



IUL School of Social Sciences

Department of Social and Organizational Psychology

Evidence of Racial Stereotyping in White Portuguese Children's
Judgments:
The shifting standards model tested with children

Laura Boavida Mégevand

Dissertation submitted as partial requirement for the conferral of
Master in Psychology of Intercultural Relations

Supervisor:
Dr. Elizabeth Claire Collins, Professor
ISCTE – University Institute of Lisbon

June, 2016

Acknowledgements

I would like to address my sincere thankfulness to:

my dissertation supervisor Prof. Elizabeth Collins, for challenging my thinking and supporting me throughout the completion of this work;

Prof. José Mourão, director of the Agrupamento de Escolas, for kindly allowing me to conduct my research in the schools and helping me to make it happen;

Prof. Isabel Bravo, Prof. Rodrigo Gomes and Prof. Ester for going beyond the necessary in helping me setting up the interviews; to the 1st and 4th grade teachers of the Escola de Rana, Escola de Santo António and Escola do Murtal for their kindness and help to accommodate the interviews;

the children who participated in my study, for making this work fun and interesting; and their parents for authorizing them to participate in the study;

my family and friends for their never ending encouragement and support.

Abstract

The shifting standards model (Biernat, Manis & Nelson, 1991) suggests that although frequently members of stereotyped groups are judged as having the characteristics attributed to their group, sometimes they can be judged in comparison to standards specific to their group, thus contrasting from the group stereotype. Objective scales may reveal assimilation effects because they have the same criterion of evaluation for everybody, whereas subjective scales may show contrast or null effects, because they allow the use of different standards of comparison for people from different groups, based on group stereotypes. We examined this model among a novel group: children. As early as 6, children start developing racial stereotypes and may use them to form judgments. We tested the influence of objective versus subjective response scales on white Portuguese first and fourth graders' evaluations of Blacks versus Whites on mathematics ability and basketball performance. No significant results were found in the domain of basketball ability. We found a suggestion of the shifting standards effect with first graders. Unexpectedly, fourth graders' rated Blacks higher objectively than Whites on mathematics ability. For fourth graders, both blacks and whites were rated differently depending on whether the scale was subjective or objective. This raises questions about how to interpret data on stereotyping in children.

Keywords: Social Perception & Cognition; Developmental Psychology; Stereotypes; Shifting Standards; Children

Resumo

O modelo de mudança de *standards* - shifting standards model – (Biernat, Manis & Nelson, 1991) sugere que membros de grupos estereotipados podem ser avaliados como tendo as características atribuídas aos seus grupos mas também, podem por vezes ser julgados em comparação com estereótipos relativos aos seus grupos, e consequentemente, serem avaliados de uma maneira que contrasta com esses estereótipos. Escalas objectivas podem revelar efeitos de assimilação na medida em que o critério de avaliação mantém o mesmo significado independentemente do alvo avaliado. Já escalas subjectivas permitem que um dado alvo seja comparado a *standards* relativos ao grupo ao qual este pertence, podendo revelar efeitos nulos ou de contraste. Examinámos este modelo junto a uma nova população: crianças. A partir dos 6 anos de idade, as crianças começam a desenvolver estereótipos raciais e a utilizá-los para avaliarem os outros. Investigámos a maneira como a avaliação de Negros e Brancos em duas dimensões estereotípicas (matemática e basquetebol) é influenciada pelo tipo de escalada de avaliação (objectiva versus subjectiva), em crianças portuguesas brancas nos 1º e no 4º anos de escolaridade. Não houve resultados significativos no domínio de basquetebol. Para matemática, as crianças no 1º ano de escolaridade avaliaram os alvos negros versus brancos de acordo com o estereótipo de que os Negros são piores a matemática do que os Brancos, atribuindo objetivamente pontuações menos elevadas ao Negros do que aos Brancos. O padrão das avaliações por parte destas crianças sugere um efeito de mudança de *standards*, quando as avaliações são feitas em escalas subjectiva. Contrariamente ao esperado, as crianças no 4º ano de escolaridade não avaliaram os alvos de acordo com o estereótipo acima referido. No entanto, para este grupo etário, tanto as avaliações dos alvos negros como as dos alvos brancos foram diferentes em função do tipo de escala utilizada, o que sugere que o tipo de escala de resposta influencia a forma como os estereótipos são visíveis, apontando para a necessidade de se ser prudente ao interpretar resultados de estudos sobre estereótipos em crianças,

Palavras chave: Percepção Social & Cognição; Psicologia do Desenvolvimento; Estereótipos; Mudança de Standards; Crianças

Table of Contents

CHAPTER I: THEORETICAL FRAMING	1
General Introduction	1
1.1 Stereotypes	2
1.1.1 Stereotyping	2
1.1.2 Stereotypes as standards of judgments	4
1.2. Development of stereotypes and cognitive abilities in children	5
1.2.1 Developmental approaches to stereotypes	5
1.2.2. Racial awareness and stereotyping	6
1.2.3. Stereotypes and prejudice in children	6
1.2.4 Flexible judgments in children	7
CHAPTER II: CURRENT RESEARCH	9
2.1. Method	11
2.1.1. Participants.....	11
2.1.2. Design	11
2.1.3. Materials	11
2.1.4. Procedure.....	12
2.2. Results.....	14
2.2.1 Standardizing scores	14
2.2.2. Mixed Analysis of Variance	15
2.3. Discussion	19
2.3.1. Mathematics Ability Stereotype	19
2.3.2. Basketball Performance Stereotype.....	22
2.3.3. Limitations	23
2.4. Future Directions	24
2.5. Conclusion	26
REFERENCES	27
APPENDIXES	30
Appendix A - Example of the objective questionnaire.....	30
Appendix B - Example of the subjective questionnaire.....	31
Appendix C - Example of presentation order of photographs	32
Appendix D – Parental informed consent form	33

FIGURES INDEX

1. Standardized mathematics ratings for first graders, by scale and race	17
2. Standardized mathematics ratings for fourth graders, by scale and race	18
3. Standardized mathematics Ratings for fourth graders, by race and scale.....	18

TABLES INDEX

1. Reliability analyses for white and black targets, by scale and stereotype 15

Chapter I: Theoretical Framing

Everyday we are confronted by an enormous quantity of information and it would be too heavy to process all of it – thus, we may sort people into groups to reduce the complexity of our world (Mackie, Hamilton, Susskind & Rosseli, 1996). When we categorize a person into a group, we tend to assume that he or she will possess the characteristics that we believe people from that group have (Stangor, 2009). For instance, if we give a hard mathematics problem to solve to a small child, we likely will expect that he will not be able to solve it, even if we have probably not seen that child before nor know anything about him. Nonetheless, because we have a belief that children - as a group - are not smart enough to solve hard mathematics problem, when we see a child, we instantly assume he will have that same characteristic. In some situations, we may use these beliefs about groups – or stereotypes – to judge a person from a certain group or category. Typically, we do so by attributing to a person the characteristics assigned to his group (Biernat, 2003). For instance, in our example, because we believe children, as a group, are not smart enough to solve difficult mathematics problems, we may judge any given child accordingly. Nonetheless, there are times when these stereotypes about groups are not directly visible in people’s evaluations: for instance when asked to evaluate a person on a scale from “a little” to “a lot” of a certain trait, we may compare that person to the beliefs we have about their own group, concerning the trait we want to evaluate. Thus, by comparing different persons to different beliefs based on their own groups, our beliefs may appear masked in such scales (Biernat, 2003).

From age 6, children show evidence of being aware of racial stereotypes – or beliefs - and use them to evaluate members of stereotyped groups (Killen & Rultand, 2011). Between the age of 6 and 10, children’s understanding of groups existing in society and their knowledge concerning stereotypes as well as their expression of these stereotypes changes (McKown & Weinstein, 2003). For instance, as they develop, children start thinking using multiple categories and, under certain conditions, learn how to hide biases in order to conform to societal norms (Baron & Banaji, 2006). Given these developmental changes in how children process information and express their beliefs, the tendency to make within rather than between group comparisons on subjective scales may also change between the ages of 6 and 10.

With the present study we aim to understand (1) how White Portuguese Children

between the ages of 6 and 10 view Blacks and Whites on two stereotypic dimensions: mathematics ability and basketball performance, and (2) how children's evaluations of Blacks and Whites on these stereotypic dimensions are influenced by different types of response scales (objective versus subjective). In other words, given that children show signs of stereotyping (i.e. applying stereotypes to group members), we are interested in understanding if, like adults, they also use different standards to judge members of different groups based on the group stereotypes, when making judgments on subjective response scales, and at what age they begin to do so.

1.1. Stereotypes

Stereotypes are the beliefs we hold about the attributes of a person based on the group they belong to, such as personality traits, skills or behaviors (Yzerbyt, Schadron, Leyens & Rocher, 1994). They are cognitive structures that contain the traits associated with certain social categories (Stangor, 2009). Stereotypes inform people about what they can expect of an individual and they influence how people perceive, think about and act towards group members (Dovidio, Hewstone, Glick & Esses, 2010). People develop stereotypes concerning diverse social groups and are likely to use them to form judgments in ambiguous situations or in situations that require a quick reaction (Smith & Mackie, 2007). Devine (1989) found that people could show signs of holding stereotypes, even if these were inconsistent with their beliefs. She suggested a distinction between controlled and automatic processes: whereas personal beliefs tap into controlled processes, stereotypes tap into automatic processes, which occur outside of awareness. Devine (1989) argues that, because throughout their lives people repeatedly hear and see others use some stereotypes, these stereotypes form strong cognitive associations in people's minds, whether or not they believe them. In this way, stereotypes become automatically activated whenever appropriate cues – like information about a person's race, for instance - are present (Devine, 1989). This automatic activation of stereotypes can further influence how people judge members of stereotyped groups on stereotypic dimensions (Amodio & Devine, 2006).

1.1.1. Stereotyping. Research on the extent to which people apply stereotypes to group members (i.e. stereotyping) often shows divergent results: one line of research suggests that people use stereotypes to evaluate group members, whereas another line of

research suggests a “fading” of these stereotypes (Heine, Lehman, Peng, & Greenholtz, 2002). The first line of research shows that people use stereotypes to evaluate members of different groups by assimilating them to the stereotype concerning that group (Biernat & Eidelman, 2007). For instance, an assimilation to the stereotype that men are stronger than women would consist of specific men being judged as stronger than specific women, without empirical evidence to support the judgment. The other line of research on stereotyping does not show this assimilation to stereotypes but rather shows a reduction in the utilization of stereotypes to evaluate stereotyped group members (Biernat, Manis & Nelson, 1991). For instance, Heine, Lehman, Peng and Greenholtz (2002) showed that part of the research on cross-cultural comparisons does not support the “common view” or broad stereotypical ideas that exist concerning different cultural groups. The authors argue two possible explanations for these findings: stereotypes people hold about a specific cultural group may be false or may have vanished, or these stereotypes may be, to some extent, true but are hidden by the way in which judgments are assessed. As these authors argue, cross cultural comparisons often use subjective self-assessment instruments, which are malleable because the meaning attached to an evaluation only makes sense in a specific context. These type of scales - such as Likert-type scales, often used in these types of studies - , imply using a common norm against which people evaluate themselves or others. This norm may be different for different cultures, thus masking some cultural differences that could exist. For instance, if we ask about the frequency of a certain behavior (e.g. “going to the gym”) on a subjective scale, the answering options could range from “never” to “very often.” However, these words can mean completely different things for people in different cultures: “very often” could mean “everyday” for one person but it could also mean “twice per week” for another, depending on what the person answering thinks is the average level of that trait for other people (Heine, Lehman, Peng & Greenholtz, 2002). Moreover, according to the theory of social comparison, when comparing ourselves to other people, we do it in comparison to similar others (Festinger, 1954). Thus, people within a culture are likely to evaluate themselves in comparison with others from the same culture. If people from two cultures differ in terms of an attribute, this difference will likely not be visible in subjective language because each person will compare themselves to different reference points. Likert-type scales often rely on these within-group comparisons: people make judgments according to what they estimate to be the average level that the appropriate group for the criterion evaluated (Collins, Crandall & Biernat, 2006). For instance, if we are interested

in comparing Italians and Germans in terms of talkativeness, a typical researcher might approach Germans and Italians and ask “how talkative are you?” Both Germans and Italians may rate themselves as “very talkative.” Does this mean Germans and Italians are equally talkative, meaning that the stereotype that Italians are more talkative than German is inaccurate? Heine, Lehman, Peng and Greenholtz (2002) would argue it is rather because Italians and Germans tend to compare themselves to people from their own group, hence a German may say he is very talkative because he believes himself to be more talkative than the average German: the talkativeness of this “average person” differs across cultures.

1.1.2. Stereotypes as standards of judgments. Beliefs about groups of people are likely to color our judgments of members of those groups. This can happen either by assimilating a person to a stereotype – judging them as having the attribute attached to the group they belong to – or by contrasting them with this stereotype – judging a person (often negatively) in comparison to the stereotypical expectation about their group (Biernat, 2003). When we judge a person, we are likely to do so in reference to certain expectations or standard that we may have about that person (Biernat & Eidelman, 2007). In this context, stereotypes offer cues to set this criterion or standard to compare a target against, as they provide information on what is expected from a stereotyped group member (Biernat, 2003). If there are relevant group stereotypes for a dimension being evaluated, a target will automatically be evaluated in comparison to his own group (Biernat & Manis, 1994). This happens when a judgment is made on a subjective rating scale, such as likert-type scales, as the meaning assigned to words in such scales can fluctuate. For instance, we implicitly know that “very” heavy means different things if we talk about a sumo wrestler or about a new born child. Thus, subjective scales allow a certain flexibility to choose the standard of comparison to evaluate a target, based on that target’s group membership and the stereotypes attached to that group. (Biernat, 2003). This is the idea behind the shifting standards model (Biernat, Manis & Nelson, 1991) that suggests that when a judgment is made on a subjective rating scale, adults automatically and without awareness evaluate a target in comparison to a within category criterion or standard (if there are relevant group stereotypes for the dimension evaluated). For instance, assuming there is the belief than men and women differ in terms of physical strength, gender could be a relevant cue to assess a person’s physical strength. A man and a woman target could receive the same subjective evaluation (e.g. “very strong”),

although objectively, women are nonetheless believed to be less strong than men. This means that although stereotypes may not be directly visible when judgments are made on a subjective scale (e.g. men and women receiving the same subjective rating for physical strength), they can still influence people's responses (Biernat, 2003; Collins, Biernat & Eidelman, 2009). Evaluations made on objective scales – scales with externally fixed criteria –, however, do not allow for the shifting of the standard used to compare targets, because the criteria have the same meaning regardless of the target's group (Biernat, 2003). For instance if we measure physical strength objectively by the amount of weight in kilograms a person can lift, this criterion (i.e. number of kilograms lifted) is the same regardless of a target's group membership. Overall, the shifting standards pattern consists of a target being assimilated to stereotypes on objective scales (i.e. judged according to stereotypes), which exposes the underlying stereotype and contrasting with the stereotype on subjective measures, which gives the external impression that the stereotype is not influencing judgments (Biernat, 2003). The Shifting Standards Model has been studied with different stereotypes (e.g. work related competence, weight, financial income, mathematics ability, verbal ability, athleticism) linked to diverse social groups, for adults (see Biernat & Fuegen, 2001; Biernat & Eidelman, 2007; Biernat, Manis & Nelson, 1991; Collins, Crandall & Biernat, 2006; Collins, Biernat & Eidelman, 2009; Kobrynowicz & Biernat, 1997; Biernat & Manis, 1994). However, this model has not yet been studied with children, although there is evidence that children too use racial stereotypes to form evaluative judgments about group members.

1.2 Development of stereotypes and cognitive abilities in children

1.2.1 Developmental approaches to stereotypes. According to social learning theories, stereotypes are acquired through socialization and learned from the environment. In contrast, cognitive theories put an emphasis on the development of cognitive abilities as the basis for the acquisition of stereotypes (Bar-Tal, 1996). The first approach sees socialization as the main vehicle for children to learn about stereotypes, whereas the second posits that children's cognitive limitations lead them to show more bias when they are younger. A more recent theoretical approach for understanding the development of stereotypes and prejudice in children - the developmental intergroup theory (DIT) (Bigler & Liben, 2007) - emphasizes the influence of both types of processes. According to the DIT, the formation of social stereotyping requires three types

of mechanisms. First, it requires a “psychological salience of person attributes.” This means that, before starting to stereotyping, children must observe different attributes in people. They may observe physical differences such as gender or skin color, and may infer that there must be some inherent differences between people, based on the differences they notice in how the world seems to be organized environment – for instance, they may notice that some groups are smaller than others or notice segregation. Second, in order to hold a stereotype a person must be capable of seeing another person as belonging to a certain category or group (Mackie, Hamilton, Susskind & Rosselli, 1996). According to the DIT, children need to be able to categorize individuals based on relevant salient attributes before they are able to stereotype. Once they are able to organize people into categories based on salient attributes, children may then develop stereotypes and prejudice through processes like essentialist thinking, that consists of thinking that there are innate features that are fixed from birth (Bigler & Liben, 2007). Along with gender or skin color, race too becomes a salient category that pre-school children discover. Soon after, they start seeing members of such categories as sharing the same “essence” (Killen & Rutland, 2011).

1.2.2 Racial awareness and stereotyping in children. Research suggests that stereotype awareness (i.e. knowing that a stereotype concerning a specific group exists) starts developing at a young age. For instance, 4 years old children are aware of gender stereotypes assigned to boys and girls (Killen & Rutland, 2011) and 3 to 4 years old children are able to sort people by race, suggesting that racial awareness starts at an early age (Aboud, 1988). Later on, children start showing signs of essentialist thinking about out-groups (i.e. groups other than their own group) and start learning about the stereotypes associated with these out-groups (Pauker, Apfelbaum & Ambady, 2010). Pauker, Apfelbaum and Ambady (2010) found that 6 to 7 and 7-10 year old children apply racial stereotypes to out-group members. As they grow, children are more likely to show evidence of knowing about broadly held stereotypes: in fact, 10 year old children seem to possess knowledge about broadly held racial stereotypes (McKown & Wienstein, 2003).

1.2.3 Stereotypes and prejudice in children. Research on stereotypes in children is often confounded with the affective component of bias – prejudice (Pauker, Ambady & Apfelbaum, 2010). Research in North America suggests that white majority children

show more prejudice towards Blacks than Whites between age 3 to 5 but that this racial bias is weaker from 7 years onward (Williams, Best & Boswell, 1975; Aboud, 1998). Because of the lack of cognitive flexibility, younger children are not able to weight multiple categories at the same time, and are thus more prone to showing biases and stereotypes than older ones (Winkler, 2009). Around age 7-8, children start being able to weigh multiple categories at the same time, which leads to fewer signs of prejudice in older children, but only when attitudes are measured explicitly (Killen & Rutland, 2011). When racial attitudes are measured indirectly (i.e. with measures that can tap into people's attitudes without them being aware or able to control responses) they seem to stay constant throughout life (Baron & Banaji, 2006). An explanation for such findings is that between the ages of 9 and 12, children from majority groups develop the ability to respond to the dual social norms of ingroup favoritism and outgroup non-discrimination, as well as social needs to appear unbiased in some contexts (Rodrigues, Rutland & Collins, in press) and thus in many contexts, may "mask" their beliefs when they are expressed explicitly.

1.2.4 Flexible judgments in children. There is a small amount of research examining the flexibility of judgments in children. In self-evaluations of their performance in mathematics, elementary school children evaluated their performance by comparing it to their performance in other school subjects – not by comparing it to other children. Thus, a child that performed higher than another in a given subject might, nonetheless, rate himself subjectively lower because of his own performance on another task (Ehm, Lindberg & Hasselhorn, 2014). Although this does not tell us whether children use different references for comparisons for others (and is probably just a sign of being self-centered), these intra- rather than inter-individual comparisons children suggest that they may somehow have a sense that they should use a special standard – at least for themselves - and not compare themselves to the everyone.

Given the cognitive changes between the ages of 6 and 10, in terms of the development of the ability to understand and conform to different social norms depending on the context, as well as the development of the cognitive structures that allow children to think using more categories (Killen & Rutland, 2011), it is possible that across that age range, children also develop an understanding of different standards of judgments for other people and likely begin to automatically shift these standards to judge people from

Racial Stereotyping in White Portuguese Children

different groups and start making within-group comparisons when asked to make subjective judgments.

Chapter II: Current Research

The goal of the present study is to understand how children's judgments are influenced by racial stereotypes and how different response scales (objective versus subjective) affect these judgments. Specifically, the aim is to (1) confirm that white Portuguese children possess the stereotypes about Black and White children that adults hold, and, in case they do, (2) investigate if they show evidence of shifting standards and at what age.

Children learn different theories from the society they live in, which serve as a basis for developing social categories, stereotypes and prejudice (Bar-Tal, 1996), showing that to a certain extent, stereotypes may be transmitted by society. Pre-tests on the content of stereotypes in Portugal indicate that black Portuguese adults are seen by white Portuguese adults as significantly more athletic than white Portuguese adults (Bianchi - unpublished data, 2008). For this study, we assumed that white Portuguese children would also hold the belief that black children are more athletic than white children. Furthermore, a study from Rodrigues, Monteiro and Rutland (2012) with white Portuguese children shows that white children are seen as intelligent, whereas black children are less so. We also assume for the present study that white Portuguese children generally believe this stereotype that white children are more intelligent than black children. In order to see if children show evidence of shifting standards, we operationalized the stereotypes mentioned above by choosing characteristics related to these general stereotypes that could easily be translated in objective units and that would be relevant to a school setting. For athleticism we chose to focus on basketball performance, measured objectively with the number of balls children thought a target could throw into a basketball hoop. We assumed children would hold the stereotype that Blacks are better than Whites in Basketball. Because intelligence is a broad construct, we chose to focus specifically on mathematics performance, assessed on objective units by the number of mathematics test questions a given target would answer correctly. We assumed children would hold the stereotype that Whites are better in mathematics than Blacks.

Considering the developmental changes in children between the age of 6 and 10, the study targeted two different age groups: first graders (5 to 6 years olds) and fourth graders (9 to 10 years olds).

In line with the reasoning presented earlier and because we understand stereotypes as cognitive structures that may influence people's judgments regardless of their attitudes (Devine, 1989), if children possess the stereotypes mentioned above, these should be visible in their evaluations on objective scales, regardless of age, such that black male targets should be given higher ratings than white ones on basketball performance (i.e., more baskets scored) and that white male targets should be given higher scores than black ones on mathematics (i.e., more test questions correct). Second, we expected that first graders may not yet shift standards, as it requires greater cognitive flexibility. Thus, for this age group, we expect no significant differences on the objective versus subjective scales, with white targets rated higher on mathematics ability and black targets rated higher on basketball performance, in both scale conditions. Lastly, because 9-10 year old children seem to be aware of racial stereotypes and have developed stronger cognitive skills that would allow them to shift standards, we expected to see a shifting standards effect in fourth grader's judgments. Specifically, for this age group we expect to see differences in the response scales used such that stereotypes should be visible on the objective scale (cf. Whites rated higher than Blacks on the mathematics ability but lower on basketball) but hidden (cf. no differences between Blacks and Whites) or contrasting with the stereotype (cf. Whites rated lower than Blacks on mathematics but higher on Basketball) on the subjective rating scale.

There is evidence that stereotypes and prejudice are two different constructs (Amodio & Devine, 2006), and we expect children's evaluations to be influenced by the stereotypes presented earlier and to be directly visible on the objective scale condition (by an assimilation effect). Still, we recognize that saying a person is "good in mathematics," can have an affective connotation (and tap into prejudice), especially for children, as "being good in mathematics" may be a desirable skill for them. For this reason, because 4th graders are at an age where they start learning about non-discrimination norms (Rodrigues, Rutland & Collins, in press), we recognized the possibility that they may show high sensitivity to such norms, thus attenuating effects seen for this age group.

2.1. Method

2.1.1 Participants

Participants in this study were 134 children from the first and fourth grade of primary school. For convenience reasons, children were selected from 3 schools in the same school district (Agrupamento de Escolas de Santo António) in Parede, a relatively affluent suburb of Lisbon. Participants were randomly assigned to the objective or the subjective rating scale condition.

All non-white participants were removed from the analyses, as we were interested only in white children's responses. Two participants who gave the same answer across all questions were removed from the analysis, as well as three other participants that were outliers on two or more variables.

The final sample consisted of 112 children. There were 44 first graders (5 to 11¹ years; $M = 6.11$ years, $SD = .94$) and 68 fourth graders (9 to 11 years, $M = 9.04$, $SD = .58$). Around 39% of children were female ($n=44$) and 61% male ($n=68$). Half of the participants ($n=56$) were assigned to the objective scale condition and the other half to the subjective condition.

2.1.2. Design

The study had a 2 (target race) x 2 (stereotypic skill) x 2 (scale type) x 2 (participant school grade) x 6 (order of presentation) design, within subject on the first two variables. Participants were asked two questions about each of ten targets, which corresponded to one stereotypical skill positively associated with Whites (mathematics ability) and one positively associated with Blacks (basketball performance). We chose to have "target" as a within subject variable and "scale" as a between subjects variable, to correspond to the studies reported in Biernat, Manis and Nelson, 1991.

The dependent variable was the evaluations made for each target. Because scale was between subjects, each participant completed either the subjective or the objective scales, for both stereotypic skills.

2.1.3. Materials

Each child was shown a booklet containing 10 photographs of faces of white or black

¹ Only two children were above the age of 7 and neither was an outlier on any variable.

children (2 white girls, 2 black girls, 3 white boys, and 3 black boys). Pictures were chosen from the Child Affective Facial Expression set (CAFE; LoBue, 2015), a database of photographs of ethnically diverse 2 to 8 year-old children enacting emotional facial expressions (anger, fear, sadness, happiness, surprise, disgust, neutral). Data suggests it is a valid and reliable tool for studying emotional facial expressions (LoBue, 2015). In the pictures used for the current research, all children appear smiling and are wearing the same white sheet. Six booklets with the 10 pictures were created, each with a different presentation order, but always starting with the same white girl.

Each participant received either the subjective or the objective version of a questionnaire, which included two questions for each of the 10 targets. The two questions tapped into stereotypical skills (cf. mathematics ability and basketball performance). In the subjective version, children had to answer the question “how good is this person in mathematics” and “how good is this person in basketball” on a Likert-type scale, ranging from 1 to 5 where 1= “not good at all” and 5 = “very good.” In the objective version, for the “good in mathematics” trait, children had to estimate how many questions each target would answer correctly, if he/she were to take a 10-question mathematics test. They answered on a 6-point scale, where 0 = zero questions answered correctly, and 5 = 9 to 10 questions correct. For the “good in basketball” stereotype, children were asked how many basketballs (out of 5) they thought each target would be able to throw in a hoop. They were also given 6 options to answer: from 0 to 5 basketballs.

In order to characterize the sample, at the end of each questionnaire, participants were asked to respond to demographic questions, including age, class and own and parent’s birth country. Children’s ethnicity and gender was identified and subsequently recorded by the experimenter.

2.1.4. Procedure

After distributing informed consent forms to parents and getting parental authorization, each child was told they would take part in a small activity, held in the school, during classes. Each child was individually taken into a quiet room and asked to sit at a table, next to the interviewer. After putting children at ease with a brief introduction explaining what they would be asked to do, they were given an unrelated example where they had to judge the number or amount of balloons shown in a picture, using the same rating scales used in the questionnaires, to ensure that instructions were understood well. Each child was randomly assigned to a type of questionnaire (subjective

versus objective) and given a booklet containing the ten selected pictures of the targets (described earlier), in one of the 6 previously defined orders of presentation. Each child was shown one target at a time from the booklet and asked to rate him or her on the questionnaire on the two stereotypical traits. At the end of each session, which took around 10 minutes, the child was thanked and given a brief explanation, before he or she returned to the classroom.

2.2 Results

2.2.1 Standardizing scores

Following Biernat, Manis and Nelson (1991), to make the data comparable across scale type (subjective versus objective) we created standardized z scores based on participant's own responses across all pictures. We only used data for male targets for three reasons. First, participants saw fewer female targets (4 female versus 6 male); second, one was used as a practice trial – the first target for all participants-; third, stereotypes concerning basketball ability and mental ability, such as mathematics performance, are usually more strongly associated with males than females (Pauker, Ambady & Apfelbaum, 2010). Given these factors, we believe it was best to keep only male targets for our analyses.

Reliability. We examined the reliability of each type of rating for each target type using Cronbach's Alpha. We ran separate tests for the pictures of white boys and of black boys, for each stereotypic skill (i.e. mathematics ability and basketball performance) on each type of rating scale (i.e. objective and subjective), totalling 8 tests. Reliability results were low compared with standard cut-off levels (see table 1). Given the wide variety of possible photos of children, the assumptions people make about others based on just a picture, and the small number of targets (only 3), we were not surprised by these low reliability levels. Although these low reliability ratings raise questions about whether responses to the different targets should be averaged and analysed together, we strongly believe that averaging across images gives a more accurate result than examining the images individually, and thus did so. Analyses indicated that presentation order of the pictures interacted with many variables, so we included it as a covariate in our analyses. Participant gender did not interact with any variable but had a main effect, so we also included it as a covariate, in order to control for gender effects. Without participant gender as a covariate, results were similar (we report in a footnote the cases where marginal results were further from significance, without gender as a covariate).

Table 1

Reliability analyses for white and black targets, by scale and stereotype

Scale type	Stereotype	Race	<i>Cronbach's Alpha</i>	N of items
Objective	Mathematics	White	,387	3
		Black	,478	3
	Basketball	White	,468	3
		Black	,594	3
Subjective	Mathematics	White	,304	3
		Black	,140	3
	Basketball	White	,032	3
		Black	,543	3

2.2.2. Mixed Analysis of Variance

To examine differences between objective and subjective evaluations of white and black targets on the two dimensions (mathematics and basketball performance) among first and fourth graders, we conducted a 2 (target's race: black or white) x 2 (stereotypic skill: basketball or mathematics) x 2 (scale: subjective or objective) x 2 (participant's school grade: 1st or 4th) mixed analysis of variance, within-subjects on the first two variables (ANOVA).

Results showed a marginally significant effect of stereotype ($F(1,106) = 3.905, p = .051$), with higher scores given overall on basketball ($M = .115, SE = .024$) than on mathematics ($M = -.055, SE = .023$). There was also a significant main effect of scale on evaluations, $F(1, 106) = 14.318, p < .001$. Higher scores were given on the subjective scale ($M = .095, SE = .024$) than on the objective one ($M = -.035, SE = .024$). This main effect was qualified by a significant scale by grade two-way interaction, $F(1, 106) = 8.466, p = .004$. Both age groups gave higher ratings on the subjective than on the objective scale, however first graders' ratings were overall more extreme than fourth graders' ($M_{FirstGradersSubjective} = .157, SE = .039; M_{FirstGraders Objective} = -.072, SE = .036; M_{FourthGraderSubjective} = .032, SE = .029; M_{FourthGraders Objective} = .002, SE = .031$). There was also a significant two-way interaction between stereotypic skill and participant school grade, $F(1, 106) = 4.188, p = .043$. Overall, both first and fourth graders gave higher ratings on basketball ($M_{FirstGraders Basketball} = .094, SE = .037; M_{FourthGraders Basketball} = .135, SE = .030$) than on mathematics ($M_{FirstGradersMathematics} = -.009, SE = .036; M_{FourthGraders Mathematics} = -.101, SE = .029$), and these differences were more extreme for older children.

All the lower order interactions were qualified by two higher order interactions that were marginally significant. These were the three-way interaction between

stereotypic skill, race and scale, $F(1, 106) = 3.748, p = .056$, and the four-way interaction between stereotypic skill, race, scale and grade, $F(1, 106) = 3.755, p = .055^2$. All other main and interaction effects were not significant ($F_s < 2.856$ and $p_s > .094$).

We were interested in a stereotype typically attributed to Whites (cf. being good in mathematics) and another typically attributed to Blacks (cf. being good in basketball). We expected these to function independent of one another. Thus, to understand our results better, we analysed the ratings of each stereotypic skill separately. Thus, to examine whether there were differences in how children at different ages used the different scales to judge the different targets on each skill we conducted a 2 (grade) x 2 (race) x 2 (scale) mixed ANOVA separately for mathematics and basketball performance, with order of presentation and participant gender as covariates.

Mathematics. There was a significant main effect of participant grade on participants' evaluations of target's mathematics performance, $F(1, 106) = 3.924, p = .050$, with first graders attributing overall higher scores than fourth graders ($M_{FirstGraders} = -.009, SE = .036$; $M_{FourthGraders} = -.101, SE = .029$). This was qualified by a significant three-way interaction of race, scale and grade; $F(1, 106) = 5.021, p = .027$. All other main and interaction effects were not significant ($F_s < 2.882$ and $p_s > .093$). Again, to ease interpretation of the 3-way interaction, we separated the data, this time by participant grade. We chose this variable because we hypothesized that we might see just stereotyping in 1st graders, but shifting standards in 4th graders. We conducted a 2 (race) x 2 (scale) ANOVA, separately for first graders and fourth graders.

First graders. For first graders the main effect of race was not significant, $F(1,40) = 0.006, p = .940$, nor was the two-way interaction between scale and race, $F(1, 40) = 0.596, p = .445$; nonetheless, the pattern of results from these participants followed the classic shifting standards effect, as can be seen on figure 1. Specifically, Blacks were rated slightly lower than Whites on mathematics performance on the objective scale ($M_{Blacks} = -.138, SE = .089$; $M_{Whites} = -.030, SE = .115$); whereas, on the subjective scale, Blacks were rated slightly higher than Whites ($M_{Blacks} = .116, SE = .098$; $M_{Whites} = .013, SE = .126$).

² Without participant gender as a covariate, the stereotype by race by scale 3-way and the stereotype by race by grade by scale 4-way interactions were slightly further from significant ($F = 3.245, p = .074$; $F = 3.42, p = .067$, respectively).

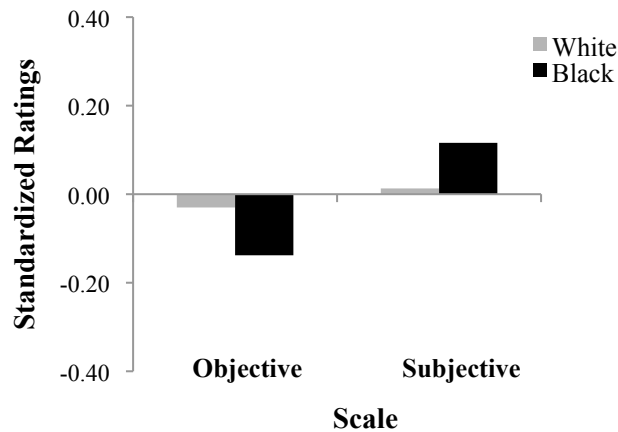


Figure 1. Mean standardized mathematics ratings for First graders by scale and race

Fourth graders. For fourth graders, the main effect of race was also not significant, $F(1, 64) = .220, p = .641$. There was a significant two-way interaction between scale and race, $F(1, 64) = 6.377, p = .014$, however, the pattern of results found does not suggest the classic shifting standards effect (see Figure 2). Instead, for mathematics performance, fourth graders rated Blacks significantly higher than Whites on the objective scale ($M_{Blacks} = .133, SD = .088; M_{Whites} = -.333, SD = .081$), $t(31) = -3.234, p = .003$. On the other hand, the difference between ratings of Blacks and Whites on the subjective scale was not significant, ($t(35) = 0.161, p = .873$): black targets were rated slightly lower than white targets ($M_{Blacks} = -.120, SD = .083; M_{Whites} = -.082, SD = .076$). Furthermore, simple effect tests to compare objective versus subjective ratings indicated that ratings of Blacks and of Whites on mathematics performance differed significantly (respectively, $F(1, 106) = 4.02, p < .05; F(1, 106) = 4.02, p < .05$). (See Figure 3 for a visual display of this result.)

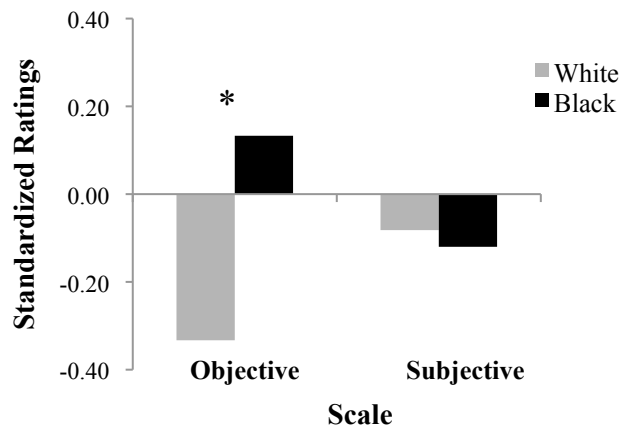


Figure 2. Mean standardized mathematics ratings for Fourth graders by scale and race
* indicates a significant difference ($p < 0.05$) between adjacent columns

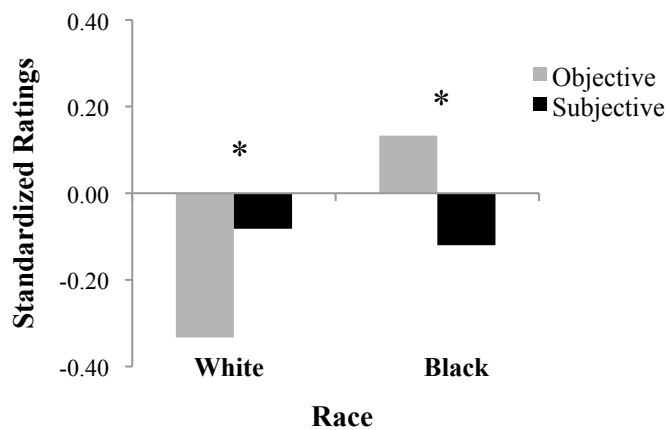


Figure 3. Mean standardized mathematics ratings for Fourth graders by race and scale
* indicates a significant difference ($p < 0.05$) between adjacent columns

Basketball. Looking only at basketball performance, we found a main effect of scale condition, $F(1, 106) = 14.873, p < .001$, with targets rated overall higher on the subjective than on the objective scale ($M_{Subjective} = .207, SE = .034$; $M_{Objective} = .022, SE = .033$). There was also a significant main effect of participant grade, $F(1, 106) = 6.243, p = .014$, such that fourth graders gave higher ratings on basketball than first graders ($M_{FourthGraders} = .135, SE = .030$; $M_{FirstGraders} = .094, SE = .037$). No other main nor interaction effects were significant ($F_s < 2.178, p_s > .143$).

2.3 Discussion

The main goal of this study was to understand how racial stereotypes influence children's judgments of Whites and Blacks and how different response scales (objective versus subjective) affect these judgments. We compared two different age groups (first and fourth graders), as we believed there might be changes in the utilization of stereotypes and shifts based on response scale across that age range. Additionally, we were interested in testing if Portuguese children possess the stereotypes that Whites are better than Blacks in mathematics but worse in basketball.

2.3.1. Mathematics Ability Stereotype

First, we expected first graders to rate targets according to the stereotype that Whites are better than Blacks in mathematics on the objective scale, thus attributing higher ratings to Whites than Blacks. Second, because of their lack of cognitive flexibility, children from this age group were not expected to shift standards when answering on subjective scales, thus we predicted that stereotypes would also be visible in this type of scale.

On one hand, as expected, results for first graders on the objective scale indicated a tendency to attribute higher scores to Whites than Blacks on mathematics ability, although these differences were small and non significant. Still, these results may suggest a certain degree of assimilation to the stereotype that Blacks are worse than Whites in mathematics. In fact, research on stereotype development suggests that 6 year old children are aware of racial stereotypes and start applying them to different group members (Pauker, Appelbaum & Ambady, 2010). Hence, we believe the lack of significant differences could be due to methodology shortcomings (that we will review later) rather than a true indication that children do not hold the stereotype that Blacks are worse than Whites in mathematics.

Furthermore, we predicted no shifting standards for first graders. Contrary to this prediction, these mostly 5 and 6 year old participants rated the Black targets somewhat higher than the White targets on the subjective scale, contrasting with the stereotype. The classic shifting standards pattern consists of an assimilation to stereotypes on objective scales and a null or contrast effect on subjective scales. Although the results were not significant, the pattern of results suggests there may have been this shifting standard

effect for first graders. The shifting standards effect is predicated on the fact that objective scales provide the same anchor for anything that is measured and keep the same meaning across situations; whereas subjective scales can have very different meanings, depending on the context (Biernat, Manis & Nelson, 1991). According to this model, because subjective scales do not have a common reference to judge different targets, targets may be inadvertently compared to standards about their own group, i.e. to within-category standards (Collins, Crandall & Biernat, 2006). Hence, subjective scales can “mask” people’s views and opinions, by allowing the standard against which a target is compared to shift, based on that target’s group membership (Biernat, 2003). In our study, when asked to make a subjective judgment such as how good in mathematics a target is, 1st grade children may have been comparing targets to other ingroup members – white children to other white children, and black children to other black children. Because of the stereotypes concerning Blacks and Whites in terms of mathematics ability, when evaluated subjectively, each target could then be seen as “good.” Even if Blacks are believed to be worse in mathematics than Whites (as seen by the tendency of judgments on the objective scale), Blacks and Whites can receive the opposite subjective evaluation, with a black target seeming “very good” in mathematics - compared to the low standard for Blacks-, and a white one seeming only “good” - compared to the higher standard for Whites. Overall, this pattern of results found suggest it is possible that, on subjective scales, children as young as 6 may already be shifting the standard used to evaluate members of different racial groups, based the racial stereotype concerning mathematics performance.

Fourth graders were expected to rate Whites higher than Blacks on mathematics on the objective scale (i.e. assimilation to the stereotype) and we expected no differences on the subjective scale. Contrary to these predictions, fourth graders rated Blacks higher than Whites on the objective scale (i.e. contrasting with the stereotype on the objective scale), and rated them almost the same on the subjective one. Nonetheless, the results for 4th graders showed that ratings of Whites and Blacks differed significantly depending on the type of response scale, which suggests that how questions are asked may influence how children evaluate targets.

Research into prejudice and stereotype development in children finds that, although they do not fully comply with adult’s views, they are still influenced by what they learn from society and pick up cues about how society works (Bar-Tal, 1996). Also, most 10 year olds (i.e. fourth grader’s age) are aware of broadly held stereotypes

(McKown & Weinstein, 2003). Thus, given that White Portuguese adults possess the stereotype that Whites are more intelligent than Blacks (Rodrigues, Monteiro & Rutland, 2012), it seems odd that children would hold the opposite stereotype, as suggested by our findings on the objective scale. Instead, it seems more likely that the 4th graders of our study hold the stereotypic belief that Whites are better than Blacks in mathematics, but because around age 10 children typically start incorporating anti-discrimination norms into their reasoning (Baron & Banaji, 2006), they may have tried to mask their beliefs in order to appear unbiased, ending up favouring Blacks over Whites on objective scales.

In fact, in our study, children had to rate targets in front of an unknown white experimenter, which could cause a contextual demand to appear socially unprejudiced. In fact, people's judgments are often contaminated by contextual cues, without people being aware of such influence (Wilson & Brekke, 1994). Moreover, the fact that this study utilized a within subjects design for race and had the same number of black and white targets, which does not represent proportions found in the life of these participants, could have made the racial nature of the study salient. We suspect that because of this, the older children in our study, who were at the age where children start incorporating non-discrimination norms (Killen & Rutland, 2011), may have been suspicious that the study was about race and may have therefore tried to appear unbiased, ending up favouring Blacks over Whites. Although race was a salient dimension in both our objective and subjective versions of the study (for the reasons mentioned above), we believe that anti-discrimination norms may have played a stronger role in judgments made on the objective scale, because the objective character of such scales may make children more aware that they can appear biased. On the other hand, because of an unconscious sense that subjective scales are malleable, differentiating between Blacks and Whites on this type of scale may seem less important. In any case, further studies would benefit from testing the same paradigm using a between group manipulation for race.

Interestingly, on the subjective scale, older children did not differentiate between targets. This is what we predicted, and is similar to results found in the literature concerning the expression of explicit racial prejudice for older children (Williams, Best & Boswell, 1975; Killen & Rutland, 2011). Baron & Banaji (2006) found evidence that implicit preferences for Whites over Blacks, measured indirectly and often unconsciously, stay constant throughout life, whereas explicit preference, conscious and more controllable, diminishes with age, often appearing as if they have no explicit preference for either. These authors argue that such differences in implicit and explicit

attitudes lies in the development of moral reasoning in children, as children begin incorporating moral rules, such as anti-discriminatory norms in their judgments. As a consequence, when they explicitly express their attitudes, they may control their responses, to accommodate for this need to appear unbiased, even if their implicit attitudes, which are less controllable, remain the same (Killen & Rutland, 2011). The Shifting Standard Model provides an interesting complementary account for such results, as it suggests that stereotypes can play a role in people's judgments even when they are not directly visible (Biernat, 2003), as is the case for instance when there is no explicit preference for Whites over Blacks in older children's responses in Baron and Banaji's study. In our study, it can be that the fourth graders tried deliberately to mask their stereotypical beliefs by not showing signs of differentiation between Blacks and Whites on the subjective scale. Still, rather than a deliberate choice to appear unprejudiced, it could be that on the subjective scale, 4th graders unconsciously shifted the standard used to judge targets from different groups, as suggested by the shifting standards model and by results for first graders (that suggest shifting standards for participants who answered on the subjective scales).

Nonetheless, as noted earlier, research on stereotypes and prejudice is often confounded in children, and the instruments used to measure stereotyping often measure evaluative judgments (i.e. racial attitudes) instead of measuring stereotypes (Pauker, Ambady & Apfelbaum, 2010). Although we considered stereotypes as cognitive associations that can be automatically activated regardless of stereotype endorsement (Devine, 1989) and expected them to influence children's judgments regardless of personal endorsement, we are aware that we might have been measuring prejudice. In fact, mathematics ability can seem like a characteristic one would like to have, thus tapping into the affective component of stereotypes (i.e. prejudice). If this is true, it would be even more likely that non-discrimination norms may have influenced participants' judgements. Still, we need to understand to what extent moral norms or the shift in standards influence the display of stereotypes. In all cases, this highlights the need to further examine the differences or links between stereotyping and prejudice in children, in order to better understand our results.

2.3.2. Basketball Performance Stereotype

The results did not support our predictions for basketball performance. It is difficult to know whether this was due to a lack of stereotype endorsement, or to

methodological issues with the study. Schwartz & Hippler (1995) found that all questions asked in questionnaires can affect people's responses to the other questions. In our study children were guided by an experimenter through the mathematics questions before the basketball questions. Thus, responses to the mathematics questions may have contaminated the subsequent basketball questions and we should be cautious in interpreting this lack of significant results for basketball. Moreover, basketball is probably a less relevant stereotype for White Portuguese children than is mathematics since they do not encounter it daily. This could explain why no interpretable effect was found for basketball performance.

2.3.3. Limitations

For this study, the photographs of black and white children were selected from a database for research on facial emotions. Pre-tests should have been conducted to make sure all black targets were perceived as Black and that there were overall no substantial differences in terms of targets' likeability, which can have influenced our results.

Additionally, to overcome the reliability problems found in our study, future research should include more images of targets. In addition, a larger sample of participants would also allow greater confidence in the results. Unfortunately, for practical reasons, this was not possible in the framework of this study.

Furthermore, the lack of significant findings, specifically for first graders, could be due to methodology shortcomings. For instance, people are more likely to apply stereotypes based on someone's group membership if they are provided with information related to the specific ability being evaluated, even if that information per se is not relevant for making a fair judgment (Darley & Gross, 1983). This was not the case in our study: the only information participants had to make their judgements was the face of the child and whether the child was Black or White. Perhaps, giving more information, for instance by showing the targets performing the relevant ability (e.g. taking a mathematics test or standing holding a basketball) participants might have been more likely to rate targets more stereotypically. Moreover, simple priming may not be enough to activate stereotypes (Muller & Rothermund, 2014), which could also explain why we did not observe significant findings.

According to the developmental intergroup theory (Bigler & Liben, 2007), the formation of stereotyping and prejudice in children is influenced not only by the development of cognitive capacities (e.g. categorization) but also by external cues

concerning how society is organized. Children use these cues to make inferences about differences between social groups (Bigler & Liben, 2007). For instance, children raised in homogenous environments tend to favour their own race more than children raised in heterogeneous environments (Killen & Rutland, 2011). In this study, we did not take into account children's social environment but future studies would benefit from taking this variable into account. Interestingly, in our study, over 1/3 of parents did not authorize their child to participate in our study. Although we do not have official data about these children, such information might provide us with some insight about who they were and why they were not given approval. It may be that more strict parents, and maybe less open minded, were the ones who did not allow their children to participate. If this is true, it may be that children who participated in our study were brought up in more open and heterogeneous environments.

Lastly, any variance in responses caused by specific characteristics of the target pictures, as well as from the order of target picture presentation was lost in the analysis of the data using repeated measures ANOVA. It is possible that analyses using mixed model ANOVA, which takes into account all the factors from each target that can influence responses might have given different results (Judd, Westfall & Kenny, 2012). This more proper data analysis technique is, unfortunately, also substantially more complicated and thus beyond the scope of this master's thesis.

2.4. Future Directions

This study is the first that we know of examining the importance of response scale in children's judgments of other children on stereotypical dimensions and is an important jumping off point for future research on children's stereotyping, as it shows the way questions are asked may influence children's expressions of stereotypes. Nonetheless, there were many shortfalls that can be addressed in future studies.

For instance, both Black and White targets were presented to each participant, making the racial nature of the study likely to be accessible to children participating. Future research would benefit from using a between participants design for this variable, in order to see if stereotypes would be expressed when children are not aware that they might appear biased.

Future research could focus on disentangling whether evidence of less bias in children as they age is due to social anti-discrimination norms or other processes, such as

the shift of standards based on stereotypes associated with groups. For instance, future research could follow the paradigm used in Biernat, Manis and Nelson (1991) and specifically ask children to evaluate targets in comparison to either the average person (implying same rule for different categories) or in comparison to the average Black or White person. This would help to elucidate to what extent the shift of standards based on group stereotypes plays a role in children's evaluations of stereotyped group members.

Also, this shifting standards effect could be tested in children using stereotypes that are more socially accepted and maybe even accurate, for instance using the stereotype that men are taller than women (Biernat, Manis & Nelson, 1991). This would help disentangle the extent to which the fact that there is no assimilation to stereotypes is due to social norms or a shift in standards, as participants would not have to worry about appearing non-prejudiced.

Furthermore, in this study only the data for White children were analysed but it would be interesting to include children from ethnic minority groups, since the development of stereotypes and stereotyping does not follow the same development trajectory for majority versus minority children (McKown & Weinstein, 2003). This would contribute to our understanding of children from minority groups shift the standard of reference and when they begin to do so.

2.5. Conclusion

This study extends research on the Shifting Standards Model to a previously unstudied population—children—and begins to examine how racial stereotypes are used by children to form evaluative judgments. The study results highlight the need for more research in this area: it remains unclear to what extent shifting standards contributed to fourth graders evaluations of targets on subjective scales versus anti-discrimination norms. Nonetheless, the pattern of responses found for first graders suggests that subjective rating scales may “mask” children’s stereotypical beliefs, as it does in adults. Similarly, fourth graders’ evaluations of targets changed based on the type of response scale. Thus, because subjective versus objective evaluations seem to lead to different types of responses, studies on evidence of biases and stereotypes in children’s judgments should be carefully interpreted. Typically, research on racial attitudes evaluates explicit attitudes through assigning certain types of behaviours or traits to either one of two racially different targets, or to both (Killen, McGothlin & Henning, 2010). Although measures of stereotyping are often confounded with the affective component of the stereotypes, prejudice, some measures ask participants to rate targets on a certain stereotypic dimension. For instance, in one study on stereotypes in children, Augoustinos and Roserwane (2001) asked children to rate targets on a list of 12 traits on a likert-type scale. If we find further evidence that children also shift standards when evaluating members of stereotyped groups on that stereotypical dimension, then it means this could come to play whenever their beliefs are evaluated subjectively.

References

- Aboud, F. E. (1988). *Children and prejudice*. Oxford: Basil Blackwell.
- Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race Bias: evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, Vol. 91 (4), 652–661. DOI: 10.1037/0022-3514.91.4.652
- Augoustinos, M., & Rosewarne, D. L. (2001). Stereotype knowledge and prejudice in children. *British Journal of Developmental Psychology*, 19, 143–156.
- Baron, A., S., & Banaji, M., R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and Adulthood. *Psychological Science*, Vol 17 (1), 53-58.
- Bar-Tal, D. (1996). Development of social categories and stereotypes in early childhood: the case of “the Arab” concept formation, stereotype and attitudes by Jewish children in Israel. *International Journal of Intercultural Relations*, 20, 341-370.
- Bianchi, M. (2008). *Unpublished raw data*.
- Biernat, M. (2003). Toward a broader view of social stereotyping. *American Psychologist*, Vol 58 (12), 1019-1027.
- Biernat, M., & Eidelman, S. (2007). Standards, chap. 13. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of Basic Principles*, 308-333. New York: Guilford Press.
- Biernat, M., & Fuegen, K. (2001), Shifting standards and the evaluation of competence: Complexity in gender-based judgment and decision making. *Journal of Social Issues*, 57, 707–724. DOI: 10.1111/0022-4537.00237
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, Vol 66 (1), 5-20.
- Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology*, 60, 485-499.
- Bigler, R. S., & Liben L. S. (2007). Developmental intergroup theory: Explaining and reducing children’s social stereotyping and prejudice. *Current Directions in Psychological Science*, 16, 162–166.
- Collins, E. C., Biernat, M. R., & Eidelman, S. (2009). Stereotypes in the communication and translation of person impressions. *Journal of Experimental Social Psychology*, 45, 368-374. DOI:10.1016/j.jesp.2008.09.005
- Collins, E. C., Crandall, C. S., & Biernat, M. (2006). Stereotypes and implicit social comparison: Shifts in comparison-group focus. *Journal of Experimental Social Psychology*, 42, 452-45.
- Darley, J.M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, 44, 20-33.
- Devine, P., G., (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social Psychology*, Vol. 56 (1), 5-18.
- Dovidio, J. F., Hewstone, M., Glick, P., & Esses, V. M. (2010). Prejudice, stereotyping and discrimination: Theoretical and empirical overview, chap. 1. In *The Sage Handbook of Prejudice, Stereotyping and Discrimination*.
- Ehm, J. H., Lindberg, S., & Hasselhorn, M. (2014). Reading, writing, and math self concept in elementary school children: influence of dimensional processes. *European Journal of Psychology of Education*, vol. 29 (2), 277-294. DOI: 10.1007/s10212-013-0198-x

- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, vol. 7 (2), 117-140. DOI: 10.1177/001872675400700202
- Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert Scales?: The reference-group effect. *Journal of Personality and Social Psychology*, Vol. 82 (6), 903–918. DOI: 10.1037//0022-3514.82.6.903
- Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103 (1), 54-69.
- Killen, M., McGothlin, H., & Henning, A. (2010). Explicit judgments and implicit bias – a developmental perspective, chap. 8. In S. R. Levy, & M. Killen (Eds.), *Intergroup attitudes and Relations in Childhood through Adulthood*. Oxford: University Press.
- Killen, M., & Rutland, A. (2011). Emergence of social categorization and prejudice, chap. 3, in *Children and Social Exclusion: Morality, Prejudice, and Group Identity*. Willey-Blackwell.
- Kobryniewicz, D., & Biernat, M. (1997). Decoding subjective evaluations: How stereotypes provide shifting standards. *Journal of Experimental Social Psychology*, 33, 579-601.
- LoBue, V. (2015). The Child Affective Facial Expression (CAFE) set. *Databrary*. Retrieved July 10, 2015 from <http://databrary.org/volume/30>.
- Mackie, D. M., Hamilton, D. L., Susskind, J., & Rosselli, F. (1996). Social psychological foundations of stereotype formation. In C. N. MacCrae, C. Stangor & M. Hewstone (Ed.), *Stereotypes and Stereotyping*. The Guilford Press: New York.
- McKown, C., & Weinstein, R. S. (2003). The development and consequences of stereotype consciousness in Middle Childhood. *Child Development*, Vol. 74 (2), 498-515.
- Müller, F., & Rothermund, K. (2014). What does it take to activate stereotypes? Simple primes don't seem enough: A replication of stereotype activation. *Social Psychology*, Vol. 45 (3), 187–193. DOI: 10.1027/1864-9335/a000183.
- Pauker, K., Apfelbaum, E. P., & Ambady, N. (2010). Race salience and essentialist thinking in racial stereotype development. *Child Development*, Vol. 81(6), 1799-1813.
- Rodrigues, R. B., Rutland, A. & Collins, E. C. (in press). A Multi-Norm Structural Social-Developmental Model of Children's Intergroup Attitudes: Integrating Ingroup-loyalty and Outgroup Fairness Norms. In J. Vala, M. Calheiros, & S. Waldzus (Eds.), *The Social Construction of Violence and Conflict Resolution*. New York, NY: Springer.
- Rodrigues, R. B., Monteiro, M. B., & Rutland, A. (2012). O que pensam e dizem crianças portuguesas “Branças” sobre grupos “raciais”? O papel de duas normas grupais conflitantes na expressão e inibição do preconceito na infância. In J. H. António & V. Policarpo (Eds.), *Os imigrantes e a imigração aos olhos dos Portugueses: Manifestações de preconceito e perspectivas sobre a inserção de imigrantes*. Lisboa: Fundação Calouste Gulbenkian.
- Schwartz, N., & Hippler, H. J. (1995). Subsequent questions may influence answers to preceding questions in mail surveys. *Public Opinion Quarterly*, Vol. 59, 93-97. DOI: 0033-362X/95/5901-0004S02.50
- Smith, E. R., & Mackie, D. M. (2007). *Social Psychology* (3rd edition). Philadelphia: Psychology Press.
- Stangor, C. (2009). The study of stereotyping, prejudice and discrimination within social psychology. In T. D. Nelson (Ed.), *Handbook of Prejudice, Stereotyping and Discrimination*. New York: Psychology Press.

Racial Stereotyping in White Portuguese Children

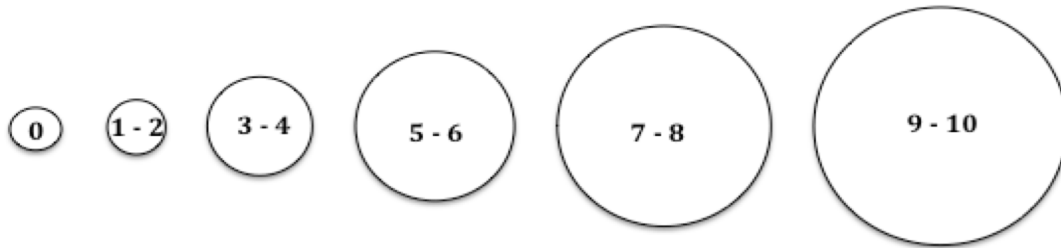
- Williams, J. E., Best, D. L. & Boswell, D. A. (1975). The measurement of children's racial attitudes in the early school years. *Child Development*, 45, 494-500.
- Wilson, T. D. & Brekke, N. (1994). Mental contamination and mental correction: unwanted influences on judgments and evaluations. *Psychological Bulletin*, Vol. 116 (1), 117-142. DOI: 0033-2909/94/S3.00
- Winkler, E. N. (2009). Children are not colorblind: How young children learn race. *Pace: Practical Approaches for Continuing Education*, Vol. 3 (3), 1-8.
- Yzerbyt, V. Y., Schadron, G., Leyens, J. P., & Rocher, S. (1994). Social judgeability: the impact of meta-informed cues on the use of stereotypes. *Journal of Personality and Social Psychology*, Vol. 66 (1), 48-55. DOI: 0022-3514/94/S3.00

APPENDIX A

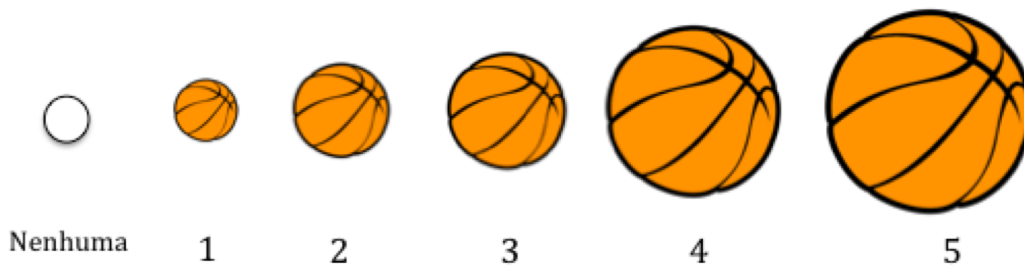
Example of the objective questionnaire.

Primeira pessoa

1. Se _____ fizesse um teste de **matemática** com 10 perguntas, quantas achas que iria acertar?



2. Se _____ pudesse atirar 5 vezes uma bola de **basquetebol**, quantas vezes achas que iria acertar no cesto?



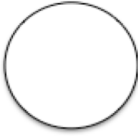
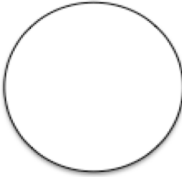
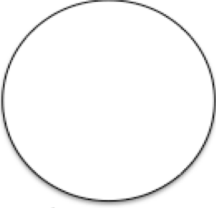


APPENDIX B

Example of subjective questionnaire.



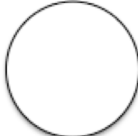
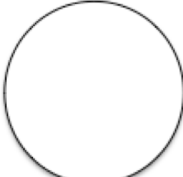
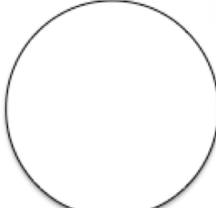
Primeira pessoa

1. Achas que _____ é bom/boa a **matemática**?

				
<u>Não é nada bom/boa</u> a matemática	<u>Não é bom/boa</u> a matemática	<u>É mais ou menos bom/boa</u> a matemática	<u>É bom/boa</u> a matemática	<u>É muito bom/boa</u> a matemática

2. Achas que _____ é bom/boa a jogar **basquetebol**?



				
<u>Não é nada bom/boa</u> a basquetebol	<u>Não é bom/boa</u> a basquetebol	<u>É mais ou menos bom/boa</u> a basquetebol	<u>É bom/boa</u> a basquetebol	<u>É muito bom/boa</u> a basquetebol

APPENDIX C

Example of presentation order of photographs.



Esta é a Marta



Este é o Tiago



Este é o Nuno



Este é o Miguel



Esta é a Rita



Esta é a Maria



Este é o João



Esta é a Paula



Este é o Pedro



Este é o Duarte

APPENDIX D

Parental informed consent form.



PEDIDO DE AUTORIZAÇÃO

Exmo.(a) Sr.(a) Encarregado(a) de Educação,

Gostaria de convidar o seu filho a participar num estudo no âmbito da minha tese de Mestrado em Psicologia das Relações Interculturais, no ISCTE - Instituto Universitário de Lisboa, sob a supervisão da prof. Dr^a Elizabeth Collins.

Objectivo do estudo

O estudo tem como objectivo aprofundar o conhecimento sobre o desenvolvimento das representações sociais acerca de grupos étnicos, em crianças entre os 5 e os 11 anos.

Condições de participação

Cada criança será entrevistada individualmente. Serão mostradas imagens com fotografias de outras crianças e colocadas algumas perguntas de resposta simples e intuitiva. Serão registados alguns dados sobre a criança (e.g. idade, sexo), de modo a poder caracterizar a amostra. O tempo previsto de duração de cada entrevista é de 10 minutos. A participação no estudo é de carácter voluntário.

Confidencialidade e anonimato

As respostas são anónimas e a criança apenas será identificada através de um código. Os dados serão utilizados exclusivamente para o estudo em questão e a sua eventual publicação só poderá ter lugar em revistas da especialidade.

Muito obrigada pela sua colaboração!

Caso queira esclarecer qualquer questão relativa ao estudo poderá fazê-lo através dos seguintes contactos:

Laura Mégevand: email: lbmda@iscte-iul.pt; tel.: 961 632 997

Elizabeth Collins: email: eccse@iscte.pt; tel.: 217 903 031

Laura Mégevand
Aluna do Mestrado em Psicologia

Prof. Dr^a Elizabeth Collins
Orientadora e Investigadora do CIS-IUL

✂----- (devolver assinado nos próximos 2 dias) -----

SIM autorizo / NÃO autorizo o meu/minha filho/a _____
a participar no estudo.

Assinatura do Encarregado de Educação: _____

Data: ___/___/___