the participation is selective by randomly allocating economic rewards of various sizes if participating. We find that it is not.

The remainder of this article is outlined as follows. Section 2 presents evidence of the explicit measure not being affected by measurement bias and that the IAT is not just a better measure of a persons attitudes and stereotypes compared to explicit measures, but rather measures related but different constructs. Section 3 presents the correspondence test and its results in more detail, while section 4 presents the implicit association test. Section 5 presents results where the main question concerns the extent to which the job interview callback rate can be predicted from the implicit performance stereotypes (the IAT score), explicit attitudes and stereotypes, respectively, while section 6 concludes.

2. Implicit and explicit measures and processes – the same, related or distinct?
The IAT is of special value in acts related to stereotyping due to its high resistance to faking. In fact, participants typically cannot successfully fake their IAT score even if explicitly instructed to by researchers (see Kim, 2003; Steffens, 2004). However, this can not be expected to be the case for self-reported measures. In the context of prejudice people might be unwilling to report an evaluative response that comes to mind because (i) they do not want others to know about it, or (ii) the feeling is unwanted in the sense that it is not endorsed or accepted as one's evaluation. However, while the former creates

measurement error in a self-reported measure the latter does not, since it might be genuine in the sense that he/she believes the response is unfair and attempts to replace it with a more positive explicit response. As argued above, this might be especially relevant when analyzing ethnic discrimination being prohibited by law, but less so in the context of differential treatment against the obese. Interestingly, more than half of the recruiters in this study explicitly report being prejudiced. Hence, there is less scope for expecting them not to have disclosed their true prejudiced attitudes/stereotypes.

Several meta studies have shown a weak correlation between IAT scores and explicit measures, see e.g., Greenwald et al., in press. In the present study the share having at least moderate negative automatic associations of obese and the share having explicit preferences for hiring a normal weight job applicant are both high making one to expect that they measure the same underlying construct. However, the correlation coefficient is no larger than 0.08. Hence, either the two measures tap related but distinct constructs or there is measurement issues involved.

Intuitively one might expect automatic activations of general associations or attitudes, as assessed by the IAT, and self-reported prejudice to be rooted in the same experiences and thus be directly related. Research in social psychology, and more recently in neuroscience, typically does not support this expectation. Instead, the evidence support that automatic (as measured by the IAT) and explicitly measured attitudes are related, but distinct, constructs (for reviews see Fazio and Olson, 2003; Nosek and Smyth, 2007). First, the IAT score and corresponding explicit measures of attitudes have a low correlation even in domains were they are expected to be highly correlated, i.e., in a situation in which there is little reason to believe that participants would have been concerned with controlling their explicit attitudes, see Karpinski and Hilton (2001).[10] Second, research in social psychology using subliminal priming methods have shown that automatically triggered attitudes can influence behaviour without our awareness (Fazio, Dunton, and Williams, 1995; Greenwald and Banaji, 1995). Third, and perhaps most convincing, is the identification of distinct neural systems for the expression and regulation of automatic/implicit attitudes and the involvement of those

---

[10] In their experiment respondents' attitude towards candy bars and apples were measured both indirectly (by IAT) and directly. As in this study, while both measures had similar means, they had a low correlation.

systems in implicit and explicit processes (Stanley, Phelps, and Banaji, 2008). This neurological evidence states that implicit measures reflect more automatic than controlled reactions to social groups (Cunningham et al., 2004; Phelps and Banaji, 2005). For example, a study conducted by Phelps et al. (2000) found that participants' performance on an IAT correlated positively with activity in the amygdala, a structure in the brain that is recruited when fear learning is induced indirectly (Phelps and LeDoux, 2005), and hence, can trigger behaviors before people are consciously aware of this, indicating that those participants who displayed more implicit negative attitudes toward unfamiliar black faces also evidenced heightened levels of amygdala activity.[11] However, there were no association between the explicit measures and the activation of the amygdala.

Could automatically activated stereotypes then affect recruiter's hiring decisions of minorities? Given that automatic associations are thought to be more influential in situations where there is a high degree of automaticity, and where deliberate and more time consuming cognitive processes are mitigated (Greenwald et al., in press), it may be that they play a more important role when, for example, recruiters have to do a quick initial screening among a number of different job applications, as well as when cognitive load, fatigue and "gut feelings" enter the equation. In (a small number of) interviews with recruiters we often fell upon the comment that they instantaneously get a "feeling" of the application. Interestingly, Ranganath, Smith, and Nosek (2006) find that the IAT score correlate with an explicitly reported, but spontaneous, "gut feeling" of gay people relative to straight people, but not with classic self-reported attitudes of the same targets. Ideally, we would like to measure (i) under what conditions a negative automatic response is activated when reading the job application, and (ii) if and when more controlled processes override such initial automatic associations. However, in the context of a real hiring situation, being outside the black box of the hiring decision, we can only observe the reduced form of such a process, i.e., whether automatically activated associations correlates with discriminatory behavior.

---

[11] Olsson, Nearing, and Phelps (2007) show that fear acquired indirectly through social observation, with no personal experience of the aversive event, engages similar neural mechanisms as fear conditioning, suggesting that indirectly attained fears may be as powerful as fears originating from direct experiences.

3. Correspondence testing – differential treatment in hiring

Rooth (in press) tested for differential treatment in hiring based on the job applicants weight (obesity) using the correspondence testing method (see Riach and Rich, 2002). When "correspondence testing" in conjunction with obesity, the researcher typically sends two job applications containing applicants with identical skills to advertised job openings with the only difference being group membership, which is signalled by a face photograph of the applicant – one randomly assigned a normalweight and the other an obese face. The extent of differential treatment is quantified by the relative callback rates for interview between the two groups.[12]

For the present study, the field experimental data from Rooth (in press) served as the behavioral criterion variable on which the employers' IAT scores were regressed. The behavioral data were collected between January and August 2006 by sending applications to job openings pertaining to seven different occupations in the Stockholm and Gothenburg labor market areas. The selected occupations were both skilled and semi/unskilled. The selected occupations were computer professionals, business sales assistants, preschool teachers, accountants, nurses, restaurant workers, and shop sales assistants.[13]

The applications used in the experiment had to be realistic and yet not refer to any real persons. Also, because the competition from other applicants was considerable, the testers had to be well qualified, that is, to have skills comparable to above average applicants for that job. Thus, applicants had identical human capital within occupations and were of the same age (varied between 25-30 years old across occupations), had the same amount work experience (varied between two to four years across occupations), and had obtained their education in the same type of school, but at different locations. Further, the application consisted of a quite general biography on the first page and a

---

[12] A detailed description of the process in constructing the obesity signal in the job applications is found in Appendix B.
[13] To make satisfactory progress in the collection of cases, it was necessary for the demand for labor to be relatively high in the chosen occupations. In addition, the skill requirement and the degree of customer contact needed to vary across occupations. Hence, the selected occupations were both skilled and semi/unskilled and included relatively high as well as low contact with customers.

detailed curriculum vitae, including education and work experience, on the second page.[14]

In total Rooth (in press) replied to 985 job ads posted at the home-page of the Swedish Employment Agency, see Table A1 in appendix. The first column tells us that the two applications were sent for a total of 985 different job openings, 527 for the men and 458 for the women. Because correspondence testing focuses only on the first step of the hiring process -- being called for an interview -- and thus neglects the second step of actually getting the job, there are four possible interview outcomes: neither invited, both invited, or only the normal weight or obese individual being invited for an interview. In 317/249 cases, neither male/female applicant was invited. In the remaining 210/209 cases, at least one of the applicants was invited to interview. Both applicants were invited in 130/118 cases, while only the normal weight applicant was invited in 56/63 cases, and only the obese applicant in 24/28 cases. From this information we can calculate separately the callback rate of normal weight and obese applicants, respectively, as well as the the difference in callback rates for men and women. The callback rate is found to be significantly lower for obese applicants. On average, it is 6 and 8 percentage points lower than for normal weight male and female applicants, respectively.

3.1 Participants in the IAT experiment

The IAT data was collected, among the recruiters in the experiment discussed in the previous section, between October 2006 and March 2007. Hence, there is a time delay between the employers' completion of the IAT test and their hiring process decision. However, Lane et al (2007) show that an individual's IAT scores are strongly correlated over time (r>0.5).

The first step was to locate the firm's recruiter and, more exactly, the person who was responsible for selecting candidates to invite for an interview for exactly the job we applied for in the field experiment. This was achieved by calling the firm at the number included in the job ad. This was an extremely laborious task, reaching approximately only

---

[14] It should be pointed out that written job applications being sent by e-mail are very common in Sweden and we have used job application exemplars available at the web page of the Swedish Employment Agency as templates, see Appendix B for an example of a job application in the computer specialist occupation.

four recruiters a day. When reached they were informed that we were interested in this specific job vacancy and that we had followed its progress. However, we did not reveal that we had sent fictitious applications. We then informed them that our project intended to study the recruitment process in general and that their participation included first taking a "sorting test" on the computer and then answering a short, also computer-based, questionnaire, without revealing that we intended to measure their stereotypes of obese individuals.

In an attempt to study the selectiveness of the participation decision we randomly altered a participation reward of 0 and 300 SEK (100 SEK is approximately 11 euros or 14 dollars) among the recruiters. Everyone was told that the task required a devotion to our study for approximately ten minutes during which they were not allowed to be disturbed. They were allowed to choose whether to participate on a computer at work or at home. Still, recruiters were found to be a very busy group and were hard to convince to participate in the study. A total of 679 employers/recruiters were invited to participate, see Table 1.[15] Thirty-seven percent, or 252 employers, were not located, stated that the recruiting decision was not his/hers or not his/hers exclusively or were simply not interested in participating. Hence, 427 employers agreed to participate. However, in the end only 153 employers, or twenty-two percent, finally completed the IAT and the questionnaire. If focusing on each group/category (rows 1 through 4) separately we find some variation in the share, varying from twenty-one to thirty-three percent, of the firms that have participated in taking the IAT, except for the category *neither invited* for which only thirteen percent participated.[16] Also, since a much smaller fraction of the group *neither invited* were invited to participate in the study we need to use weights in the analysis of the data when analysing the marginal effect of implicit stereotypes on the difference in the callback rate. These weights are calculated according to the within group occupational distribution of the field experiment and therefore also correct for participation differences across groups and occupations.

---

[15] Attempts were made to contact all employers in the first three categories (*only obese applicant*, *only normalweight applicant* and *both*), while a random sample of 293 employers were selected for contacting in the last group (*neither invited*).

[16] This could be expected since this group also was found to be especially difficult to get in contact with.

3.2 Selective participation – using economic incentives to increase participation

To analyze whether selective participation exists and affects the implicit and explicit measures we offered approximately half of the sample (49 percent) in group 1 through 3 an economic incentive of 300 SEK for participating in the interview. Initially we were not as interested in group 4 and hence, only 14 percent in that group were given the incentive.[17] The first two columns in Table 2 show the proportion of participants who received the incentive within each group, while columns 4 and 5 give the same distribution for those 153 recruiters who finally participated in the interview. When the economic incentive was offered, the proportion of participants who decided to participate increased in groups 1-3, while that was not the case for group 4. However, a Fischer test shows that participation is significantly altered by the economic incentive for group 1 and 3, but not for group 2 and 4.[18]

The next step is to analyze whether the *means* (or more correctly the *counts*) of the implicit and explicit measure differ between those receiving an economic incentive and those who did not within groups. Again, we employ the Fisher test for the explicit measures, where we merge a "2" response into the "1" response category since too few responded with a "2", see Table 3. For neither measure do we find any indications of that participation is selective.[19] For the implicit measure, e.g., the IAT score, the Wilcoxon test is used and no evidence is found of an association between participation incentive and the IAT score, see Table 3. Even if the analysis of group 2 and 4 is not as convincing as we would have wished, the lack of selective participation in groups 1 and 3 makes us believe that selectivity is not an issue for groups 2 and 4.

---

[17] At the initial stages of the project we only made comparisons among groups 1 through 3. The reason for neglecting the fourth group was that we were unsure whether "neither invited" actually was an act of equal treatment. Other researchers have since then convinced us of including also the group "neither invited" in the analysis.

[18] The p-values for group 1 through 4 are <0.001, 0.3, <0.001 and 1, respectively. Given the small cell sizes we have to employ nonparametric tests such as the Fisher test (and later the Wilcoxon test, see below). These calculations are available upon request.

[19] However, there is a tendency that the performance stereotype differs between those recruiters who received an incentive and those who did not at the ten percent significance level.

To conclude, using a zero or 300 SEK incentive, we find that participation in the study increases significantly when having received a greater economic reward for participation, while the IAT scores and responses to the explicit measures are not affected. Hence, it seems as if those individuals whose participation increases after an incentive has been offered do not differ from those whose participation does not increase with respect to the endorsement of explicit/implicit attitudes/stereotypes.

The reason that participation is not selective might not be that surprising since non-participation seems to be unrelated to the employer/recruiter knowing what we intended to study. We failed to locate, or reach, approximately twenty-five percent of the recruiters (we stopped trying to contact them if not reached within one month's time) and another twenty-five percent directly stated, without us having informed them about the project, that they did not have time to participate. Hence, these groups were probably unaware of our purpose. What about those who agreed to participate but never did? It is clear that no one started taking the IAT (which is needed in order to realize what the test is about) and then decided to withdraw.[20] Instead, almost half of the recruiters not participating stated that they experienced problems with starting the program for firm Internet security reasons.[21] Hence, since they did not start taking the IAT, they were also unaware of our intentions.

3.3 Explicitly stated hiring preferences and productivity stereotypes

The explicit measures used in the experiment were developed based on the measures used in Nosek, Greenwald, and Banaji (2005), and included the *hiring preferences* task, and the *performance stereotype rating* task. The former is expected to capture recruiters' explicit attitudes, while the latter their explicit performance stereotype, of obese compared to normal weight individuals.[22] These questions appeared on the screen after the IAT had been completed.

---

[20] Any such withdrawal will show in the returning files of the IAT.

[21] Many companies have Internet firewalls that do not allow for plug-ins as required by the web-based IAT.

[22] The choice of explicit measures is a difficult one. We decided to use these questions since they are very direct and could not be misunderstood. If using more subtle questions, as for example situation based ones, then we would move toward questions being somewhat implicit in character.

The *Hiring preference rating* asked the participants to choose which groups they prefer when hiring people. Rather than being directly related to the IAT, hiring preferences were supposed to measure actual explicit preferences in hiring, which could be related to differential treatment. Employers had to choose one of five alternatives (e.g. "When hiring staff I strongly prefer obese to normal weight") ranging from strong preference for obese relative to normal weight (coded as -2), to a strong preference for normal weight relative to obese (coded as +2), with a neutral intermediate alternative (coded as 0).

The *performance stereotype rating* task was constructed in the same manner as the *hiring preference* task, thus consisting of five alternatives (e.g. "Normal weight perform much better at work than obese") with a neutral intermediate.

Table A2 in the appendix shows that half of the employers explicitly state that they prefer hiring (56%) a normal weight person over an obese person, while a clear majority (75 percent) state that there are no productivity differences between the two. Hence, these answers indicate the existence of a quite strong explicit negative hiring preference toward the obese, but less of a negative explicit productivity stereotype toward the same group. Since the share reporting negative hiring preferences toward the obese is quite high we expect measurement error in this variable to be less of an issue when interpreting the empirical results in section 5.

4. The IAT – measuring implicit stereotypes

The Implicit Association Test (IAT) has been developed within social psychology to measure automatic/implicit attitudes and stereotypes towards various groups. It is a computer based test designed to specifically measure individual differences in relative associations between two concepts. It was first introduced by Greenwald, McGhee, and Schwartz (1998) and has since become a widely used measure in order to capture individual's implicit attitudes and stereotypes toward social groups.[23]

The specific IAT test used in the present research was the Obese performance stereotype IAT found in Agerström, Carlsson and Rooth (2007). In this computer based

---

[23] See http://implicit.harvard.edu for a test version of the IAT.

version participants first classify/sort, as fast as possible, obese and normal weight faces appearing in the middle of the screen according to the (target) category to which they belong, "Obese" or "Normal weight". In the next part of the IAT the participants now instead classify words that are found being associated with high and low work productivity. These include words such as "lazy", "slow", "efficient" and "hard-working". In the third step faces and high/low work productivity words appear at random. The intuitive idea is that it will be easier, and hence, go faster, to classify names and words that are compatible than those that are incompatible. The IAT measures every latency in response to the presented stimulus. For example, when two categories are "easily" associated in terms of their nominal features (Obese + low work productivity and Normal weight + high work productivity) the participant classifies the stimuli much faster and with fewer errors than when they are not associated. A total of sixty stimuli are presented for the compatible and incompatible part, respectively. The difference in response latencies, or rather a recalculation of this difference gives the association strength, e.g. the IAT effect/score, between the compatible and incompatible parts.

Greenwald, Nosek, and Banaji (2003) have given bounds for the power of the IAT score, where below 0.15 is non-existent, 0.15-0.35 slight, 0.35-0.60 moderate, and an IAT-score over 0.60 is a strong effect.[24] The IAT scores of the 153 recruiters participating in this study show that a very clear majority associate words signalling negative productivity, such as "lazy" and "slow", with belonging to the obese category. In fact, as many as ninety percent have an IAT score above 0.15 (see Figure 1 using the the non-standardised IAT score).

5. Results

Differential treatment in the hiring process could possibly be a result of both consciously controlled, as well as automatically activated, negative stereotyping of the obese applicant group. This section attempts to determine which form of prejudice, the consciously controlled or the automatically activated, that is the most important determinant in the hiring situation. The empirical analysis starts by investigating the

---

[24] These bounds closely correspond to a conservative version of the bounds of the more familiar Cohen's d.

correlation between the implicit stereotype measure, i.e., the standardised IAT score, and the standardised explicit hiring preference and productivity stereotype measures, see Table 4.[25] We find a very weak positive correlation, ranging between 0.04 and 0.08, between the IAT score and the explicit statements. Hence, these weak correlations indicate that the obese performance stereotype IAT measures something that is far from related to the explicit measures. The correlation between the two explicit measures is around 0.5.

*** Table 4 about here***

The next step is to analyze to what extent the implicit and explicit measures correlate with behavior, e.g., the employers' decisions regarding whom to call for an interview. In doing so we analyze the obesity differences in the probability of a callback for interview using probit regressions (reporting marginal effects) and how this difference varies when introducing the implicit and/or explicit discrimination measure(s) into the regression.[26] The data include 153 observations for normal weight and obese applications, respectively, which are stacked together into the 306 applications being analyzed. All estimations are clustered on the level of the firm and weighted according to scheme discussed in section 3.1.[27]

When we regress the callback dummy on only the obesity indicator variable we find an almost identical result as in Rooth (in press) using the full sample of 1,970 observations, indicating that the weighting scheme is correct. Applications signalling an obese job applicant have seven percentage points lower probability of being called for interview compared to applications signalling a normal weight job applicant.

---

[25] Table A3 gives the descriptives of the raw data being used in the regression analysis.

[26] Since the explanatory variables are truly exogenous (due to the experimental design) we can rely on the probit model to return consistent estimates. In fact, in a comparison between the estimates using the probit model (the marginal effects), the linear probability model, and a fixed effects dito we receive very similar estimates for the first two, and (as expected) exactly the same estimates for the last two, models. Also the standard errors are almost identical across the models and, hence, there is no efficiency loss of using the probit model.

[27] For comparison the results using the unweighted data are found in Table A4. Since these data downplay the numerator, that is, there are too few observations in the group *neither invited*, the results become even stronger than when using the weighted data.

To investigate which, if any, of the implicit and explicit attitude and stereotype variables are correlated with the differential treatment in callbacks, we regress the callback dummy on the explicit and implicit attitude/stereotype measures and the interaction of those measures with the obesity dummy, see Model A through F in Table 5. The extent to which the included measures are associated with differential treatment in callbacks is captured by the interaction effect, reported as a marginal effect. Model D through F also control for occupation fixed effects and the interaction of these with the obesity dummy.[28] One could expect that the inclusion of occupation fixed effects would weaken the association strength of the implicit/explicit stereotype measures and the callback rate since stereotypes might be more important in certain occupations.

*** Table 5 about here***

The estimate of the implicit stereotype measure, and its interaction with the obesity dummy, implies that recruiters with an implicit negative stereotype appear to treat applications of obese more negatively, see Model A. When comparing recruiters with a one standard deviation difference in the implicit stereotype score the probability of a callback for interview is statistically insignificant for normal weight applicants, but eight percentage points lower and statistically significant for obese job applicants. The economic interpretation of this estimate is that the probability to invite obese job applicants is, on average, eight percentage points lower for recruiters that have at least a moderate implicit low-performance stereotype toward obese as compared to recruiters who have no such stereotypes.

This result is very robust, and is basically unchanged, when the explicit attitude and stereotype measures are included into the model, see Model C. The results are also unchanged when occupation fixed effects are included, see Models D and F, which

---

[28] We also have a number of other variables at our disposal as the BMI (weight and length) of the recruiter and his or hers gender and age. However, BMI had to little variation with only three percent of the recruiters being obese, while gender and age (interacted with the obesity dummy) had no statistically significant impact and were left out of the final analysis.

implies that the association between the implicit stereotype measure and the differential treatment in callbacks for interview is the same within as between occupations.[29]

In the second column, Model B, only the explicit hiring preference and productivity stereotype measures are included. The point estimates of the explicit measures are small, indicating that they are economically unimportant, and they are also estimated with a low precision. In fact, none of the explicit measures are statistically significant in Table 5.[30] This result is basically unchanged when we include the implicit stereotype measure (Model C) and include occupation fixed effects (Model E and F).

6. Conclusion

This study is the second to examine the relation among automatically activated stereotypes, explicit attitudes/stereotypes, and differential treatment (behavior) in a real hiring situation, but the first to do so for obese job applicants. We find strong and consistent negative correlations between the IAT score and the probability that the firm/recruiter invited the obese applicant for an interview. The results show that the probability to invite obese job applicants decreases by eight percentage points when the recruiter has at least a moderate negative implicit stereotype toward the obese. On the contrary, no such statistically significant correlations are found for the explicit measures and the probability to invite obese job applicants for a job interview.

An important issue is whether these results imply the emergence of an additional discrimination theory not yet proved to be important in the field of economic research or if it is just the "old" theories working but in a new disguise? Our answer must be a cautious yes and no to the first and second statement, respectively. A "Yes" because, even if there exists an(other) explicit measure that can explain discriminatory behaviour when hiring, we have provided evidence for an additional explanation to such behaviour in that an implicit measure, e.g., automatically activated stereotypes toward the obese, predict the probability that employers will invite an obese applicant less often for a job interview. And "No", because we, and others before us, have provided evidence of that

---

[29] We have also elaborated with non-linearities including a cubic of the IAT measure being interacted with the obesity dummy, but non-linearities in the IAT-effect was not found.

[30] This does not change if we include the explicit variables one by one or if we use a dummy variable were a "2" response is merged with the "1" response category.

explicit and implicit attitudes/stereotypes are in fact different ones. Especially, recent advances in neuroscience show that explicit and implicit measures of attitudes/stereotypes are identified in distinct neural systems.

Since it could be assumed that the decisions related to the hiring process are far from always under cognitive control but rather automatic in nature when, for instance, gut-feelings are activated, and stress enters the equation, this makes an even more convincing case for using the IAT. The central idea behind this is that a test that measures automatic associations should be a good predictor of behavior that takes place in situations where deliberate and more time consuming cognitive processes are mitigated. When the IAT is a particularly useful predictor in hiring situations is a topic for future research to find out.

In conclusion, our study highlights that much can be gained by combining research methods in social psychology with applied research in economics. Our research demonstrates the utility of the IAT when it comes to predicting ecologically valid behavior with "real-world" relevance for the labor market and the economy. Elucidating the role played by automatic/implicit processes in the hiring process offers new insights into labor market discrimination. Understanding all "forces" underlying discrimination is crucial in order to eventually be able to create a fair labor market.

References:

Agerström, J., Carlsson, R., and Rooth, D. (2007) "Ethnicity and Obesity – Evidence of Implicit Work Performance Stereotypes in Sweden", IFAU WP 2007:20, IFAU, Uppsala, Sweden.

Altonji, J., and Blank, R. (1999) Race and Gender in the Labor Market. In Ashenfelter O, Card D (eds) Handbook of Labor Economics, 3143–3259. Elsevier. Amsterdam.

Åslund, O. and Rooth, D. (2005) "Shifts in Attitudes and Labor Market Discrimination: Swedish Experiences after 9-11", *Journal of Population Economics* 18 (4), 602-629.

Bargh J., Chen, A., and Burrows, L. (1996) "Automaticity of Social Behaviour: Direct Effects of Trait Construct and Stereotype Activation of Action", *Journal of Personality and Social Psychology*, 1, 1-40.

Bertrand, M., Chugh, D., and Mullainathan, S. (2005) "New Approaches to Discrimination: Implicit Discrimination", *American Economic Review*, 95 (2), 94-98.

Bertrand M., and Mullainathan, S. (2004) "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination", *American Economic Review*, 94, 991-1013.

Carlsson, M., and Rooth, D. (2007) "Evidence of Ethnic Discrimination in the Swedish Labor Market Using Experimental Data", *Labor Economics*, 14, 716-729.

Chaiken, S., and Trope, Y. (1999) Dual-process theories in social psychology, New York: Guilford Press.

Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., and Banaji, M. R. (2004) "Separable neural components in the processing of Black and White Faces", *Psychological Science*, 15, 806-813.

Devine, P. (1989) "Stereotypes and Prejudice: Their Automatic and Controlled Components", *Journal of Personality and Social Psychology*, 56(1), 5–18.

De Houwer, J., Thomas, S., and Baeyens, F. (2001) "Associative Learning of Likes and Dislikes: A review of 25 Years of Research on Human Evaluative Conditioning". *Psychological Bulletin*, 127, 853–869.

Eurobarometer (2006) Public Opinion in the European Union. Report from the European Commission, Brussels.

European Social Survey (2003) Majorities' Attitudes Towards Minorities: Key Findings from the Eurobarometer and the European Social Survey. Report from the EUMC, Vienna.

Fazio, R.H., Dunton, J.R., and Williams, C.J. (1995) "Variability in Automatic Activation of an Unobtrusive Measure of Racial Attitudes: A Bona Fide Pipeline?", *Journal of Personality and Social Psychology*, 69, 1013-1027.

Fazio, R., and Olson, M. (2003) "Implicit Measures in Social Cognition Research: Their Meaning and Use", *Annual Review of Psychology*, 54, 297-327.

Green, A., Carney, D. R., Pallin, D., Iezzoni, L., and Banaji, M. (2007) "Implicit Bias among Physicians and its Prediction of Thrombolysis Decisions for Black and White Patients", *Journal of General Internal Medicine*, 22(9), 1231-1238.

Greenwald, A.G., and Banaji, M.R. (1995) "Implicit Social Cognition:Attitudes, Self-Esteem, and Stereotypes", *Psychological Review*, 102, 4-27.

Greenwald, A., McGhee, D., and Schwartz, J. (1998) "Measuring individual differences in implicit cognition: The Implicit Association Test", *Journal of Personality and Social Psychology*, 74, 1464–1480.

Greenwald, A., Nosek, B., and Banaji, M. (2003) "Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm", *Journal of Personality and Social Psychology*, 85, 197-216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E., and Banaji, M. R. (in press) "Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity", *Journal of Personality and Social Psychology*.

Hamermesh, D., and Biddle, J. (1994) "Beauty and the Labor Market", *American Economic Review*, 84 (5), 1174-1194.

Heckman, J. (1998) "Detecting discrimination", *Journal of Economic Perspectives*, 12, 101-116.

Heckman, J., and Rubinstein, Y. (2001) "Importance of Noncognitive Skills: Lessons from the GED Testing Program", *American Economic Review*, 91 (2), 145-149.

Karpinski, A., and Hilton, J.L. (2001) "Attitudes and the Implicit Association Test", *Journal of Personality and Social Psychology*, 81, 774–788.

Kim, D. (2003) "Voluntary Controllability of the Implicit Association Test (IAT)", *Social Psychology Quarterly*, 66, 83–96.

Lane, K., Banaji, M., Nosek, B., and Greenwald, A (2007) "Understanding and Using the Implicit Association Test: IV. What We Know (So Far) About the Method", in B. Wittenbrink and Schwarz, N. (Eds), Implicit Measures of Attitudes: Procedures and controversies, 59-102. New York: Guilford Press.

Lange, A. (2000) Diskriminering, integration och etniska relationer, The Integration Board (Integrationsverket), Norrköping.

McConnell, A. R., and Leibold, J. M. (2001) "Relations Among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Racial Attitudes", *Journal of Experimental social Psychology*, 37, 435-442.

Nosek, B. (2005) "Moderators of the Relationship Between Implicit and Explicit Evaluation", *Journal of Experimental Psychology*, 134, 565-584.

Nosek, B., Greenwald, A., and Banaji, M. (2007) "The Implicit Association Test at Age 7: A Methodological and Conceptual Review", in J. Bargh (Ed.), Automatic processes in social thinking and behavior, 265-292. Psychology Press.

Nosek B., Greenwald, A., and Banaji, M. (2005) "Understanding and Using the Implicit Association Test: II. Method Variables and Construct Validity", *Personality and Social Psychology Bulletin*, 31 (2), 166–180.

Nosek, B., and Smyth, F. (2007) "A Multitrait-Multimethod Validation of the Implicit Association Test: Implicit and Explicit Attitudes are Related But Distinct Constructs", *Experimental Psychology*, 54, 14-29.

Olsson, A., Nearing, K.I., and Phelps, E.A. (2007) "Learning Fears by Observing Others: the Neural Systems of Social Fear Transmission", *Social Cognitive and Affective Neuroscience Advance Access,* 2, 3-11.

Pew (2007) "World Publics Welcome Global Trade – But Not Immigration: 47-Nation Pew Global Attitudes Survey", Report from the Pew Global Attitudes Project, Washington.

Phelps, E.A. (2006) "Emotion and Cognition: Insights from Studies of the Human Amygdala", *Annual Review of Psychology*, 24, 27-53.

Phelps, E.A., and Banaji, M.R. (2005) "Animal models of human attitudes: Integrations across behavioral, cognitive, and social neuroscience", In Cacioppo, J.T., Visser, P.S., and Pickett, C.L. (Eds.) *Social Neuroscience: People Thinking about Thinking People* (229-243). Cambridge, MA : MIT Press.

Phelps, E.A., and LeDoux, J.E. (2005) "Contributions of the Amygdala to Emotion Processing: From Animal Models to Human Behavior", *Neuron*, 48, 175-187.

Phelps, E., K., O'Conner, W., Cunningham, E., Funayama, J., Gatenby, J., Gore, and Banaji, M. (2000) "Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activation", *Journal of Cognitive Neuroscience*, 12 (5), 729-738.

Price, J., and Wolfers, J. (2007) "Racial Discrimination Among NBA referees", IZA DP#2863, IZA, Bonn.

Ranganath, K.A., Smith, C.T., and Nosek, B.A. (2006) "Distinguishing Automatic and Controlled Components of Attitudes from Direct and Indirect Measurement Methods", *Journal of Experimental Social Psychology*, 44, 386-396.

Riach, P. A., and Rich, J. (2002), "Field experiments of Discrimination in the Market Place", *The Economic Journal*, 112, F480-F518.

Rooth, D-O., and Ekberg, J. (2003) "Unemployment and Earnings for Second Generation Immigrants in Sweden. Ethnic Background and Parent Composition", *Journal of Population Economics*, 16 (4), 787–814.

Rudman, L. A., and Lee, M. R. (2002). "Implicit and Explicit Consequences of Exposure to Violent and Misogynous Rap Music", *Group Processes and Intergroup Relations*, 5, 133-150.

Stanley, D, Phelps, E.A, and Banaji, M.R. (2008) "The Neural Basis of Implicit Attitudes", Current Directions in Psychological Science, 17(2), 164-170.

Steffens, M.C. (2004) "Is the Implicit Association Test Immune to Faking?", *Experimental Psychology*, 51, 165–179.

Uhlmann, E. L., and Cohen, G. (2007) "I Think, Therefore It's True: Effects of Self-Perceived Objectivity on Hiring Discrimination", *Organizational Behavior and Decision Processes*, 104, 207-223.

Ziegert, J., and Hanges, P. (2005) "Employment Discrimination: The Role of Implicit Attitudes, Motivation, and a Climate for Racial Bias", *Journal of Applied Psychology*, 90 (3), 553-562.

Figures:

Figure 1. The distribution of the non-standardized IAT score for implicit stereotypes against obese relative to normal weighted men and women.

Mean = 0,7219
Std. Dev. = 0,37682
N = 153
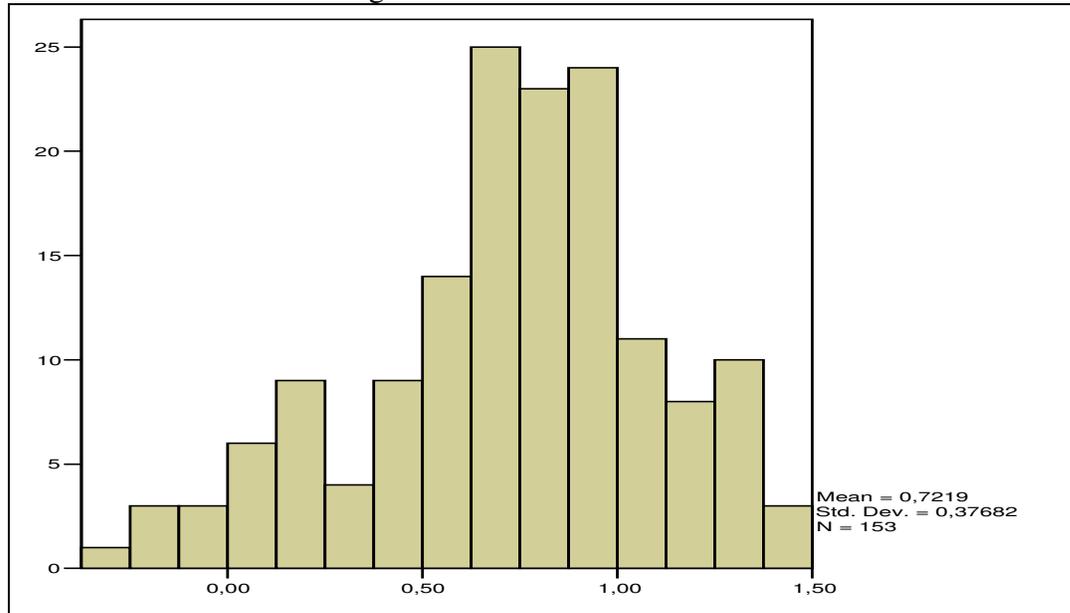
Tables:

Table 1. Participation in the two experiments.

Called for interview Different outcomes	Correspondence testing	IAT: Invited	Participated (share %)
Only normal weight applicant	119	103	35 (34)
Only obese applicant	52	48	11 (23)
Both	248	229	68 (30)
Neither invited	566	293	39 (13)
Total	985	673	153 (23)

*Note***:** From the field experiment we included all recruiters/companies from the original data (985) that were found in Statistics Sweden's firm register. The reason why all firms were not invited to participate is because they have not been located in the register (due for example to close downs). Since only a fraction (52%) of the group "neither invited" were invited to participate in taking the IAT we need to use weights in the analysis. The weights are group and occupation specific. For instance, a specific weight has been calculated for recruiters who belonged to the group "Only obese applicant" and the occupation "computer specialist". Hence, these weights also correct for participation differences across groups and occupations.

Table 2. Number of recruiters being offered an incentive of 300 SEK to participate in the experiment and the number of participating recruiters receiving and not receiving the 300 SEK incentive.

	Offered a 300 SEK incentive			Participating recruiters		
	300 SEK incentive		Total	300 SEK incentive		Total
	no	yes		no	yes	
1 Only normal weight applicant	51	52	103	5 (10)	30 (58)	35
2. Only obese applicant	20	28	48	3 (15)	8 (29)	11
3. Both	123	106	229	17 (14)	51 (48)	68
4. Neither invited	251	42	293	34 (14)	5 (12)	39
Total	445	228	673	59 (13)	94 (41)	153

Note: In parentheses is the particiaption rate.

Table 3. Tests of the association (p-values) between financial incentives to participate and explicit and implicit measures (IAT score).

	Explicit measures		**Implicit measure**
	Hiring preference	Performance stereotype	
1 Only normal weight applicant	1	0.64	0.47
2. Only obese applicant	0.49	0.24	0.38
3. Both	0.78	0.09	0.84
4. Neither invited	1	1	0.95
Test used	Fisher	Fisher	Wilcoxon

Table 4. Correlation matrix showing the relationship between the IAT score and the explicit measures. 153 cases.

Variables	Weighted data			Unweighted data		
	1	2	3	1	2	3
(1) IAT score	1	0.08*	0.04	1	0.02	0.03
(2) Hiring preference	0.08*	1	0.45*	0.02	1	0.45*
(3) Performance stereotype	0.04	0.45*	1	0.03	0.45*	1

Note: * p-value<0.05.

Table 5. The correlation between the callback rate for interview and the implicit and explicit attitude and stereotype measures. Percentage points. Weighted data.

	Model:					
	A	B	C	D	E	F
Obese (job applicant)	-0.077	-0.029	0.112	-	-	-
	(0.088)	(0.048)	(0.094)			
Implicit measures:						
IAT	-0.045	-	-0.043	-0.044	-	-0.044
	(0.041)		(0.042)	(0.045)		(0.045)
IAT*Obese	-0.079**	-	-0.077**	-0.084**	-	-0.083**
	(0.038)		(0.038)	(0.041)		(0.041)
Explicit measures:						
Hiring preference	-	-0.068	-0.065	-	-0.017	-0.014
		(0.048)	(0.047)		(0.051)	(0.050)
Hiring preference*Obese	-	-0.025	-0.023	-	-0.018	-0.017
		(0.032)	(0.032)		(0.038)	(0.037)
Performance stereotype	-	0.049	0.049	-	0.042	0.041
		(0.051)	(0.051)		(0.052)	(0.052)
Performance stereotype* Obese	-	-0.035	-0.035	-	-0.033	-0.034
		(0.029)	(0.031)		(0.034)	(0.036)
Occupation fixed effects	No	No	No	Yes	Yes	Yes
Occupation fixed effects *Obese	No	No	No	Yes	Yes	Yes
No of cases	306	306	306	306	306	306

Notes: (***), (**) and (*) indicate the one, five and ten percent level of significance, respectively. Each column shows the estimates from a regression of the callback rate on the set of variables given in the rows of the first column. For instance, the Model A being estimated is Prob (Callback=1) = a + b*Obese + c*IAT + d*[IAT*Obese]. However, the estimates in the table are marginal effects reported after the dprobit command in STATA 9. Weights are used in all models. In columns D through G the estimate of the Obese indicator variable is discarded since it now corresponds to only one of the included occupations.

Appendix A:

Table A1. Aggregated results for the correspondence testing. Men and women separately.

| | Jobs No. (1) | Neither Invited No. (2) | At least one invited No. (3) | Equal Treatment No. (4) | Only normal weighted invited No. (5) | Only obese invited No. (6) | Callback rates | | Relative callback rate (8/7) (9) | Difference in callbacks (8-7) (10) |
							Normal Weighted (7)	Obese (8)		
Men:										
Computer professionals	36	23	13	6	5	2	0.31	0.22	0.73	-0.08
Business sales assistants	80	37	43	28	13	2	0.51	0.38	0.73	-0.11***
Preschool teachers	75	27	48	35	9	4	0.59	0.52	0.89	-0.06
Accountants	53	47	6	1	3	2	0.08	0.06	0.75	-0.05
Nurses	79	24	55	42	6	7	0.61	0.62	1.02	0.01
Restaurant workers	93	66	27	9	13	5	0.24	0.15	0.64	-0.10**
Shop sales assistants	111	93	18	9	7	2	0.14	0.10	0.69	-0.07*
Men total:	527	317	210	130	56	24	0.35	0.29	0.83	-0.06***
Women:										
Computer professionals	37	24	13	6	2	5	0.22	0.30	1.37	0.09
Business sales assistants	113	57	56	27	18	11	0.40	0.34	0.84	-0.06
Preschool teachers	109	42	67	46	15	6	0.56	0.48	0.85	-0.07**
Accountants	50	34	16	7	8	1	0.30	0.16	0.53	-0.15**
Nurses	45	20	25	19	2	4	0.47	0.51	1.10	0.04
Restaurant workers	71	42	29	11	17	1	0.39	0.17	0.43	-0.21***
Shop sales assistants	33	30	3	2	1	0	0.09	0.06	0.67	-0.08
Women total:	458	249	209	118	63	28	0.40	0.32	0.81	-0.08***

Notes: The null hypothesis for the statistic in the 10th column is (7) = (8). It is calculated using the *signtest* command in Stata 9.0. *, **, and *** denote the ten, five and one percent significance level, respectively. *Source*: Table from Rooth (in press).

Table A2. Explicit measures and their discrete counterparts. 153 cases.

		Values		
Explicit measure	<0	0	1	2
Hiring preference:	-	65	76	12
Performance stereotype:	-	115	33	5

Note: Not weighted.

Table A3. Implicit and explicit variable means by group.

		Implicit measure	Explicit measure	
Called for interview	No. of	IAT	Hiring	Performance
Different outcomes	obs		preference	stereotype
1. Only normal weight applicant	35	2.15 (0.93)	1.11 (1.09)	0.69 (1.11)
2. Only obese applicant	11	0.93 (0.95)	1.17 (1.04)	0.67 (0.93)
3. Both	68	1.79 (0.99)	0.90 (0.94)	0.44 (0.85)
4. Neither invited	39	2.14 (0.89)	1.24 (1.00)	0.47 (1.01)
Total	153	1.98 (1.00)	1.12 (1.00)	0.51 (1.00)

Note: All variables are presented as standardized means of their continuous/discrete version.

Table A4. The correlation between the callback rate for interview and the implicit and explicit attitude and stereotype measures. Percentage points. Unweighted data.

	Model:					
	A	B	C	D	E	F
Obese (job applicant)	0.090	-0.159	0.092	-	-	-
	(0.105)	(0.064)	(0.113)			
Implicit measures:						
IAT	0.006	-	0.007	0.001	-	0.001
	(0.041)		(0.042)	(0.044)		(0.044)
IAT*Obese	-0.130***	-	-0.131***	-0.141***	-	-0.141***
	(0.047)		(0.048)	(0.053)		(0.053)
Explicit measures:						
Hiring preference	-	-0.076*	-0.076*	-	-0.025	-0.025
		(0.046)	(0.046)		(0.048)	(0.047)
Hiring preference*Obese	-	0.018	0.015	-	0.006	0.002
		(0.045)	(0.045)		(0.051)	(0.051)
Performance stereotype	-	0.036	0.036	-	0.029	0.029
		(0.049)	(0.049)		(0.049)	(0.049)
Performance stereotype* Obese	-	-0.038	-0.032	-	-0.030	-0.024
		(0.049)	(0.050)		(0.054)	(0.055)
Occupation fixed effects	No	No	No	Yes	Yes	Yes
Occupation fixed effects *Obese	No	No	No	Yes	Yes	Yes
No of cases	306	306	306	306	306	306

Notes: (***), (**) and (*) indicate the one, five and ten percent level of significance, respectively. Each column shows the estimates from a regression of the callback rate on the set of variables given in the rows of the first column and corresponds to the models being estimated in Table 5 using weights. The estimates are marginal effects reported after the dprobit command in STATA 9.

Appendix B: Design of job applications

Signalling obesity through photos

We signal obesity by attaching a portrait photograph to the job application. In this section, we describe the procedure by which the photos were chosen. First, approximately 100 photos of young men and women (approximately 20-30 years old) who were of normal weight but varied attractiveness were collected from an Internet photo site (www.photosearch.com). From this pool of photos, approximately 50 were chosen by a group of five evaluators (three male and two female researchers). The criteria was to choose sets of at least five photos of individuals with similar looks where each set varied according to attractiveness, from being perceived as unattractive to very attractive, and neglecting the far end category of very unattractive applicants. A total of 25 male and 23 female photos were selected. These photos then were distributed randomly and eventually graded on attractiveness (on a 9-point scale) and age by 150 students at Kalmar University.[31] In order to control for differences in average grading across students, the deviance of an individual student's average grading of the photos was calculated for each of the 48 photos. We (the evaluators) then used this information, together with "eye-balling" to choose photo pairs that were valued approximately the same on attractiveness. Our aim was to choose pairs as identical as possible with regards to such features as clothing, hair color, hair length, smile, and facial shape. The image of the (somewhat) more attractive individual of the two then was sent to a photo firm, www.mikeelliottfineart.com, for manipulation into being obese. The background of the photos also was changed to be the same within pairs. This strategy thus minimizes the

---

[31] In practice, five students evaluated the 48 photos in the same order, then another five students in another order, and so forth. Hence, 30 different orders of the photos were used.

probability that photo characteristics other than appearance through the weight manipulation are driving the result.[32]

Finally, seven pairs of photos (four male and three female) were used in the study (see Table B1 in this appendix). The obvious drawback with using only a few photos is that the results could be specific to just these photo pairs and not generalized to a greater population.[33] By using photo pairs that vary according to attractiveness this is acknowledged to some extent in the experimental design. Still, there are only a small number of photo pairs being used.

---

[32] One could be worried that employers observe that the photos have been manipulated. However, when distributing the manipulated photos to 87 students for evaluation (see below), none replied that the photos looked odd or manipulated.

[33] At first glance there are alternative strategies available. However, since the personal attribute one would like to measure needs to be strongly signaled when sending job applications, our choice was to focus on a clear obesity signal instead of a continuum of weights. Hence, we believed the treatment strategy, turning one of two equally attractive individuals into obese, to be a more successful and transparent one.

Table B1. Photos sent to employers.

Note: The four male and three female faces to the left have been manipulated into being obese. The photo pairs sent to the same employer are for men A-H, B-E, C-G and D-F, while the photo pairs for women are A-D, B-F and C-E.

Application exemplars (translated into English): Computer specialists

<div align="center">

Application A (in Verdana 10)

</div>

Hi,

 My name is Karl Johansson and I am 30 years old. Previously I worked as a system designer at Telenor
 AB between 1998 and 2005 in an environment based on win2000/SQL Server. I then participated in three
 different projects and my work contained development, maintenance and every day problem solving.
 Development work was done in ASP, C++ and Visual Basic and we used the development platform .Net
 and MS SQL. In addition I have experience in HTML, XML, J2EE and JavaScript.

I enjoy working with development and problem solving. And I now hope that I will develop further at your
company. To my personal characteristics one could add that I find it easy to work both on my own and in
a group. I am a dynamic person that likes challenges. I really like my occupation which I think is mirrored
in the work I do. I have a degree in computer engineering. I graduated with good grades from Stockholm
university.

In my spare time I like to read and listen to music. Me and my wife also like to socialize with our friends
during weekends.

I look forward to being invited for interview and I will then bring my good certificates and diplomas.

Best regards

Karl Johansson

<div align="center">

Application B (in Times new roman 12)

</div>

Hi,

First of all I would like to introduce myself. My name is Erik Nilsson, and I am 29 years old and live in Stockholm.

I previously worked at EssNet AB for about 7 years and my work tasks included designing, implementing as well as testing different
financial applications. The programming was mainly done in a Java J2EE environment. A lot of the development was also done in
C/C++ and Visual Basic. Sometimes I also participated in projects which were more web based and linked to data bases. This work
gave me experience in ASP, JSP and MS SQL among other things.

To my skills one could also add a great work joy and that I do not want to finish up work only half done. I work independently and am
good at taking own incentives. At my previous job I also learned how to cooperate and to listen to others and together create great
results.

As regards my educational background I have an engineering degree in technical physics with a focus on computer science at Uppsala
University.

Since I am very social person a great deal of my spare time is spent together with my friends and my wife. Traveling is also a great
interest of mine. I hope I will get the opportunity to come for an interview and then get the chance of telling you more about myself
and to show my good grades and credentials.

Best regards Erik