

The relationship between teacher perceptions of pupil attractiveness and academic ability.

Kirstine Hansen

Department of Social Science, UCL Institute of Education,
University College London.

Abstract: There is an established literature that suggests teacher perceptions of pupils affect how they interact with them, how they teach them and how they rate their ability and behaviour. Evidence also indicates that a teacher's perception of a child is often based on ascriptive characteristics such as gender, ethnicity and socio-economic background independent of a child's ability. This paper builds on the literature by examining the relationship between teacher perceptions of a child's ability and behaviour and their perceptions of a particular ascriptive characteristic of those children - their attractiveness. Using data from the National Child Development Study (NCDS) results show that not only do teachers rate the academic ability of pupils they perceive to be attractive more highly than less attractive students both in terms of their performance across different areas of learning (general knowledge, numbers, books and oral ability) and whether they show any outstanding potential but they are also more likely to over rate and less likely to under rate their ability than other students. This is true even after controlling for a wide range of other factors related to the child, their family, their teacher and their school that could influence the relationship, many of which are shown to be independently related to teacher's ratings.

Key words: teacher perceptions, beauty, child outcomes.

Acknowledgements: I would like to thank Lorraine Dearden, Jane Elliott, John Jerrim, Stephen Machin, Lindsey Macmillan, John Micklewright Anna Vignoles and an anonymous referee for useful discussions about, and comments on, this paper.

Introduction and background

There is an established literature that suggests teacher perceptions of children affect how they interact with them, how and what they teach them and how they rate their ability and behaviour (Alvidrez and Weinstein, 1999). Evidence also indicates that a teacher's perception of a child is often based on ascriptive characteristics such as gender (Doherty and Connelly, 1985; Cotton, 1989; Jussim and Eccles, 1992), ethnicity (Dusek and Joseph, 1983, Tom and Cooper, 1985; Noguera, 2003; Crozier, 2005) and socio-economic background (Cotton, 1989; Alvidrez and Weinstein, 1999) independent of a child's ability. This is potentially problematic because every stage of the English education system relies on teacher assessment of a pupils ability, either directly as in the Foundation Stage Profile which measures teacher perception of the ability of 5 year olds across a range of subjects or through (usually non-blind) marking of course work which contributes to a large proportion of GCSE examination marks and even A-level, undergraduate and postgraduate degrees. This paper builds on this literature by examining whether teacher ratings of a child's ability and behaviour are associated with their perception of a particular ascriptive characteristic of the children they teach – their attractiveness.

The hypothesis is that attractiveness, or beauty as it is often referred to in the literature, triggers a positive response in teachers which leads them to evaluate the ability and behaviour of some students more favourably than others. This is in line with evidence that teachers rate the ability of children they perceive to be tidy (Doherty and Conolly, 1985), well groomed and well dressed (Cotton, 1989) above other children. It is also in line with a range of other evidence which shows that people are biased in favour of attractive people in terms of both their perceptions and behaviours (Hatfield and Rapson, 2000). As early as the 1970s Dion et al. (1972) showed a sample of men and women college year book photos and asked them to comment on the people in the photos. The reviewers of the pictures assumed the good looking people were more sociable, outgoing, interesting and exciting than their less attractive peers. The reviewers also thought the attractive people would have more fulfilling lives, be happier, have more successful marriages, better jobs and more satisfying lives than the less attractive people in the yearbook.

There is a great deal of evidence to suggest that these perceptions affect behaviour and attractive people are treated differently to less attractive people. For example, attractive people have more success in romantic relationships (Kleck and Rubenstein, 1975) and the romantic partners of attractive people express more commitment to the relationship than partners of less attractive people (Sangrador and Yela, 2000). The attractive are also at an advantage in the social arena where they are more likely to be approached and spoken to in social settings (Garcia et al., 1991); are more likely to receive help and have their possessions returned if lost (Benson et al., 1976) and be given directions if they themselves become lost (Wilson, 1978). In the economic arena attractive people are more likely to be hired (Hatfield and Sprecher, 1986; Hatfield and Rapson,

2000), be promoted (Morrow et al., 1990; Hatfield and Rapson, 2000), have higher incomes (Quinn, 1978; Hamermesh and Biddle, 1994; Hatfield and Rapson, 2000; Biddle and Hamermesh, 1998, Harper, 2000) and have faster income growth (Roszell et al., 1989) than less attractive people. In terms of health and wellbeing, children rated as unattractive report less good health outcomes at age 50 (Braakman, 2011). Attractive college students report greater happiness than other students (Diener et al., 1999) and in the general population attractive people are more likely to report being happy (Umberson and Hughes, 1987; Hamermesh and Abrevaya, 2012) and satisfied with life (Peterson and Miller, 1980; Hamermesh and Abrevaya, 2012).

Of most relevance to this present study is Hamermesh and Parker's (2012) study of student feedback on their university teachers. With a sample of around 17,000 student ratings, along with 6 independent measures of attractiveness, Hamermesh and Parker showed that better looking teachers received higher student ratings even after controlling for a range of other factors. This allowed the authors to conclude that 'measures of perceived beauty have a substantial independent positive impact on instructional ratings by undergraduate students' (p13). The idea is that the positive ratings were triggered by positive responses to beauty by the students. This is in line with evidence that suggests that people feel more positive emotions after interactions with attractive people (Garcia et al., 1991; Pataki and Clark, 2004).

The current paper draws on both the teacher perception literature and the beauty literature and examines whether teacher's perceptions of the attractiveness of their pupils is related to how they rate the academic ability and behaviour of those pupils. Based on the evidence presented above we would expect that teachers would rate both the ability and behaviour of attractive students more favourably than that of their less attractive peers. This hypothesis is tested using data from a sample of 9,233 11 year old children born in a single week in England, Scotland and Wales in the National Child Development Study (NCDS). These data are described in the following section before descriptive - and then regression analyses are carried out and discussed. A summary section draws the paper to a close with a discussion of the results and the implications they may have.

Data

This paper uses data from the NCDS which follows the lives of 17,000 people born in England, Scotland and Wales in a single week of 1958. Since the birth survey in 1958, there have been eight further sweeps of cohort members at ages 7, 11, 16, 23, 33, 42, 46 and 50. This paper focuses on the age 11 survey, where teachers were asked to rate the child in terms of attractiveness. Teachers could rate children as either attractive, not attractive, undernourished, slovenly or dirty or as having abnormal features. Due to issues related to abnormal features such as the link with foetal alcohol syndrome this group have been discarded from the analysis, leaving a sample of 9,233 children who were rated as either attractive

or unattractive by the teacher (including those who were undernourished and slovenly or dirty).

Teachers were also asked a number of questions related to their perception of the child's academic ability (a binary response question of whether the child showed any outstanding ability) and their behaviour (whether they were delinquent, rebellious, aggressive or showed other bad behaviours). Teachers were further asked to rate each child's ability on a scale from 1-5 with 1 being exceptional, 2 above average, 3 average, 4 below average and 5 very limited in different areas of learning – general knowledge, numbers, books and oral ability.

During the same survey the children took a General Ability Test (Douglas, 1964) containing verbal and non-verbal items. In addition, the longitudinal nature of the NCDS means that it collects a large range of data related to the children themselves and their families which will allow us to control for background characteristics that could play a role in how a teacher perceives a child, the child's ability and attractiveness - such as gender of the child, a child's prior achievement and behaviour, mother's age, father's education and occupation and family socio-economic group. In this paper these variables are measured at a previous sweep to the age 11 outcomes we are interested in when the children were aged 7 years old.

Results

At age 11 the majority of students are rated as attractive by their teachers. Indeed, as Table 1 shows, of the 9233 children we have data on, 81 percent were rated as attractive, 15 percent were rated as unattractive, 4 percent as undernourished and a further .3 percent as slovenly or dirty. Due to the low numbers in the latter two groups they have been incorporated into the unattractive category for the remainder of the analysis. Whether children rated as attractive vary from those rated as unattractive in other characteristics is a question examined in Table 2. It appears that girls are more likely to be rated as attractive than boys (83.5% compared to 78.7%). And female teachers are more likely to rate students as attractive compared to male teachers (82.2% versus 79.9%). Children rated as unattractive are slightly shorter than attractive children (47.7 versus 48.3 inches) but there is no difference in terms of what they weigh. Those rated as attractive come from slightly smaller families (where the average number of children is 2.9 compared to 3.7 children in families where children were rated as unattractive). Attractive children are also more likely to have a father with post compulsory education (23.7% compared to 14.8%) and have a father in the top social group (5.2% compared to 2.1%).

Tables 1 and 2 here***

Teachers are given no guidance on what constitutes attractiveness but Hakim (2012), in a discussion of what she terms erotic capital, discusses the importance of beauty as the central element of attractiveness, other aspects including sexual attractiveness and sexuality can be ruled out as the children are too young.

Social presentation may play a role and, as discussed above, those children rated as undernourished, slovenly or dirty have been categorised as unattractive in this analysis. We may also suspect there to be ethnic differences in attitudes to attractiveness, as historically different cultures have different perceptions of beauty. Unfortunately, we cannot study the role of ethnic differences in this analysis as the NCDS has so little ethnic variation amongst the cohort children (with 98 percent of children at age 7 being White¹) and we have no information on the ethnicity of the teachers. But research by Cunningham et al. (1995), suggests this may be less of an issue than we think. Their research examined ideas of beauty across different ethnic and cultural groups and found little difference in cross cultural perceptions of beauty. Their analysis produced extremely high correlations between the ratings of different groups (ranging from .91 to .94).

It is important for this analysis that teachers have a similar conception of what is meant by attractiveness. Hammermesh and Biddle (1994) present a range of evidence that shows “within a culture at a point in time there is tremendous agreement on standards of beauty, and these standards change quite slowly” (p1175). We can examine this in the NCDS as there are two measures of attractiveness rated by the teacher when the children are aged 7 and 11. As this paper is interested in is the relationship between a teacher’s perception of attractiveness and how that same teacher rates a child’s ability we need to look at these measures at the same point in time, so we focus on the age 11 measures. However, we do have a measure of attractiveness at age 7 rated by a previous teacher and when we examine that, consistent with the literature, we find considerable consistency in which children are rated as attractive at age 7 and 11. Indeed, 73 per cent of children who are rated in terms of attractiveness at both 7 and 11 are rated as attractive by their teachers at both ages. And 93 percent of children rated as attractive at age 11 were also rated as attractive at age 7 by a previous teacher.

Table 3 here****

The relationship between teacher rating of attraction and their rating of the pupil’s ability and behaviour

The rich NCDS data allow us to look at whether teachers rate the ability and behaviour of children more favourably if they think they are attractive. Looking at a simple cross tabulation in the upper rows of Table 4 it would seem that this is indeed the case. The teacher rated scores in the different areas of learning – general knowledge, numbers, books and oral ability are higher for children perceived as attractive. This is consistently true across the different areas of learning. The mean score for general knowledge among children perceived as attractive by the teacher is 3.1 compared to an average score of 2.5 among the less attractive. For numbers the difference is 3.0 compared to 2.4; books 3.3

¹ 1.1 percent Black, 0.6 percent Indian or Pakistani and 0.9 percent of children identified as having an other ethnicity.

compared to 2.7; and oral ability 3.2 compared to 2.6. These differences may not seem large but when we consider scores for each subject can only vary from 1 to 5 these gaps are considerable. When we look at the total gap across all subjects where scores can range from 1 to 20 attractive children have a mean score of 12.6 compared to 10.2 achieved by the less attractive children. The lower rows of Table 4 show the same scores this time transformed into Z scores (with a mean of 0 and a standard deviation of 1) giving standard deviation differences in the ratings of children across the different areas of learning. This lower part of the table shows less attractive children are rated by the teacher around half a standard deviation beneath the mean across all four areas of learning (-.50 for general knowledge, -.46 for numbers, -.44 for books, -.58 for oral and .56 in total). The attractive children on the other hand receive ratings of around .2 standard deviations above the mean across the different areas of learning (.18 for general knowledge, .17 for numbers, .19 for books, .22 for oral and .22 in total). Moreover, examining Table 5 indicates teachers are much more likely to say a child has outstanding ability if they perceive that child as being attractive (27%) compared to unattractive (13%). And the converse appears to be true for teacher rating of bad behaviour, with teachers more likely to rate a child's behaviour as delinquent, rebellious, aggressive or having other bad behaviour if they perceive that child to be unattractive (13%) compared to attractive (5%).

Tables 4 and 5 here....

These results appear to indicate that teachers' perceptions of a child's ability and behaviour are influenced by their perceptions of a child's appearance, with children rated by teachers as attractive also being rated as more able and less badly behaved. However, it may well be that these teachers are picking up genuine differences in the children. In terms of ability we can test this by comparing the teacher ratings of the children's ability with the actual scores the children achieved in the tests administered to them as part of the survey. In order to do this the teacher ratings of 1 to 5 for each of the 4 areas of learning (general knowledge, numbers, books and oral ability) were summed to give a total teacher rating. These scores were then ranked from highest to lowest. The same was done for the scores achieved by the children in the General Ability Test taken as part of the age 11 survey. Then the location of each child's score in the distributions was compared across the teacher ratings and the different tests. If the child's score in the distribution is lower on the teacher rating than on the survey administered test the child's ability is under rated by the teacher. If the teacher rated assessment of the child is higher than the survey administered tests the teacher has over rated the child's ability. If the scores on the two distributions are within a tenth of a standard deviation of one another the scores are deemed equal. Of course there may be errors and we would not necessarily assume the child's score would match perfectly across the teacher rated ability and the survey test scores but we would expect any error to be unrelated to whether the teacher perceives a child to be attractive or not. However, this does not appear to be the case. Looking at Table 6 we can see that teachers are much

more likely to over rate a child's ability if they perceive that child to be attractive (47%) than unattractive (39%). It is also true that children perceived by the teacher as unattractive are more likely to have their ability under rated by the teacher (80%) than their more attractive counter parts (77%). However, the gap between children perceived as attractive and unattractive is smaller in terms of under rating than over rating.

Table 6 here

Regression Analysis

The analysis so far indicates that a teacher's perception of a child's ability and behaviour may be influenced by their perception of whether that child is attractive or not. However, it may be the case that the relationship can be explained by other factors. For example, we know that the parental education levels are related to the educational achievement of children but parental education may also be related to how attractive a child is due to more educated parents providing better diet and nutrition to their children for example. For this reason the rest of this paper uses regression techniques to examine the relationship between attractiveness and ability controlling for other factors that might be influencing this relationship. These other factors, which include a range of child, family, school and teacher variables, are added to the models in an effort to control for variables that may be related to differences in children's ability, behaviour and attractiveness and are detailed below:

Child and family characteristics

The models control for the gender of the child as that has been found to be related both to academic achievement (Machin and McNally, 2005; Hansen and Jones, 2012) and teacher perceptions of ability (Doherty and Conelly, 1985; Ross and Jackson, 1991). Models also control for the height and weight of a child as this will influence perceptions of beauty. Obesity in particular has been found to be associated with a range of negative perceptions (Puhl and Brownell, 2002). There is also a control for whether the child has experienced pre-school as this has been found to be positively related to cognitive ability in other studies (Hansen and Hawkes, 2009). Mothers age is also controlled for due to the range of negative outcomes that are associated with young mothers (Hawkes and Joshi, 2012) as is the number of children in the household due to the detrimental effects large families have been found to have on children's academic ability in other studies (Iacovou, 2001). There are also controls for father's education and occupation which have repeatedly been shown to be related to a child's academic ability (Dearden et al., 1997; Vignoles and Galinda-Rueda, 2005; Blanden et al., 2012). The models include the region the families live in as there may be regional difference in perceptions of attractiveness (Hamermesh and Abrevaya, 2012) and also regional differences in teacher ratings of ability and

student achievement in tests. Finally, added to the regression models is a measure of parental aspirations for the child to stay on past the compulsory school leaving age which has been shown to be a good predictor of later educational achievement in the NCDS (Gutman, 2008) and also a measure of whether the mother and father read books, newspapers and magazines on a weekly basis as parental reading has been found to be a good proxy measure for parental cultural capital and strongly associated with achievement of children (De Graaf et al., 2000).

Teacher and School controls

In addition to the child and family variables discussed above the models also control for a range of teacher and school variables. The gender of the class teacher is included in the models as previous work has found the attributes of teachers important in teacher ratings (Hoge and Coladarci, 2012) and we may also expect that male and female teacher have different perceptions of what constitutes an attractive student. Models also control for class size as evidence suggests that smaller classes are positively associated with academic achievement (Mosteller, 1995; Angrist and Lavy, 1999) and we might also think that class size may affect teacher perceptions of students – for example in bigger classes with more students teacher perceptions of ability may be less accurate than in smaller class rooms where teachers can get to know individual children better. Using the same rationale the models also control for school size which has been found to influence both student achievement and teacher attitudes (Lee and Loeb, 2000).

The models include a measure of whether the class the child is in is streamed and if so what stream they are in. This inclusion is important as evidence suggests that teachers' perceptions of children's ability is influenced by how able they are told the children are. The classic example of this is Rosenthal and Jacobson's (1968) Pygmalion study which involved giving teachers false information about the learning potential of certain students. Teachers were told that these students had been tested and found to be on the brink of a period of rapid intellectual growth. In reality, the students had been selected at random. At the end of the experiment a number of students achieved higher IQ tests than expected and higher than students of similar ability in the control group. Models also control for whether the children have special educational needs as this will affect not only their ability but also how teachers perceive them, and whether they have free school meals as a proxy for children from disadvantaged backgrounds (Hobbs and Vignoles, 2009).

All these variables are measured in a previous sweep of the NCDS survey when they children were aged 7. It is hoped by including these factors in the regression models we will be controlling for variables that may be related to differences in children's ability, behaviour and attractiveness. Also included in the regression models is a measure of children's behaviour at age 7 as measured by the mother and previous cognitive ability measured in tests at age 7. These variables are important to include as they control for the possibility that some

teachers may have a tendency to mark children highly across both dimensions of attractiveness and ability because they are generous and not because the child is attractive or academically able. These models will effectively be asking what is the marginal effect of attractiveness after accounting for other possible causes of variation in academic ability.

Regression Results

The variables detailed above are added to a model of our summary measure of teacher rating of ability across the four areas of learning (general knowledge, numbers, books and oral ability). This outcome has been transformed into Z scores so that the coefficients from a standard Ordinary Least Squares regression can be interpreted as standard deviation differences in teacher rating of ability given a certain set of characteristics. Table 7 shows 5 specifications, the first, in column one, is simply a regression of ability on beauty and the coefficient of .774 equates to the standard deviation difference between the mean scores for children rated by the teacher as attractive and those rated as less attractive that we saw previously in Table 4. We also know from Table 4 that a standard deviation roughly equates to 3 marks in the teacher rated assessment so a difference of .774 corresponds to a difference of about 2.3 marks.

Column 2 adds child and family level controls, which reduces the coefficient to .576, which is further reduced to .316 when previous ability and behavior are added to that specification in column 3. All of the child and family variables entered into the model are statistically significant. Females, those who attended pre-school, whose parents aspire for them to gain post compulsory education and those whose parents read on a weekly basis receive higher ratings of their ability than other students. So too do children with older mothers and those whose fathers have post compulsory education. Children from larger families and those in single mother households receive lower ratings than other children as do children in lower SES households. Perhaps unexpectedly, heavier children receive higher teacher ratings and taller children lower ratings. In terms of magnitude, the greatest penalty is associated with being in a single mother household (with a coefficient of -.742, equating to a reduction of around 2.2 marks for children in single parent households). On the positive side, parental aspirations attracts a coefficient of .312 standard deviations or an increase in scores of just under one mark. Both previous ability and behaviour attract the expected coefficients. The higher cognitive scores at age 7 are associated with an increase in teacher rated ability at age 11 of .055 standard deviations and bad behaviour at age 7 is associated with a .016 standard deviation reduction in teacher rating of ability at age 11.

When the school level variables are added to the model in column 4 the coefficient on the attraction dummy is .524 standard deviations (or an increase in scores of around 1.6 for children rated as attractive compared to less attractive). This is very similar to the coefficient in the model with child and family controls (column 2). However, when we compare the R squared for both models, the school characteristics explain slightly more variation in teacher ratings of ability

(29.8%) than the child and family model (22.0%). The results show that children with special educational needs receive lower ratings (-.771) as do those with female teachers (-.041) and children who are eligible for free school meals (-.402). Compared to children in un-streamed classrooms, children in middle or lower streams also receive lower ratings (-.109 and -.693 respectively). On the other hand children in the top stream receive higher teacher ratings (.585) than children in un-streamed classes.

The effect of all of these characteristics is examined in column 5 of Table 7 which shows the coefficient on the attractiveness measure along with all the child, family, teacher and school variables as well as measures of previous behaviour and ability to control for the fact that some children are more able and better behaved than others. The results show that a teacher's perception of a child as attractive is positively associated with their ratings of that child's academic ability, even after controlling for a range of other factors. Children perceived as attractive by the teacher receive higher ratings of their academic ability by around .3 standard deviations. This equates to around one point score difference. This may not seem much but when we consider the total scores for this outcome range from 0 to 20 an increase or decrease of 1 point is relatively small but non-negligible.

Unsurprisingly, there are a range of other factors in the models that are statistically significantly related to teacher ratings of ability. Indeed, all the controls that were significant in the previous models remain statistically significant in this full model. So girls receive lower teacher ratings than boys, as do children in households with a larger number of children living there. Children with a female teacher and those identified by their teacher as having or requiring SEN, children in lower streamed classes and those who displayed previous bad behaviour also receive lower ratings. On the other hand children who went to pre-school, those whose parents have aspirations for them to stay on past the compulsory school leaving age and those with parents who read every week receive higher ratings of their academic ability than other children. Children with older mothers, fathers who have post compulsory education and those whose fathers are in SES I (professional occupations) also receive higher teacher ratings of their ability. Children with male class teachers, those in top streamed classes and those who scored higher in ability tests at a younger age also score more highly according to their teacher at age 11.

Table 7 here

Over rating and under rating

The results so far indicate that teachers rate the academic ability of attractive students more highly than less attractive students in terms of their performance across the four areas of learning (general knowledge, numbers, books and oral ability). This is true even after controlling for a range of other factors that could influence the relationship, many of which have been shown to be independently related to teacher's ratings. Before we conclude that teachers are biased towards

attractive children we need to consider that perhaps attractive children are more academically able. Fortunately, we can examine this by using the over rating and under rating variables that were constructed by comparing the teaching ratings with performance in the General Ability Test as previously used in the descriptive analysis.²

Both the over rating and under rating variables are binary variables (either a child's ability is over rated or under rated or not) their analyses therefore requires probit models which take the value of 1 if the child is over rated (under rated) or 0 otherwise. As a probit model the coefficients are marginal effects that can be interpreted at percentage point differences. These marginal effects from the full model are shown in Table 8, the first column shows the characteristics associated with the probability a teacher will over rate a student's performance the second, the characteristics associated with under rating a student's performance. The coefficients on the attractiveness variable show that even after controlling for a range of other factors teachers are more likely to over rate an attractive child's ability (by 22 percentage points) and less likely to under rate their ability (by 13 percentage points) than less attractive class mates. Again, indicating the premium associated with being attractive is greater than the penalty for being less attractive. Interestingly, very few of the control variables are significantly related to over or under rating.

Table 8 here

Robustness checks

So far the results have shown that even after controlling for a range of factors that are related to outcomes including family background measures, school variables and previous ability and behaviour, teachers rate the ability of students they perceive to be attractive above other students. Moreover, these same students are more likely to have their ability over rated, and less likely to have their ability under rated by teachers.

However, we have to bear in mind that our models measure perceived attractiveness and ability at the same point in time and by the same teacher. In order to test the hypothesis that teacher perceptions of attractiveness affect their perceptions of ability it is obviously crucial that both perceptions are measured at

² As a reminder we compare the position of each child in the percentile ranking in the summed teacher ratings of ability in the 4 areas of learning (general knowledge, numbers, books and oral ability) to their position in the percentile rankings of the General Ability Test the children sat at age 11. If the child's score in the distribution is lower on the teacher rating than on the GA test the child's ability is under rated by the teacher. If the teacher rated assessment of the child is higher than the GA test the teacher has over rated the child's ability. If the scores on the two distributions are within a tenth of a standard deviation of one another the scores are deemed equal. As noted previously there may be errors and we would not necessarily assume the child's score would match perfectly across the teacher rated ability and the survey test scores but we would expect any error to be unrelated to whether the teacher perceives a child to be attractive or not.

the same time and by the time teacher. However, this approach does mean we may be concerned that our attractiveness measure is endogenous. In other words as the teacher rates both attractiveness and ability we may be worried that unobserved characteristics of the teacher may cause him or her to rate pupils, both in terms of their attractiveness and ability, in specific ways. Including measures of previous ability and behaviour (rated by a different teacher) as we have done goes some way to addressing this issue. But we can take this further by incorporating data on attractiveness measured at age 7 to carry out 2 robustness checks.

The first test simply redefines the attraction variable to incorporate the age 7 data on attractiveness rated by a different teacher and re-runs the same regressions but selecting only children who are rated as attractive (or not) at both ages 7 and 11 where they are rated at both ages and taking the age 11 information where there is none at age 7. The results of this exercise, summarised in Table 9 (with the full set of controls shown in Appendix 1), produce reassuringly similar results.

The second robustness check goes one step further and adopts an instrumental variable approach, instrumenting the potentially endogenous attractiveness measure at age 11 with an exogenous attractiveness measure rated by a different teacher at age 7. If the measure is endogenous by running an instrumental variable regression we should capture only the effects on the outcomes on shifts in beauty whereas OLS not only captures the direct effect of attractiveness on ability but also the effect of the endogeneity. For an instrument to be valid it must be highly correlated with the endogenous variable, and not correlated with the outcome or error term. We would expect this age 7 variable to be a good instrument as we have already seen it is highly correlated with the age 11 measure of attractiveness but is unlikely to be correlated with the outcome measures at age 11 as the measures are rated by different teachers.

When the instrumental variable models are run (the coefficients are summarised in Table 10, and the full results displayed in Appendix 2) we get similar result to the original regressions. Pupils who are perceived to be more attractive by their teachers are also rated more highly in terms of their academic ability across the four areas of learning. The coefficient is larger in magnitude (.556) than the original OLS results (.298), equating to a difference of almost 2 marks rather than the 1 mark using OLS estimates. Like the probit regressions the IV results also show that attractive pupils are more likely to have their ability over rated (.375) and less likely to have their ability under rated (-.105) than less attractive pupils.

Tables 9 and 10 here....

Summary and discussion

This paper tested whether a teacher's perception of a child's attractiveness is related to their perception of the child's ability. Results show that not only do teachers rate the academic ability of attractive students more highly than less

attractive students in terms of their performance across the four areas of learning (general knowledge, numbers, books and oral ability) but they are also more likely to over rate and less likely to under rate their ability than other students. This is true even after controlling for a wide range of other factors related to the child, their family, their teacher and their school that could influence the relationship, many of which have been shown to be independently related to teacher's ratings (less so for over and under rating). It is also true if we use the age 7 measure of attractiveness to define a tighter definition of beauty and when we use it to instrument for any potential endogeneity.

Should we be worried about these results? On the one hand they show teachers are no different from anyone else. They respond more favorably to attractive people. This fits with the extensive literature that attractive people are perceived, and treated, more positively than others. These findings are also in line with evidence that teachers' perceptions of children are often based on ascriptive characteristics independent of a child's ability and this affects both teacher and child (Alvidrez and Weinstein, 1999). If teachers respond differently to children due to their gender (Doherty and Connelly, 1985; Cotton, 1989; Jussim and Eccles, 1992), ethnicity (Dusek and Joseph, 1983, Tom and Cooper, 1985; Noguera, 2003; Crozier, 2005) and socio-economic background (Cotton, 1989; Alvidrez and Weinstein, 1999) then coupled with what we know about reactions to beauty we would indeed expect the results we have found.

However, it is equally true that teacher ratings of ability play an important part in children's education and influence things like decisions about streaming, course content, examination entries and coursework marking. Where any decisions about education are influenced by anything other than academic ability, as this paper indicates is the case, this would seem less than ideal. The push for a move away from our current system of heavy focus on teacher assessments and coursework towards a more exam based system could possibly go some way to improving the situation. However, teacher perceptions about ability are still likely to influence children in more subtle ways. Teacher perceptions filter down to students through classroom interaction which in turn influences children's perceptions of their own ability, their motivation, engagement in school and, ultimately, their achievement. Therefore, even in a situation where all teacher assessments were removed from the system teacher perceptions of attractiveness would still put some children at an advantage in the education system. In order to tackle this problem efforts need to be directed at helping teachers identify and address the role their perceptions play in creating inequalities.

It should be borne in mind that the results we are looking at here refer to perceptions of attractiveness and teacher ratings made in 1969 and it is very likely that societal views on beauty and how teachers approach rating pupils have changed over this period. However, this does not mean the results are any less relevant for policy makers, teachers, or pupils today. Indeed, we would argue that in today's media rich society, where media consumption and screen time is high, and where media content is increasingly concerned with beauty and

image, perceptions of attractiveness held by teachers and others in society are likely to be more firmly entrenched now than in the past.

At the same time the educational system and teacher training have seen huge changes: Whether such changes would be enough to overcome the issues around teacher perceptions and the negative effects they may have on children is, unfortunately, a question we cannot answer. This is because newer surveys, like the Millennium Cohort Study (MCS), would not be allowed to collect information on the attractiveness of their child cohort members as ethical concerns have also increased over time. However, recent work by Campbell (2015) using the MCS to examine stereotyping of pupils by teachers as recently as 2008 found 7 year old pupils who performed equivalently on cognitive tests were over- and under-rated by their teachers depending on ascriptive characteristics such as ethnicity, family income, gender and SEN status. Campbell concluded that historical attempts to promote equity in the education system had not resulted in parity of perception. This evidence, based on different types of ascriptive characteristics not only mirrors the findings in this paper but supports our conclusions that the problems identified in this paper, using older data, are still likely to be a problem today. And despite efforts to promote greater equality in the education system and massive improvements in teacher training since the 1960's it remains the case that more needs to be done to reduce inequalities created by teachers' perceptions and judgements.

References

- Alvidrez, J., Weinstein, R. (1999). Early teacher perceptions and later student academic achievement. *Journal of Educational Psychology*, 91(4): 731-746.
- Angrist, J. and Lavy, V. (1999) Using Maimonides Rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, 114, 2:533-541.
- Benson, P., Karabenick, S and Lerner, R. (1976) Pretty Pleases: The effects of physical attractiveness, race and sex on receiving help. *Journal of experimental Social Psychology*, 12, 409-415.
- Biddle, J., and Hamermesh, D. (1998) Beauty, productivity and discrimination: Lawyers' looks and lucre. *Journal of Labor Economics* 16(1), pp. 172-201.
- Blanden, J. Gregg, P. And Macmillan, L. (2012) Intergenerational persistence in income and social class: The impact of increased inequality. *Journal of the Royal Statistical Society: Series A-Statistics in Society*.
- Braakman, N. (2011) Too ugly to be healthy? Physical attractiveness, health and why they are related Newcastle University Business School, Working Paper No 2011/04.
- Brophy, J. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, 75, 631-661.
- Campbell, T. (2015) Stereotyped at seven? Biases in teacher judgments of pupils' ability and attainment. Forthcoming in the *Journal of Social Policy*.
- Cotton, 1989; Expectations and Student outcomes. *School Improvement Research Series No 7*.
- Crozier G. (2005) There is a war against our children: Black parents' views on their children's education. *British Journal of Sociology of Education*, 26(5): 585-598.
- Cunningham, M., Roberts, A., Barbee, A., Druen, P. and Wu, C. (1995) Their idea of beauty are, on the whole, the same as ours: Consistency and variability in the cross cultural perception of female physical attractiveness. *Journal of Personality and Social Psychology*, 68: 261-279.
- Diener, E., Suh, E., Lucas, R. and Smith, H (1999) Subjective Wellbeing: Three decades of progress. *Psychological Bulletin*, 125, 2: 276-302.
- Dion, K., Berscheid, E. and Walster, E. (1972). What is beautiful is what is good. *Journal of Personality and Social Psychology*, 24: 285-290.
- De Graaf, N., De Graaf, P. and Kraaykamp, G. (2000) Parental cultural capital and educational attainment in the Netherlands : A refinement of the cultural capital perspective. *Sociology of Education*, 73: 92-111.
- Dearden, L., Machin, S. and Reed, H. (1997) Intergenerational mobility in Britain. *Economic Journal*, 107(440):47-66.

Doherty, J., and Conolly, M. (1985). How accurately can primary school teachers predict the scores of their pupils in standardised tests of attainment? A study of some non-cognitive factors that influence specific judgments. *Educational Studies*, 11: 41-60.

Dusek, J., and Joseph, G. (1983). The bases of teacher expectancies: A meta-analysis. *Journal of Educational Psychology*, 75: 327-346.

Garcia, S., Stinson, L., Ickes, W., Bissonnette, V. and Briggs, S. (1991) Shyness and physical attractiveness in mixed-sex dyads. *Journal of Personality and Social Psychology*, 61:35-49.

Gutman, L and Akerman, R. 2008 Determinants of aspirations. Centre for Research on the Wider Benefits of Learning research Report 27.

Hansen, K. and E. Jones (2010) Age 5 cognitive development in England, *Child Indicators Research*, 3: 105-126.

Hansen, K. and D. Hawkes (2009) Childcare and child development, *Journal of Social Policy*, 38: 211-239.

Hamermesh, Daniel S. (2011) *Beauty Pays*. Princeton, NJ: Princeton University Press, 2011.

Hamermesh, D. and Abrevaya, J. (2011) *Beauty is the Promise of Happiness?* IZA Discussion Paper 5600.

Hamermesh, D. and Biddle, J. (1994) Beauty and the labor market. *American Economic Review* 84(5):1174-1194.

Hamermesh D. and Parker A. (2012) Beauty in the classroom: Instructors' pulchritude and putative pedagogical productivity. *Economics of Education Review*.

Harper, B. (2000) Beauty, stature and the labour market: A British Cohort Study. *Oxford Bulletin of Economics and Statistics* 62: 771-800.

Hatfield, E. and Sprecher, S. (1986) *Mirror, Mirror...: The Importance of Looks in Everyday Life*, Albany, NY: State University of New York Press.

Hatfield, E. and Rapson, R. (2000). Physical attractiveness. In W. E. Craighead and C. B. Nemeroff (Eds.). *The Corsini encyclopedia of psychology and behavioral science*. Vol. 3. New York: John Wiley and Sons.

Hawkes, D. and Joshi, H. (2012) Age at motherhood and child development: Evidence from the UK Millennium Cohort, *National Institute Economic Review* 222 (1): R52-R66.

Hobbs, G. and Vignoles, A. (2009) Is children's free school meal eligibility a good proxy for family income? *British Educational Research Journal*, 36(4); 1469-1518.

Hoge, R. and Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research*, 59: 297-313.

Iacovou M. (2001) Family composition and children's educational outcomes. mimeo

Jussim, L. and Eccles, J. (1995). Are teacher expectations biased by student's gender, social class, or ethnicity? *Stereotype accuracy: Towards appreciating group differences* Washington, DC: American Psychological Association.

Kleck, R. and Rubenstein, C. (1975) Physical attractiveness, perceived attitude similarity and interpersonal attraction in an opposite sex encounter. *Journal of Personality and Social Psychology*. 31(1):107-114.

Lee, V. and Loeb, S. (2000) School size in Chicago elementary schools: Effects on teachers' attitudes and students' achievement. *American Educational Research Journal*. 37(1):3-31.

Edward P., Lemay, Jr, Clark, M. and Greenberg, A. (2010) What is beautiful is good because what is beautiful is desired: Physical attractiveness stereotyping as projection of interpersonal goals. *Personality and Social Psychology Bulletin* 36, 3: 339-353.

Machin, S. and McNally, S. (2005) Gender and student achievement in English schools. *Oxford review of Economic Policy* 21(3): 357-372.

Morrow P., McElroy, J., Stamper B. and Wilson, M. (1990) The Effects of physical attractiveness and other demographic characteristics. *Journal of Management*, 16,(4): 723-736

Mosteller, F. (1995) The Tennessee Study of class size in the early school grades.

Noguera, P. (2003) The trouble with black boys: The role and influence of environmental and cultural factors on the academic performance of African American males. *Urban Education*, 38: 431-459.

Pataki, S. and Clark, M. (2004) Self presentations of happiness: Sincere, polite or cautious? *Personality and Social Psychology Bulletin*, 30: 905-914.

Puhl, R. and Brownell, K (2001) Bias, discrimination and obesity. *Obesity Research*, 9: 788-805.

Quinn, R. (1978) Physical deviance and occupational mistreatment: The short, the fat and the ugly, Unpublished paper, Institute for Social Research, University of Michigan.

Rosenthal, R. and Jacobson. L. (1968) *Pygmalion in the classroom: Teacher expectation and Pupils' intellectual development*. New York: Holt, Rinehart and Winston, Inc., 1968.

Ross, S. and Jackson, J. (1991). Teachers expectations for Black males and Black females academic achievement. *Personality and Social Psychology Bulletin*, 17: 78-82.

Sangrador, J. and Yela, C. (2000) What is beautiful is loved: Physical attractiveness in love relationships in a representative sample. *Social behaviour and personality*, 28(3);207-218.

Tom, D., Cooper, H. and McGraw, M. (1984). Influences of student background and teacher authoritarianism on teacher expectations. *Journal of Educational Psychology*, 76, 259-265.

Umberson, D. and Hughes, M. (1987) The impact of physical attractiveness on achievement and psychological well-being. *Social Psychology Quarterly*, 50(3):227-236.

Vignoles, A. and Galinda-Rueda, F. (2005) The declining relative importance of ability in predicting educational attainment. *Journal of Human Resources*. 40(2):335-353.

Wilson, D. W. (1978). Helping behavior and physical attractiveness. *Journal of Social Psychology*, 104: 313-314.