

Preto no Branco? Mensuração, Relevância e Concordância Classificatória no País da Incerteza Racial

Jerônimo O. Muniz

Pesquisador e professor adjunto do Departamento de Sociologia da Universidade Federal de Minas Gerais (UFMG). E-mail: jeronimo@fafich.ufmg.br.

INTRODUÇÃO

A pergunta sobre raça/cor no censo brasileiro estabelece e institucionaliza categorias que podem se tornar um critério para diferenciação social, para estruturar relações raciais, definir identidades e esclarecer o que se entende popularmente por cor ou raça. A definição de fronteiras raciais também é importante para a implementação de políticas públicas que utilizam a variável raça como critério de inclusão. Regras informais de interação social observadas em posições de classe também contribuem para a codificação da raça/cor e para a atribuição de privilégios e direitos (Telles, 2004:222). O problema, entretanto, é que essas regras, normas, ideias e crenças têm importantes consequências no mundo real, sobretudo quando as categorias raciais utilizadas variam ao longo do tempo e do espaço, dependendo de ideologias, práticas raciais, mobilidade social e respostas estatais a demandas sociais e políticas (Carvalho, Wood e Andrade, 2003; Graham *et alii*, 1990; Jenkins, 2008; Nobles, 2000; Omi e Winant, 1994; Skidmore, 1993). A variável raça/cor funciona, portanto, como sinalizador, catalisador visível para a exclusão social e para o pertencimento de classe. Um sinalizador que serve tanto para implementar políticas públicas quanto para reificar distinções raciais e sustentar estereótipos sobre comportamentos, crenças e experiências de populações particulares (Noymer, Penner e Saperstein, 2011).

No fim de 2001, por exemplo, várias ações afirmativas foram criadas para combater a discriminação e aumentar a participação de negros¹ em órgãos públicos, ministérios e universidades. Mais especificamente, decretos foram estabelecidos para garantir que pelo menos 20% de todas as posições administrativas no Ministério da Agricultura e 30% do orçamento ministerial fossem alocados para comunidades majoritariamente negras. Em 9 de outubro de 2001, a Assembleia Legislativa do Rio de Janeiro anunciou que 40% das admissões nas universidades estaduais seriam reservadas para pretos e pardos. Em dezembro desse mesmo ano, o Ministério da Justiça e o Supremo Tribunal Federal ordenaram que pelo menos 20% dos consultores, funcionários e empregados fossem negros. Ações afirmativas também foram aprovadas para prover vinte bolsas anuais para estudantes negros no Instituto Rio Branco. Em 24 de março de 2002, o Ministério do Trabalho decretou que 20% do orçamento do Fundo de Amparo aos Trabalhadores (FAT) fosse direcionado a profissionais negros. Em maio de 2002, o então presidente Fernando Henrique Cardoso assinou o Programa Nacional de Ações Afirmativas, propondo a adoção de mecanismos que promovessem a melhoria de “populações em desvantagem”, mas não implementou cotas nem metas específicas (Htun, 2004; Telles, 2004:72).

Atualmente, 70 das 98 universidades públicas do Brasil adotam pelo menos um dos 13 tipos de ação afirmativa para o preenchimento de vagas². Dessas 70, aproximadamente 57% (40 universidades) adotam medidas de inclusão racial que abrangem cotas, bônus e acréscimo de vagas para negros. Essa modalidade de ação afirmativa só é menos prevalente do que a que beneficia alunos oriundos de escolas públicas – medida adotada em 59 das 70 universidades federais e estaduais do Brasil. Em termos de potencial inclusivo, entretanto, apenas 9,3% das vagas disponíveis nos vestibulares estão reservadas para candidatos pretos e pardos pelo regime de cotas fixas (Feres Júnior, Campos e Daflon, 2011:70-77). Supostamente, essas medidas seriam adotadas com base na informação sobre cor ou raça autodeclarada. Contudo, a ampla variabilidade racial e as diferentes mensurações de raça/cor fazem com que o critério baseado na autodeclaração seja, pelo menos, questionável.

A autodeclaração de raça/cor tem sido o método oficial de classificação racial no Brasil desde 1950 (Telles, 2002). Diferentemente dos Estados Unidos, onde a cor é definida por ascendência, no Brasil baseia-se predominantemente na aparência³. De fato, a Pesquisa Social Brasi-

leira (PESB, 2002) mostra que 42% dos entrevistados utilizam a cor da pele como principal critério de inclusão em dada categoria, enquanto 35% o fazem por causa da cor do pai, da mãe ou de ambos. Esse resultado é corroborado pelos dados de uma pesquisa recente realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2011), na qual 82,3% das pessoas entrevistadas em seis unidades da federação responderam que utilizam a cor da pele como principal critério para definir sua cor ou raça. Já o critério “origem familiar e antepassados” foi a opção escolhida por 47,6% dos entrevistados.

Em tese, a autoclassificação racial seria um processo mais reflexivo, envolvendo socialização pessoal, ao passo que a categorização feita por terceiros envolveria percepções de *status*, localização geográfica e noções relativas de cor (Cohen, 1994; Erikson, 1968:89). Entretanto, o sistema de classificação oficial inclui a autopercepção do entrevistado e do entrevistador. Este, muitas vezes, marca a cor do entrevistado por assumir que a resposta é óbvia, por se sentir desconfortável para realizar a pergunta ou por não considerar a pergunta tão importante (Telles, 2004). Além disso, o respondente é responsável pela declaração de cor de todos os residentes do domicílio, e essa declaração pode não corresponder à cor na qual os membros dessa casa se autoclassificariam. O sistema classificatório do censo utiliza, assim, uma combinação entre autodeclaração e percepção do entrevistador.

Os resultados produzidos por sistemas de classificação racial dependem do esquema classificatório utilizado e também de quem realiza a classificação (Bailey, 2008; Bailey e Telles, 2006; Telles e Lim, 1998). Pessoas diferentes podem classificar o mesmo indivíduo de maneiras distintas, e, mesmo que a classificação seja feita pela mesma pessoa, variações podem existir, dependendo da situação, da formulação e da interpretação da pergunta. Apesar dos significados sociais de raça/cor serem socialmente compartilhados, a inclusão em categorias predefinidas não é algo fixo. Isso é particularmente verdade no Brasil, onde a classificação racial tende a ser ambígua e volátil. A maneira como as pessoas classificam umas as outras e se autoidentificam é muitas vezes contraditória, variando de acordo com a situação social (Sansone, 2003; Simões e Jeronymo, 2007). A classificação racial no Brasil, portanto, “está longe de ser uma ciência exata” (Telles, 2004:88).

Diferentemente da variável sexo, a variável raça, além de apresentar múltiplas categorias, não possui uma fronteira bem-definida que per-

mita a identificação inequívoca dos beneficiários de cotas raciais, atualmente adotadas em quase dois quintos das universidades públicas brasileiras. Assim, políticas raciais puras ainda carecem de pontos de corte e de limites raciais capazes de definir a inclusão em uma ou em outra categoria (*ibidem*:262).

Concordância Racial no Brasil

Apesar de ser um entre vários, a autoclassificação, mais do que qualquer outro critério de categorização racial, é o mais aceito no Brasil e também em outras partes do mundo. Cientistas políticos, sociólogos e economistas, em particular, raramente questionam a origem dos dados de raça, assumindo que a variável é essencialmente constante⁴. Sociólogos concordam que raça é uma construção social permeada de inconsistências, mas raramente questionam ou criam soluções para incorporar essas ambiguidades em análises sociais e tornar a evidência racial mais informativa e (estatisticamente) confiável⁵. Pesquisas recentes, por exemplo, demonstram que a desigualdade entre renda e saúde depende das medidas unidimensionais e multidimensionais de raça empregadas (Saperstein, 2006; 2008; 2011; Saperstein e Sykes, 2008).

Pouco se sabe sobre a concordância entre classificações raciais no Brasil. Ainda não conhecemos o percentual de pessoas que se classificam (e são classificadas) em uma mesma categoria racial quando diferentes métodos de classificação, contendo as mesmas categorias, são utilizados, por exemplo. Quatro estudos investigaram como a categorização racial mudaria caso os dados de raça/cor fossem coletados utilizando metodologias distintas daquela imposta pelo IBGE. Telles e Lim (1998) mediram a desigualdade de renda entre brancos, pardos e pretos, e constataram que brancos ganham 26% a mais que pardos quando a classificação racial é feita pelo entrevistador e 17% a mais que pardos quando a cor é autodeclarada. Já a desigualdade racial entre pretos e pardos permanece praticamente inalterada: pretos ganham 13% a menos que pardos quando são classificados pelo entrevistador e 12% a menos quando declaram a própria cor. Assim, de acordo com essa pesquisa, dependendo de como se implementa a categorização racial, a desigualdade de renda entre brancos e pardos muda significativamente.

Em outro estudo, Bailey e Telles (2006) investigaram como a adoção dos termos “moreno” e “negro” alteraria a distribuição de raça/cor da população em comparação com as categorias do IBGE. Os autores mostram que 40% dos brasileiros escolhem o termo “moreno”, especialmente os mais jovens, os com menos anos de estudo e as pessoas que vivem em áreas com poucos brancos. Em particular, 13% dos brancos, 60% dos pardos e 38% dos pretos preferem se classificar na ampla categoria dos “morenos”. Os autores também demonstram que o termo “negro” é mais prevalente entre as pessoas com maior *status* socioeconômico e que se autodeclaram “pretas” na classificação oficial do censo. Mais recentemente, Bailey (2009) utilizou dados da PESB 2002 para investigar como o formato dicotômico – preto ou branco – afetaria a identificação de beneficiários de políticas raciais específicas e as trajetórias futuras de classificação. O autor mostra que aproximadamente metade dos pardos opta pela categoria “branco” quando as opções de resposta se restringem apenas aos termos preto ou branco. O formato dicotômico aumenta a participação dos brancos de 49% para 67%.

Simões e Jeronymo (2007) utilizaram os dados da região metropolitana de Belo Horizonte, coletados em 2005, para comparar medidas de classificação racial alternativas àquelas utilizadas pelo IBGE. Os autores utilizaram quatro medidas: a) autoclassificação espontânea; b) raça observada (controlada pela raça do entrevistador); c) percepção da ancestralidade (raça dos pais e avós); e d) autoclassificação e observação do fenótipo (cor da pele, tipo de cabelo). Os resultados encontrados indicaram “divergências significativas tanto internamente ao critério de autoclassificação quanto entre autoclassificação e raça observada” (*ibidem*:119).

Em 2008, o IBGE (2011) levou a campo a Pesquisa das Características Étnico-Raciais da População (PCERP). Além de mostrar que, para mais de 60% das pessoas entrevistadas, a cor ou a raça tem influência em suas vidas, a pesquisa mostra o tamanho da discordância classificatória regional existente quando a cor ou a raça são autodefinidas e também definidas pelo entrevistador. Todas as unidades da federação nas quais a PCERP foi conduzida possuem elevado grau de discordância entre os sistemas de auto e heteroclassificação raciais. Apenas 21% das pessoas autodenominadas pretas, por exemplo, são classificadas nessa mesma categoria pelos entrevistadores. No caso dos pardos, esse percentual é igual a 54,2%; e, no caso dos brancos, o percentual de concor-

dância entre os dois sistemas eleva-se para 91%, apesar de no Distrito Federal ser bem mais baixo (68,5%) (IBGE, 2011:Tabela 2.10)⁶.

Diante da ambiguidade e da multiplicidade racial brasileira, seria possível então traçar uma linha divisória entre as fronteiras definindo a cor ou a raça dos brasileiros? Seria possível definir a linha divisória entre beneficiários (não seriam beneficiados?) e membros do grupo dominante na ausência de regras claras de distinção racial? A presente investigação expande os estudos realizados por Telles, Lim, Bailey, Simões e Jeronymo ao considerar duas novas possibilidades de categorização racial nunca antes medidas no Brasil: a ascendência e a percepção racial do entrevistado segundo fotografias. Além disso, ainda não existem estudos combinando diferentes classificações raciais para estimar intervalos de confiança estatística por raça/cor para a composição populacional.

Os resultados deste artigo demonstram a extensão da ambiguidade das medidas de raça/cor no Brasil com base na construção de intervalos de confiança estatística para linhas de separação racial combinando múltiplos esquemas de classificação. A estimação de tais linhas mostra o grau de interseção entre esquemas classificatórios alternativos e gera os insumos necessários para a criação de um potencial esquema de categorização ampliado, combinando diferentes modos de percepção racial. Por fim, para ilustrar a incerteza envolvida no processo de classificação racial, apresentamos as probabilidades preditas de potenciais universitários se classificarem ou serem classificados como brancos ou pretos, considerando a cor de seus pais e outros determinantes relevantes em diferentes esquemas de classificação.

A seção seguinte descreve a base de dados utilizada. Em seguida, na terceira seção, são apresentados os resultados, que estão organizados em três partes. A primeira mostra a distribuição proporcional da população por cor, de acordo com metodologias alternativas (autoclassificação, entrevistador, fotografias e ascendência), e apresenta intervalos de confiança estatística para essas distribuições. A segunda avalia a concordância entre metodologias distintas de categorização racial quando três (branco, pardo ou preto) ou duas (branco ou preto) opções de classificação são oferecidas. A terceira parte dos resultados apresenta a probabilidade de classificação entre os termos branco ou preto segundo características individuais e o tipo de metodologia utilizada. A quarta seção discute a relevância dos resultados encontrados tendo

em vista a crescente adoção de políticas raciais específicas no Brasil. A seção seguinte encerra o artigo.

BASE DE DADOS

A análise proposta utiliza a Pesquisa Social Brasileira (PESB), realizada pela Universidade Federal Fluminense (UFF) em 2002. A pesquisa segue o modelo do General Social Survey (GSS) norte-americano e do British Social Attitudes (BSA). A PESB realizou 2.362 entrevistas com adultos acima de 18 anos em 102 municípios, cobrindo todos os estados brasileiros. Os municípios com menos de 20 mil habitantes nas regiões Norte e Nordeste foram excluídos para reduzir os custos da amostragem. Como o interesse deste artigo é analisar diferentes formas de classificação (entre branco, pardo e preto ou entre branco e não branco), as pessoas classificadas como amarelas ou indígenas (0,9% da população) foram excluídas da análise. A amostra da PESB é probabilística e representativa do território nacional.

Esta análise se vale do terceiro módulo da PESB 2002, que investiga as relações raciais e a concepção de raça/cor do brasileiro. Esse módulo é o único a captar a cor, conforme percepções individuais, usando: (1) categorias raciais predefinidas seguindo o formato utilizado pelo IBGE; (2) autoclassificação segundo fotografias⁷; (3) classificação de acordo com a percepção do entrevistador; e (4) ascendência ou cor dos pais. Nesse último caso, a cor do(a) filho(a) segue a cor do ascendente direto mais escuro. Cerca de 80% dos entrevistados tiveram sua cor coletada, em todas as metodologias investigadas, tanto no formato tricotômico quanto no dicotômico. A PESB, em particular, investiga a classificação racial considerando somente as alternativas “branco” ou “preto”. Esta pesquisa é a primeira a permitir a investigação da concordância e a estimação de intervalos de confiança para a declaração de cor segundo metodologias distintas. É, assim, particularmente útil para informar e contribuir para o debate sobre cotas e outras políticas raciais específicas no Brasil.

Resultados

Os métodos e os resultados são apresentados e discutidos em três partes: a primeira apresenta a distribuição da população por cor e raça utilizando diferentes esquemas de classificação racial. A segunda discute a concordância racial entre esses esquemas quando são oferecidas três

ou duas opções de escolha para a cor. A terceira parte apresenta as probabilidades de um indivíduo hipotético, com perfil universitário, classificar-se como branco ou preto em função de sua ascendência e do tipo de método utilizado na coleta.

Distribuição da População por Cor/Raça

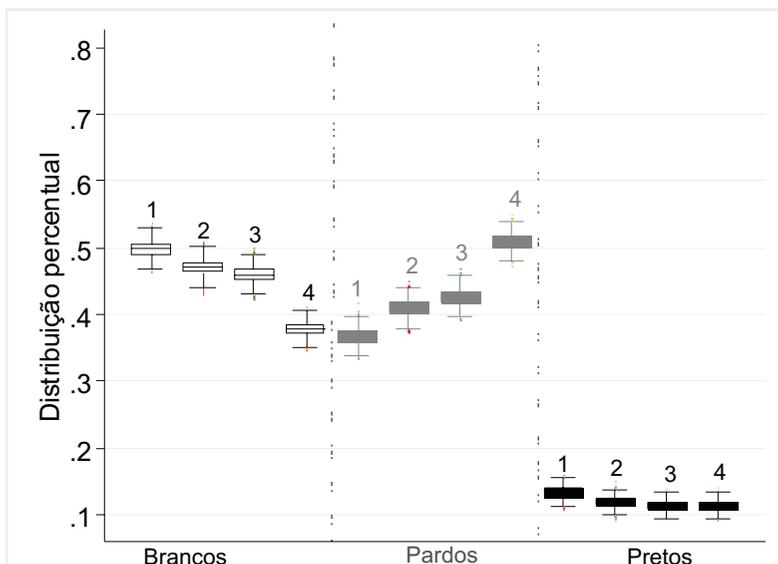
De acordo com o critério de autoclassificação do censo demográfico brasileiro de 2000, entre cada 100 pessoas havia 53 brancas, 39 pardas, 6 pretas, 0,4 amarela e 0,45 indígena (Ruggles *et alii*, 2004). Essa composição é similar à dos dados do Datafolha, coletados em 1995, quando o mesmo esquema autoclassificatório foi utilizado (Bailey e Telles, 2006; Telles e Lim, 1998). Quando apenas três opções aparecem para a autoclassificação, essa composição muda para 56% branca, 33% parda e 11% preta, segundo os dados do Datafolha. Já de acordo com a PESB, a população em 2002 era 49% branca, 39% parda e 12% preta. Alternativamente, usando um sistema de autoclassificação aberto baseado nos mesmos dados, 43% da população se classifica como branca, 15% como parda e apenas 3% como preta (Bailey, 2008). Nos dados do Datafolha, esses números são iguais a 42%, 6% e 5% para brancos, pardos e pretos, respectivamente (Bailey e Telles, 2006). Quando a população é classificada pelo entrevistador, a composição racial muda para 56% branca, 31% parda e 13% preta (Telles e Lim, 1998).

Assim, vê-se que a distribuição da população por cor varia segundo a metodologia utilizada na coleta, as opções disponíveis para a escolha dos entrevistados (ou entrevistador), o período de coleta, a base de dados e a amostra utilizadas. É preciso, portanto, controlar esses fatores de variação para que se tenha um efeito puro do modo de coleta da informação sobre cor. A Figura 1 faz isso ao apresentar os intervalos percentuais de brancos, pardos e pretos em cada uma das metodologias de coleta da PESB 2002⁸.

Essa figura mostra que, no esquema utilizado pelo IBGE (método 1), a maioria da população (entre 47% e 54%) se autoclassifica branca, enquanto uma proporção entre 33% e 40% se autoclassifica parda, e entre 11% e 16% preta. A informação sobre cor coletada segundo a autoclassificação em fotografias (método 2) e segundo a opinião do entrevistador (método 3) possui valores percentuais similares em cada categoria de cor. De modo geral, os entrevistadores tendem a subestimar o número de brancos e a superestimar o de pardos em relação à cor autode-

Figura 1

Intervalos de Confiança para a Distribuição Percentual da População entre os Termos Branco, Pardo e Preto segundo a Metodologia de Coleta – Brasil (2002)



Elaboração do autor com base na PESB 2002.

Nota: 1 = IBGE; 2 = fotografia; 3 = entrevistador; 4 = ascendência. Intervalos de confiança estatística de 95% possuem distribuição aproximadamente normal

clarada. Já a autodeclaração a partir de fotografias tende a apresentar números intermediários se comparada às outras duas metodologias. Nos métodos de coleta 2 e 3, o intervalo de confiança para o percentual de brancos varia entre 44% e 50%; o de pardos, entre 39% e 46%; e o de negros, entre 9% e 14%. Portanto, nas metodologias 2 e 3, o percentual de brancos é ligeiramente menor, e o de pardos ligeiramente maior do que no caso da metodologia empregada pelo IBGE (1). A diferença mais significativa ocorre na comparação por ascendência (4). Quando os entrevistados são forçados a adotar a cor do ascendente direto mais escuro, tal como é feito no sistema de classificação americano, a maioria da população se torna parda (entre 47% e 54%), o que reflete não apenas a elevada proporção de casais racialmente heterogâmicos e com filhos na população do Brasil (43%) como também a preferência por cores mais claras por parte desses casais, já que, quando a regra baseada na ascendência não é adotada, a maioria se torna branca. A Figura 1 também mostra a estabilidade percentual da categoria dos pretos. Percebe-se que a adoção de um ou outro método tende a afetar mais a

classificação entre branco ou pardo do que entre pardo ou preto. Independentemente da metodologia utilizada, a proporção de pretos permanece estável, com intervalos de confiança variando entre 10% e 15%. Isso sugere que as principais trocas reclassificadoras advindas de mudanças na coleta de dados ocorrem entre pardos e brancos, enquanto pretos apresentam maior certeza quanto sua cor. Contudo, será que a proporção de pretos permaneceria a mesma se o entrevistador e os entrevistados fossem obrigados a escolher sem a alternativa pardo como opção?

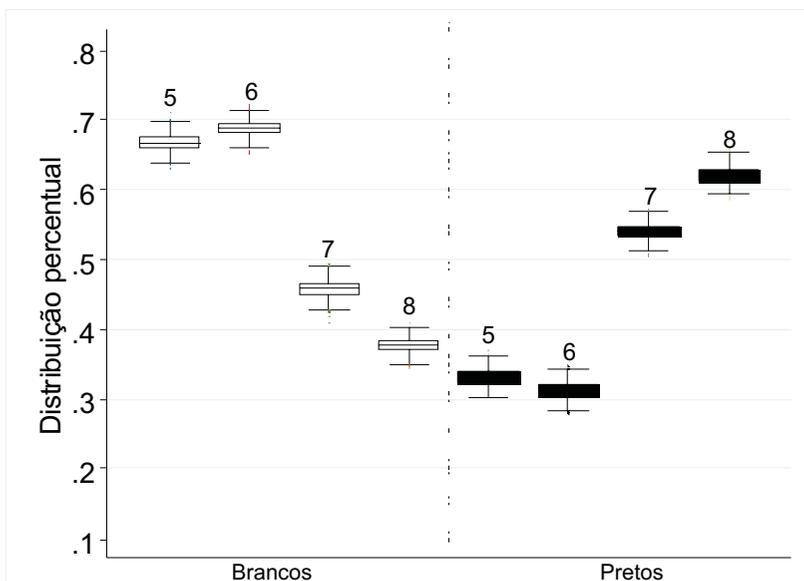
A Figura 2 responde a essa pergunta apresentando a distribuição da população por cor quando as opções de escolha se restringem apenas aos termos branco ou preto. Mostra que, caso aqueles que se classificaram como pardos no esquema do IBGE fossem submetidos a uma escolha dicotômica, a proporção de brancos se elevaria para um valor entre 64% e 70%, o que corresponde a um aumento de 17 pontos percentuais em relação à metodologia tricotômica. Essa proporção é ligeiramente superior no esquema dicotômico com fotografias (68% e 71%). Nos esquemas 7 e 8, somente a proporção de pretos se altera, já que representa a soma simples do número de pretos e pardos encontrados nos esquemas 3 e 4. Essa lógica de agrupamento (pardos + pretos) é similar àquela defendida por ativistas do movimento negro e utilizada por certas instituições como critério de inserção e de concessão de benefícios. O problema de agrupar pretos e pardos de forma não espontânea em uma única categoria – denominada “negros” – é que o agrupamento forçado se mostra inconsistente ao que seria observado caso os entrevistados escolhessem livremente pela cor branca ou preta⁹. Essa inconsistência fica clara quando se compara a proporção de autodeclarados pretos nas formas de coleta 5 e 6 com a proporção forçada de pretos nas metodologias 7 e 8. Quando a escolha pela categoria “preto” ocorre por autodeclaração (5) e via fotografias (6), a população branca representa a maioria. Quando a proporção de pretos é calculada por agrupamento (pardos + pretos), os brancos se tornam a minoria da população.

CONCORDÂNCIA E RECLASSIFICAÇÃO RACIAL

Esta seção apresenta a concordância entre modos de coleta da informação racial no Brasil segundo dois esquemas categóricos. No primeiro modo de classificação, utilizam-se três opções: branco, pardo ou preto; no segundo, apenas duas: branco ou preto. O reagrupamento induzido em apenas duas categorias permite investigar o tamanho da “válvula

Figura 2

Intervalos de Confiança para a Distribuição Percentual da População entre os Termos Branco e Preto/Não Branco segundo a Metodologia de Coleta – Brasil



Elaboração do autor com base na PESB 2002.

Nota: 5 = IBGE; 6 = fotografia; 7 = entrevistador; 8 = ascendência. Intervalos de confiança estatística de 95% possuem distribuição aproximadamente normal.

de escape” (Degler, 1986) do grupo racial intermediário em diferentes metodologias de classificação. Em particular, o esquema dicotômico permite investigar se a população de fato se divide tal como é assumido pelo Estado, por alguns acadêmicos e por movimentos institucionais: brancos de um lado, pardos e pretos de outro.

Consistência no Formato Tricotômico

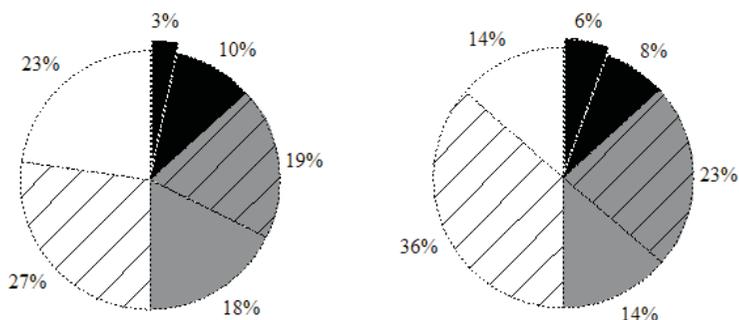
A Figura 3 ilustra a proporção de indivíduos classificados entre as cores branca, parda e preta segundo a metodologia do IBGE e ilustra ainda a proporção consistentemente classificada (ou que se classifica) em todas as metodologias. Metade (27+19+3) dos entrevistados se classifica ou é classificada da mesma maneira, isto é, de modo independente da metodologia utilizada na coleta ou na construção da variável raça/cor. Mais precisamente, 54% [27/(27+23)] das pessoas que se classificam como brancas no formato do IBGE, 51% [19/(19+18)] das que se classificam como pardas e apenas 23% (3/13) das que se classificam

Figura 3

Consistência Classificatória e Distribuição Racial segundo a Metodologia do IBGE

Consistência entre metodologias 1, 2, 3 e 4

Consistência entre metodologias 1, 2 e 3



N = 1.911

Elaboração do autor com base na PESB 2002.

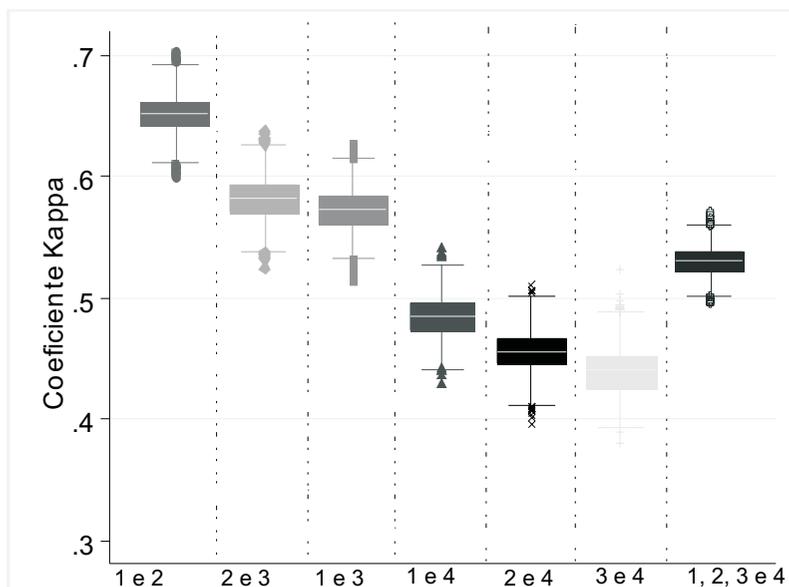
Nota: As fatias brancas (mais as hachuradas da mesma cor) correspondem ao percentual de brancos (52%); as fatias cinza (mais as hachuradas da mesma cor), ao percentual de pardos (36%); e as fatias pretas (mais a preta destacada), ao percentual de pretos (12%). As fatias hachuradas e a fatia preta destacada representam o percentual cuja cor foi declarada de maneira consistente entre metodologias.

como pretas mantêm a mesma cor em todas as outras metodologias utilizadas; seja ela baseada na autodeclaração, em fotografias, na opinião do entrevistador, seja na ascendência. Ignorando-se a metodologia baseada na ascendência (4), a consistência sobe para 65% (23+36+6). Isso representa 72% dos brancos, 43% dos pretos e 62% dos pardos que se classificam no formato tricotômico do IBGE.

Para avaliar a concordância entre os métodos de coleta da variável cor, uma alternativa interessante é a estatística Kappa, elaborada por Cohen (1960). A Figura 4 apresenta intervalos estatísticos de 95% de confiança para a concordância entre pares de métodos de coleta. Segundo Landis e Koch (1977), o grau de concordância não aleatória entre observadores ou medidas repetidas da mesma variável categórica pode ser interpretado de acordo com uma escala de seis níveis: valores abaixo de zero indicam concordância ruim; 0 a 0,2, concordâncias baixíssimas; 0,21 a 0,4, concordâncias baixas; 0,41 a 0,6, concordâncias moderadas; 0,61 a 0,8, boa concordância; 0,81 ou mais, concordância quase perfeita¹⁰.

A Figura 4 mostra que a pior concordância ocorre entre a metodologia baseada na ascendência (4) e as demais, com o coeficiente Kappa va-

Figura 4
Concordância entre Metodologias Tricotômicas de Classificação Racial segundo o Coeficiente Kappa de Cohen



Elaboração do autor com base na PESB 2002.

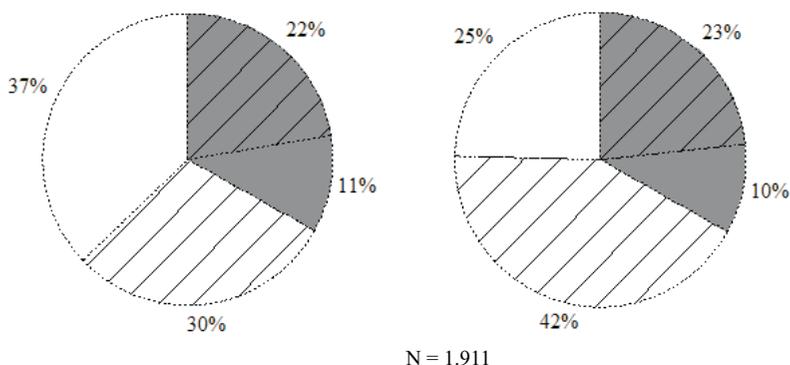
Nota: 1 = IBGE; 2 = fotografia; 3 = entrevistador; 4 = ascendência. Intervalos de confiança estatística de 95% possuem distribuição aproximadamente normal.

riando entre 0,39 e 0,54. Já a maior concordância ocorre entre a metodologia do IBGE e a baseada em fotografias. De acordo com a escala sugerida por Landis e Koch (1977), essas são as únicas medidas que apresentam boa concordância, entre 0,61 e 0,69. As outras comparações apresentam concordância apenas moderada¹¹.

Consistência no Formato Dicotômico

Quando são oferecidas apenas as opções “preto” ou “branco”, a consistência da medida racial aumenta para 52%, dos quais 30% são brancos e 22% pretos. Quando a opção baseada na ascendência (8) não é considerada, a concordância aumenta para 65%. Desses, 42% se classificam consistentemente como brancos, enquanto 23% se classificam em todas as metodologias como pretos. Brancos, portanto, apresentam maior consistência taxonômica em sua medida de cor do que os pretos (Figura 5).

Figura 5
Consistência Classificatória e Distribuição Racial entre Brancos e Pretos
segundo a Metodologia 5
 Consistência entre metodologias 5, 6, 7 e 8 Consistência entre metodologias 5, 6 e 7

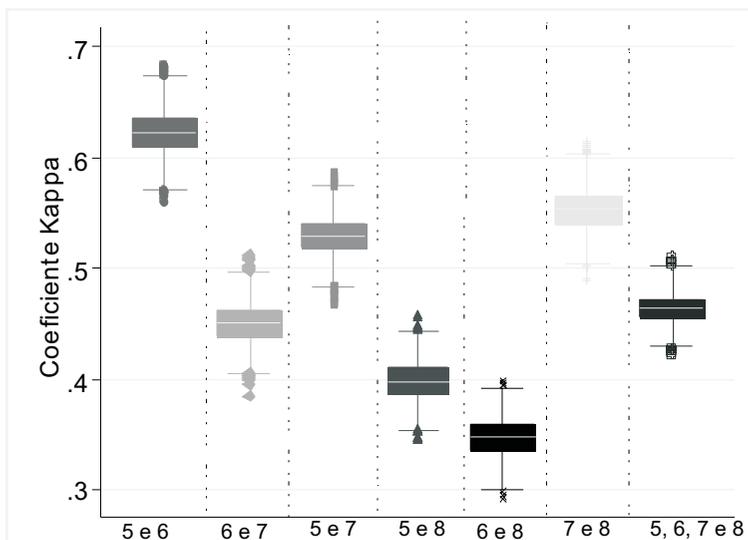


Elaboração do autor com base na PESB 2002.

Nota: As fatias brancas (mais as hachuradas da mesma cor) correspondem ao percentual de brancos (67%); as fatias cinza (mais as hachuradas da mesma cor), ao percentual de pretos (33%). As fatias hachuradas representam o percentual cuja cor foi declarada de maneira consistente entre metodologias.

A comparação entre as Figuras 3 e 5 revela a porcentagem de pardos que se reclassificaram como pretos ou brancos quando “forçados” a escolher entre os dois termos. Dos 37% que se classificaram como pardos no formato tricotômico, quando puderam escolher apenas entre duas opções, 20% se reclassificaram como pretos e 17% como brancos. Ou seja, uma vez que o indivíduo se autodenominou pardo, em uma autorreclassificação, há 17% de probabilidade de ele vir a se declarar branco e 20% de passar a se declarar preto. Caso os termos preto e pardo fossem substituídos pela categoria negros, a maioria dos pardos ($20/37 = 54\%$) migraria para a categoria mais escura. Se calcularmos as probabilidades compostas de um indivíduo ser pardo no esquema tricotômico e depois se declarar branco ou preto, os números são: 7% ($0,37 \times 0,2$) de probabilidade de ser pardo e de se reclassificar como preto, e 6% ($0,37 \times 0,17$) de probabilidade de ser pardo e depois se reclassificar como branco. A evidência mostra, portanto, que pretos e pardos pertencem a categorias distintas, uma vez que praticamente a metade dos pardos ($17/37 = 46\%$) se considera branca no formato dicotômico. A válvula de escape dos pardos parece, assim, “vazar” de maneira homogênea quando há apenas duas opções de classificação, confir-

Figura 6
Concordância entre Metodologias Dicotômicas de Classificação Racial segundo o Coeficiente Kappa de Cohen



Elaboração do autor com base na PESB 2002.

Nota: 5 = IBGE; 6 = fotografia; 7 = entrevistador; 8 = ascendência. Intervalos de confiança estatística de 95% possuem distribuição aproximadamente normal.

mando, desse modo, a evidência encontrada por Bailey (2009). Agrupar pardos e pretos em um único grupo violaria, portanto, identidades raciais individuais que, caso fossem escolhidas de maneira espontânea, seriam construídas de forma distinta.

Em termos de concordância entre as quatro medidas de classificação racial, a Figura 6 mostra que apenas os métodos baseados na reclassificação de pardos e a autoclassificação baseada em fotografias (métodos 5 e 6) concordam satisfatoriamente, apresentando valores de Kappa superiores a 0,6. Todas as demais comparações apresentam concordância baixa ou moderada, com valores de Kappa entre 0,3 e 0,6.

Nota-se também que ter apenas duas opções de escolha racial diminui a concordância entre os métodos de coleta. Todos os intervalos de confiança apresentados na Figura 6 mostram valores abaixo daqueles observados quando três opções de cor são oferecidas (Figura 4). Em relação às metodologias que utilizam categorias de cor tricotômicas, há aumento na concordância apenas entre o método de classificação pelo entrevistador (7) e o baseado na ascendência (8). Isso ocorre porque o

grupo de pretos, tanto em 7 quanto em 8, é construído simplesmente a partir da agregação de pardos e pretos advindos dos esquemas 4 e 5. Por causa desse modo forçado de construção, a concordância entre o número de pessoas que se classificam como brancas e não brancas (pretas + pardas) tende a ser mais parecida no esquema dicotômico (entre 7 e 8) do que no esquema tricotômico (entre 3 e 4).

Branco ou Preto?

A fim de ilustrar a incerteza racial classificatória, modelos logísticos binomiais foram construídos para determinar as probabilidades de ser branco ou preto quando se consideram variáveis individuais e contextuais. O modelo de regressão logístico binomial¹² controla a estimativa de probabilidades pelas seguintes variáveis explicativas: idade, escolaridade, estado civil, região (Nordeste), área de residência (capital), cor do entrevistador segundo ele próprio e segundo o entrevistado, padrão relativo de consumo¹³ e cor dos pais declarada pelo entrevistado (branco, pardo ou preto).

A média (ou o percentual) de algumas das variáveis explicativas utilizadas no modelo encontra-se na Tabela 1. A tabela revela o *status* superior dos brancos em termos educacionais, de consumo e regionais. Enquanto 21,2% dos brancos classificados pelo entrevistador (método 7) possuem ensino superior completo, apenas 5,12% dos classificados como não brancos (pretos + pardos) possuem educação superior. Além disso, apenas 14,48% dos brancos residem na região Nordeste, enquanto entre os não brancos a cifra é igual a 39,78%. Cerca de 30% dos não brancos e 25% dos brancos possuem residência em centros estaduais. Esses percentuais são similares para aqueles que se autot classificam (métodos 5 e 6).

A vantagem de utilizar modelos baseados em características individuais é que, além de mensurar a associação existente entre o evento de interesse e as variáveis relevantes, eles também permitem a construção de cenários contrafactuais que “padronizam” a análise comparativa. A padronização consiste em assumir que os atributos individuais permanecem os mesmos, enquanto aqueles cuja influência se deseja avaliar variam conforme o interesse do pesquisador. Cenários contrafactuais permitem perguntar, por exemplo, qual seria a probabilidade de ser classificado como branco ou preto assumindo que existem diferenças apenas quanto ao tipo de método utilizado na coleta da informação

Tabela 1
Distribuição Média e Percentual das Variáveis Explicativas de Acordo com a Cor e com o Método de Mensuração

Variável	Branco			Preto ou Não Branco		
	Método 5	Método 6	Método 7	Método 5	Método 6	Método 7
Idade média	40,85	40,45	41,36	37,78	38,55	38,51
Escolaridade (%)						
analfabeto	6,2	6,25	5,19	10,88	11,03	9,98
até a 4ª série	21,29	21,6	17,15	28,01	27,67	29,21
de 5ª a 8ª série	21,75	22,51	20,72	26,75	25,25	25,74
2º grau completo superior ou mais	34,4	33,7	35,73	29,09	30,44	29,94
superior ou mais	16,35	15,94	21,22	5,27	5,61	5,12
Padrão de consumo médio	0,133	0,130	0,152	0,074	0,079	0,080
Região Nordeste (%)	21,67	23,89	14,48	40,508	36,379	39,775
Residência na capital (%)	25,29	25,05	25,11	30,777	31,634	28,797
Estado conjugal: casado (%)	47,28	47,99	50,14	41,630	39,651	41,195
Cor da mãe (%)						
branca	76,01	74,06	84,69	28,08	29,91	38,8
parda	21,71	23,23	13,58	51,62	49,79	47,44
preta	2,28	2,72	1,74	20,3	20,3	13,76
N (não ponderado)	1313	1345	879	652	620	1086

Elaboração do autor com base na PESB 2002.

sobre cor e quanto à cor dos pais¹⁴. As probabilidades preditas de ser classificado como branco ou preto segundo a cor dos pais e o método de coleta encontram-se na Tabela 2¹⁵.

Nesse sentido, a Tabela 2 apresenta duas informações relevantes. A primeira diz respeito à similaridade entre as probabilidades preditas pelos métodos 5 e 6. As probabilidades de potenciais universitários se autotclassificarem brancos ou pretos, via fotografias ou não, são basicamente as mesmas. Em contrapartida, quando os entrevistados são classificados como brancos ou não brancos (pretos + pardos) pelo entrevistador (método 7), as probabilidades de serem brancos são evidentemente menores do que quando os entrevistadores escolhem “livremente” entre os termos branco ou preto.

A segunda informação importante apresentada pela Tabela 2 diz respeito às probabilidades de classificação dos descendentes de casais homogâmicos e heterogâmicos. Para os casais homogâmicos brancos e pretos, não há surpresa em observar que a maioria de seus descendentes se classifica ou é classificada como, respectivamente, branca e preta. O interessante é constatar que, quando os membros do casal são pardos, a probabilidade de eles se autotclassificarem brancos ou pretos

Tabela 2
Probabilidades Preditas de Classificação Racial de Acordo com a Cor do Pai e da Mãe e a Metodologia de Coleta

Cor do(a)		Probabilidade predita de ser classificado como...					
		Branco			Preto ou Não Branco		
Pai	Mãe	Método 5	Método 6	Método 7	Método 5	Método 6	Método 7
branco	branca	0,97	0,95	0,86	0,03	0,05	0,14
branco	parda	0,76	0,75	0,56	0,24	0,25	0,44
branco	preta	0,41	0,41	0,24	0,59	0,59	0,76
pardo	branca	0,78	0,73	0,53	0,22	0,27	0,47
pardo	parda	0,51	0,50	0,19	0,49	0,50	0,81
pardo	preta	0,17	0,26	0,05	0,83	0,74	0,95
preto	branca	0,47	0,47	0,15	0,53	0,53	0,85
preto	parda	0,22	0,25	0,03	0,78	0,75	0,97
preto	preta	0,07	0,13	0,01	0,93	0,87	0,99
Probabilidade predita média		0,67	0,68	0,45	0,33	0,32	0,55

Elaboração do autor com base na PESB 2002.

Nota 1: As demais características individuais foram fixadas a fim de refletir um “perfil universitário” e as características da maioria da população amostral na predição das probabilidades apresentadas. Exemplo: idade entre 18 e 24 anos, segundo grau completo, solteiro, vivendo fora da região Nordeste etc.

Nota 2: Intervalos de confiança de 95% para as probabilidades preditas encontram-se no Anexo 2.

é a mesma, variando em torno de 50%, com um intervalo de confiança de 95% entre 38% e 61%. No caso de casais heterogâmicos, a tabela mostra que, quando um é branco e o outro é pardo, a probabilidade de o potencial universitário se classificar como branco (em torno de 75%) é expressivamente maior do que a de se classificar como preto (em torno de 25%). No entanto, quando um dos membros do casal é pardo e o outro é preto, a probabilidade de o filho se classificar como preto se torna maior do que a de se classificar como branco. No método 5, essa probabilidade é de 83% no caso de a mãe ser preta e de 78% no caso de o pai ser preto e a mãe ser parda. No método 6, essas probabilidades são ligeiramente menores. Por fim, quando um dos pais é branco e o outro é preto, a probabilidade de ser classificado como preto é maior do que como branco. Estudos passados mostram que, caso a opção pardo fosse oferecida, certamente seria a mais prevalente entre os descendentes de casais branco/preto (Schwartzman, 2007). Entretanto, a Tabela 2 mostra que, quando apenas as opções branco ou preto são oferecidas, a probabilidade média de potenciais universitários se autotransformarem pretos se aproxima de 60% quando o pai é branco e a mãe é preta.

DISCUSSÃO

Os resultados apresentados põem em foco a confiabilidade das classificações raciais ao mostrar a inconsistência existente entre métodos baseados na autoclassificação, na opinião do entrevistador e na ascendência. Contudo, a inconsistência demonstrada não invalida a autodeclaração, uma vez que esta já está estabelecida como sistema dominante de classificação racial e sua reprodução permite a comparação de dados ao longo dos anos. A autodeclaração é ainda a única forma de não violar identidades, respeitar preferências e continuar permitindo que a cor seja “socialmente construída” e leve em conta grande parte das complexidades individuais, coletivas e circunstanciais envolvidas na construção da raça/cor. Enquanto demarcador de identidades, a autodeclaração também não apresenta nenhum viés ou incongruência. Todavia, quando utilizada, brancos se tornam a maioria. Além disso, outra dificuldade aparece quando a cor autodeclarada passa a ser vislumbrada como sinalizador para alocação de benefícios. Afinal, se a identificação racial for dependente da obtenção de privilégios, o critério cor da pele torna-se ineficaz para viabilizar a inclusão de categorias sociais excluídas de serviços e direitos. Há evidência, por exemplo, demonstrando que as cotas raciais serviram de incentivo para reforçar a identidade negra (Francis e Tannuri-Pianto, 2012; no prelo). Quando isso ocorre, é legítimo que outras medidas potencialmente menos enviesadas pela concessão de privilégios passem a ser cogitadas como opção. Daí a importância de se averiguar a compatibilidade entre medidas de cor autodeclaradas e aquelas construídas a partir de pressupostos distintos, como o que assume que pretos e pardos poderiam ser agrupados em uma única categoria por possuírem “características similares”. Entretanto, como demonstrado, a aceitação desse pressuposto criaria outro impasse: uma distribuição populacional por cor de maioria negra (tanto com base no agrupamento entre pretos e pardos segundo a classificação do entrevistador quanto pela ascendência).

Assim, em contextos nos quais a autodeclaração direta se apresenta como um indicador enviesado pela concessão de privilégios, é importante que se proponham outras medidas que respeitem as preferências individuais e que sejam concomitantemente neutras à oferta de privilégios. Nesse sentido, a autodeclaração intermediada por fotos pode vir a ser uma opção robusta por duas razões. Primeiro porque a autodeclaração direta (métodos 1 e 5) e a realizada via fotografias (métodos

2 e 6) possuem os maiores níveis de concordância. Isso significa que, caso a autodeclaração direta viesse a sofrer perturbações – por exemplo, por causa da concessão de benefícios vinculados à cor declarada –, a metodologia baseada em fotografias poderia ser utilizada como representativa desde que também não fosse enviesada pelas mesmas perturbações. A razão para acreditar que o procedimento baseado em fotos não seria igualmente afetado está no fato de se basear em duas etapas em vez de apenas em uma. Na primeira etapa, faz-se a classificação racial de terceiros, representados em fotografias; na segunda, o entrevistado escolhe a foto que se parece com ele. Melhor ainda se as duas etapas do processo de autoclassificação forem feitas em momentos afastados temporalmente, pois o viés da cor declarada gerado pela concessão do potencial benefício certamente seria menor. Atualmente, do total de 45 universidades que adotam procedimentos de definição racial para estabelecer ações afirmativas raciais, três utilizam fotografias e quatro utilizam comissões de verificação para a homologação dos candidatos (Feres Júnior, Campos e Daflon, 2011:68).

Outro ponto a ser discutido diz respeito ao uso da cor como demarcador de diferenças de escolaridade, renda, saúde, mortalidade, fecundidade ou qualquer outra variável socioeconômica que se deseje estudar. O problema de se utilizar a cor da pele como demarcador socioeconômico é que ela não é um indicador estável e varia no tempo e de acordo com a metodologia utilizada na coleta. Sendo assim, a adoção da cor para a construção de grupos e de classes de comparação é potencialmente problemática por induzir a resultados que dependem, em muito, de como a variável raça é construída ou entendida/interpretada. A partir do momento em que uma ou outra forma de pertencimento racial passa a definir grupos heterogêneos, os resultados de interesse também o serão.

Telles e Lim (1998), por exemplo, mostraram a volatilidade da desigualdade de renda ao compararem a renda de pretos e de brancos autoclassificados com a renda de pretos e brancos classificados segundo o entrevistador. Loveman, Muniz e Bailey (2011) também chegaram a conclusões semelhantes ao compararem os níveis de desigualdade de renda entre brancos e pretos a partir de metodologias distintas de construção racial. Esses estudos demonstram que o entendimento de desigualdade racial depende, antes de tudo, de como e de quem faz parte dos grupos de pretos e de brancos. Ser classificado de um jeito ou de outro pode ter algum impacto na variável de interesse, mas, à me-

didada que o entendimento do que vem a ser raça ou cor se altera, também mudam a análise realizada e os resultados previamente alcançados (Zuberi, 2008:133). Por isso, aqueles que utilizam a variável raça/cor em suas pesquisas, diferentemente dos que a estudam, devem ter cuidado ao tratá-la como característica fixa e imutável. A variável cor ou raça não é um fato social preexistente, mas sim um epifenômeno que pode ser “criado, esvaziado, transformado e destruído” (Omi e Winant, 1994:55). É preciso percebê-la como uma característica dinâmica, complexa e multifacetada da população em vez de tomá-la como neutra e dada. Quiçá, tal como sugerido por Zuberi e Bonilla-Silva (2008:329), deveria dar-se menos atenção ao infamado “efeito raça” e mais atenção a como e por que o processo de estratificação racial produz resultados diferentes entre grupos raciais. Dessa maneira, a variável raça/cor se tornaria não uma característica descritiva – e erroneamente explicativa –, mas parte da solução do problema.

CONCLUSÃO

O objetivo desta pesquisa foi contrastar o que entendemos por cor/raça a partir da comparação entre diferentes métodos de coleta que consideram três ou duas opções de cor. O primeiro método baseia-se na autodeclaração direta, o segundo mede a cor autodeclarada em dois estágios por meio de fotografias, o terceiro utiliza a percepção do entrevistador, e o quarto atribui ao entrevistado a cor do ascendente direto mais escuro, imitando assim a regra de classificação norte-americana.

Em termos de composição populacional, as maiores diferenças são observadas quando se compara a metodologia baseada na cor dos pais com as demais. A mensuração pautada na ascendência é de longe a que produz maior proporção de pardos e negros (pretos + pardos) na população. Em contrapartida, quando é dada aos pardos a opção de se auto-classificarem brancos ou pretos, metade se considera branca, e metade preta. Após a adesão dos pardos, a população branca passa de 50% para 67%.

No que tange à consistência, os resultados mostram que, caso diferentes métodos de coleta fossem utilizados para construir a variável cor, apenas metade das pessoas teria a mesma cor em todas elas; e, se a consistência entre metodologias não levasse em conta a classificação baseada na ascendência, a consistência aumentaria para em torno de

65%. Em particular, as autotransclassificações raciais diretas e aquelas feitas por meio de fotografias são muito parecidas. Sendo assim, o método por fotos apresenta grande potencial para substituir a autodeclaração caso a credibilidade desta venha a ser comprometida ou enviesada pela concessão de benefícios.

Por fim, realizou-se também um exercício de simulação para calcular as probabilidades de potenciais universitários se classificarem como brancos ou pretos de acordo com a cor dos pais e o tipo de metodologia de coleta. Os resultados revelaram que os descendentes de casais heterogâmicos, nos quais pelo menos um dos membros é preto, são mais propensos a se classificarem como pretos, mesmo quando o outro membro é branco. Quando um dos pais é pardo e o outro é branco, há maior propensão de os descendentes se classificarem como brancos. Quando ambos os pais são pardos, a probabilidade de se autodeclararem brancos ou pretos é igual.

Chega-se ao fim deste artigo, portanto, considerando que é preciso ir além da discussão sobre os condicionais e os determinantes da raça e ultrapassar as controvérsias metodológicas envolvidas na mensuração da cor humana. Nesse sentido, uma sugestão seria combinar as diferentes formas de mensurar a raça em uma única variável latente ou, alternativamente, utilizar várias metodologias taxonômicas para descompactar a complexidade do que consideramos ser a cor da pele. A coleta de múltiplas medidas de classificação racial, em particular, ajudaria pesquisadores a distinguirem e testarem diferentes hipóteses sobre as causas das disparidades raciais e a gerar novas teorias para explicar por que elas persistem (Saperstein, 2011).

O intuito deste trabalho é ainda alertar para que a variável cor não seja usada como fator causal, mas sim como demarcador imperfeito apresentado com ressalvas. Como sistemas de classificação racial representam a contabilidade oficial e sistêmica de distinções sociais salientes, a variável “cor” precisa ser utilizada considerando-se a incerteza envolvida em sua mensuração. Dados étnicos e raciais precisam ser tratados com cuidado por causa das circunstâncias diversas sob as quais as informações são coletadas ou registradas, sobretudo no caso brasileiro, em que o uso de descrições étnicas ou raciais reflete a “aparência” e os costumes sociais em vez de origens hereditárias ou genéticas, como já dizia Oracy Nogueira. Assim, a variável cor deve ser criticamente estu-

dada, e não passivamente tomada como dada para que entendamos por que grupos raciais incertos apresentam resultados distintos.

(Recebido para publicação em novembro de 2011)

(Reapresentado em março de 2012)

(Aprovado para publicação em abril de 2012)

NOTAS

1. Na maioria dos decretos ministeriais, o termo “negro” foi utilizado. A inclusão de “pardos” no termo “negros” não é clara e supostamente foi deixada a critério de autoidentificação.
2. As ações afirmativas hoje vigentes incluem a reserva de vagas para indígenas, deficientes, professores da rede pública, quilombolas, nativos do Estado, pessoas de baixa renda, mulheres, filhos de policiais, bombeiros e refugiados políticos. Ver Feres Júnior, Campos e Daflon (2011:70) para mais detalhes sobre essa distribuição.
3. Uma revisão completa sobre o uso da variável raça/cor em estudos quantitativos, sobretudo nos Estados Unidos, pode ser encontrada em Muniz (2010).
4. Veja Carvalho, Wood e Andrade (2003) para uma exceção. Esses autores sugerem uma metodologia original para mensurar a reclassificação racial utilizando técnicas de projeção demográfica.
5. Intervalos de confiança estatística são utilizados para descrever a estabilidade e a consistência de estimativas amostrais. Intervalos de confiança descrevem a margem de variabilidade dos resultados amostrais. A PESB 2002 mostra que, de acordo com o sistema autotransclassificatório, 12% da população é preta. Um intervalo de 95% de confiança estatística para essa proporção varia entre 10,5% e 13,4%. *Ceteris paribus*, intervalos de confiança mais estreitos são mais confiáveis do que intervalos mais largos.
6. Levando em conta que, semanticamente, os termos “pardo” e “moreno” significam algo muito similar, as discrepâncias classificatórias entre auto e heterotransclassificação tornam-se menores. Considerando a amostra do Brasil, pela autotransclassificação, para cada 100 pessoas haveria 49 brancas, 22 morenas, 14 pardas, 7,8 negras e 1,4 preta. Já pelo sistema de heterotransclassificação haveria 56 brancas, 9 morenas, 22 pardas, 8,4 negras e 1 preta. Restringindo-se a amostra a regiões específicas do país, a discrepância classificatória semântica entre pardos e morenos torna-se ainda mais evidente: na Paraíba, por exemplo, a proporção de pessoas que se autodenominam “morenas” é igual a 46%; e a de “pardas”, a 11%. Quando a cor é definida pelo entrevistador, a proporção de morenos diminui para 4%, enquanto a de pardos aumenta para 45%.
7. Essa variável foi construída em duas etapas. Primeiro perguntou-se aos entrevistados qual era a cor (branco, pardo ou preto e branco ou preto) dos indivíduos em oito fotografias distintas (ver Anexo 1). Em seguida, pediu-se aos entrevistados que iden-

tificassem a fotografia que possuía um indivíduo com a cor mais parecida com sua própria cor. (As fotos em cores estão disponíveis na versão eletrônica. Ver <http://scielo.br/dados>.)

8. Intervalos de confiança de 95% para a proporção de brancos, pardos e pretos (esquema tricotômico) e para a proporção de brancos e pretos (esquema dicotômico) foram construídos utilizando-se a técnica *bootstrap* a fim de garantir que as proporções obtidas não fossem um mero resultado de variações amostrais e de seletividade. Em seguida, os desvios padrão e as proporções médias encontradas no processo *bootstrap* foram utilizados para construir intervalos de confiança aproximadamente normais. O método *bootstrap* produz medidas estatísticas em diferentes subamostras com reposição da amostra original utilizada. Esse método não paramétrico fornece mais robustez para os desvios padrão e os intervalos de confiança estimados. O teorema do limite central garante que, após um número suficientemente grande de repetições, os intervalos de confiança estimados terão uma distribuição normal.
9. Há estudos mostrando que, se o termo “negro” fosse utilizado em vez do termo “preto” como uma das opções dicotômicas, o resultado não se alteraria de forma relevante (Bailey, 2008; Telles, 2004).
10. Intervalos de confiança estatística para concordância entre diferentes esquemas de classificação racial foram construídos utilizando-se a técnica *bootstrap*. Nesse caso, em vez da proporção racial, a estatística Kappa é a variável de interesse. A concordância entre os quatro tipos de categorização racial é medida pelo cruzamento de variáveis utilizando tabelas de contingência e pela estatística Kappa. Tal estatística é utilizada para mensurar o grau de concordância entre medidas alternativas e mutuamente excludentes de classificação racial. A equação que define o Kappa de Cohen é:
$$\kappa = \frac{\text{Pr}(a) - \text{Pr}(e)}{1 - \text{Pr}(e)}$$
, sendo Pr(a) o percentual de concordância observada entre diferentes medidas e Pr(e) a probabilidade de concordância aleatória. A Kappa mede quanto a concordância observada excede o esperado caso a escolha racial fosse apenas um resultado de chance. Quando a concordância entre medidas raciais é máxima, Kappa é igual a 1. Em contrapartida, quando as medidas são totalmente distintas, Kappa é igual a ou menor do que zero (Gordis, 2004). Isso ocorreria se, por exemplo, todas as pessoas que se declararam brancas fossem classificadas como pardas ou pretas pelo entrevistador.
11. A estatística Kappa para a concordância entre a cor do entrevistador autodeclarada e a cor do entrevistador declarada pelo entrevistado é apenas moderada, variando dentro de um intervalo de confiança entre 0,44 e 0,51.
12. Regressões binomiais são apropriadas para a modelagem de variáveis categóricas e para “investigar como variáveis explicativas influenciam a probabilidade de ocorrência do evento de interesse” (Long e Freese, 2003). O modelo logístico é apropriado para mensurar a verossimilhança de classificações raciais e tem sido utilizado para estudar o caso brasileiro (Bailey e Telles, 2006; Schwartzman, 2007; Telles, 2002) e a reclassificação racial em Porto Rico (Loveman e Muniz, 2007). Além disso, modelos binomiais superam várias restrições impostas pelo popular método de mínimos quadrados ordinários (MQO). Esses modelos, por exemplo, não pressupõem linearidade entre variáveis dependentes e independentes, não requerem que as variáveis sejam normalmente distribuídas e não assumem homocedasticidade do erro da regressão (Agresti, 2002; Allison, 2001; Long e Freese, 2003).

13. O padrão relativo de consumo reflete os quartis de um índice construído segundo a posse ou a presença de bens duráveis no domicílio. O índice de consumo é definido segundo a equação $\sum_{i=1}^n \frac{Di \cdot n}{Di}$, sendo Di igual a 1 quando o domicílio possui o bem durável ou serviço i (rádio, banheiro, empregada, aspirador de pó, máquina de lavar, videocassete, geladeira, congelador, telefone, telefone celular, computador, carro, internet); e 0 se o domicílio não possui o bem ou serviço i . Já \bar{D}_i é o percentual de domicílios na PESB com o bem durável i ; e n é o número absoluto de bens e serviços possuídos pelo domicílio. O índice é padronizado para variar entre 0 e 1, sendo 1 o valor máximo de consumo. Em vez da variável renda, usa-se o padrão de consumo relativo pelo fato de ele: 1) ter uma característica mais permanente do que a renda declarada; 2) apresentar um número de declarações maior e menos enviesado do que no caso da renda; e 3) em função de a variável de consumo resolver em parte o problema de endogeneidade existente entre cor declarada e padrão de renda.
14. Utiliza-se a cor dos pais por ser o principal fator associado à cor autodeclarada pelos entrevistados. Na PESB 2002, a cor da pele dos ascendentes é citada por 35% dos entrevistados como razão para classificar-se em dada cor/raça. A cor da pele dos pais é a segunda razão mais citada depois da própria cor dos entrevistados. O peso do determinismo genético sobre a cor declarada também tem sido demonstrado em outros estudos (Loveman e Muniz, 2007; Schwartzman, 2007).
15. Para facilitar a apresentação e a discussão dos resultados, as probabilidades de classificação entre as cores branca ou preta são simuladas utilizando-se a abordagem sugerida por King, Tomz e Wittenberg (2000), que utilizam simulações Monte Carlo para construir intervalos de confiança estatística. Para se obterem estimativas comparáveis, todas as características individuais foram fixadas e variaram-se apenas a cor dos pais e o método de coleta (cor autodeclarada segundo fotografias ou segundo o entrevistador). Dada a importância do debate sobre cotas universitárias, optou-se por definir um “tipo ideal” com características fixas mais prevalentes na amostra e próximas do perfil universitário. O “tipo ideal” adotado possui idade entre 18 e 24 anos, tem segundo grau completo, baixo padrão de consumo, é do sexo feminino, solteira, mora fora da região Nordeste e da capital. Ao contrário da estimativa pontual comumente apresentada na saída da regressão, na simulação Monte Carlo os valores preditos são extraídos de uma distribuição de coeficientes (ex.: *betas*) calculados a partir de 1.000 simulações *bootstraps*. Os coeficientes da estimativa encontram-se no Anexo 3.

REFERÊNCIAS BIBLIOGRÁFICAS

- AGRESTI, Alan. (2002), *Categorical Data Analysis*. New York, Wiley-Interscience.
- ALLISON, Paul D. (2001), *Logistic Regression Using the SAS System: Theory and Application*. Cary, NC/New York, SAS Institute/Wiley.
- BAILEY, Stanley R. (2008), "Unfixing for Race Making in Brazil". *American Journal of Sociology*, vol. 114, nº 3, pp. 577-614.
- . (2009), *Legacies of Race: Identities, Attitudes, and Politics in Brazil*. Stanford, CA, Stanford University Press.
- e TELLES, Edward E. (2006), "Multiracial versus Collective Black Categories. Examining Census Classification Debates in Brazil". *Ethnicities*, vol. 6, nº 1, pp. 74-101.
- CARVALHO, José Alberto Magno, WOOD, Charles H. e ANDRADE, Flávia Cristina Drumond. (2003), "Notas Acerca das Categorias de Cor dos Censos e sobre a Classificação Subjetiva de Cor no Brasil: 1980/90". *Revista Brasileira de Estudos de População*, vol. 20, nº 1, pp. 29-42.
- COHEN, Anthony P. (1994), *Self Consciousness: An Alternative Anthropology of Identity*. London/New York, Routledge.
- COHEN, Jacob. (1960), "A Coefficient of Agreement for Nominal Scales". *Educational and Psychological Measurement*, vol. 20, nº 1, pp. 37-46.
- DEGLER, Carl N. (1986), *Neither Black nor White: Slavery and Race Relations in Brazil and the United States*. Madison, WI, University of Wisconsin Press.
- ERIKSON, Erik H. (1968), *Identity, Youth, and Crisis*. New York, W. W. Norton.
- FERES JÚNIOR, João, CAMPOS, Luiz Augusto e DAFLON, Verônica Toste. (2011), "Fora de Quadro: A Ação Afirmativa nas Páginas d' O Globo". *Contemporânea – Revista de Sociologia da UFSCar*, nº 2, pp. 61-83.
- FRANCIS, Andrew M. e TANNURI-PIANTO, Maria. (2012), "Using Brazil's Racial Continuum to Examine the Short-Term Effects of Affirmative Action in Higher Education". *The Journal of Human Resources*, vol. 47, nº 3, pp. 754-784.
- . (no prelo), "Endogenous Race in Brazil: Affirmative Action and the Construction of Racial Identity among Young Adults". *Economic Development and Cultural Change*.
- GORDIS, Leon. (2004), *Epidemiology*. Philadelphia, PA, Saunders.
- GRAHAM, Richard *et alii*. (1990), *The Idea of Race in Latin America, 1870-1940*. Austin, University of Texas Press.
- HTUN, Mala. (2004), "From 'Racial Democracy' to Affirmative Action: Changing State Policy on Race in Brazil". *Latin American Research Review*, vol. 39, nº 1, pp. 60-89.
- IBGE (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA). (2011), "Características Étnico-Raciais da População: Um Estudo das Categorias de Classificação de Cor ou Raça, 2008". *Instituto Brasileiro de Geografia e Estatística, Ministério do Planejamento, Orçamento e Gestão*. Rio de Janeiro, IBGE.
- JENKINS, Richard. (2008), *Rethinking Ethnicity*. Thousand Oaks, CA, Sage.

- KING, Gary, TOMZ, Michael e WITTENBERG, Jason. (2000), "Making the Most of Statistical Analyses: Improving Interpretation and Presentation". *American Journal of Political Science*, vol. 44, nº 2, pp. 347-361.
- LANDIS, J. Richard e KOCH, Gary G. (1977), "The Measurement of Observer Agreement for Categorical Data". *Biometrics*, vol. 33, nº 1, pp. 159-174.
- LONG, J. Scott e FREESE, Jeremy. (2003), *Regression Models for Categorical Dependent Variables Using Stata*. College Station, TX, Stata Press.
- LOVEMAN, Mara e MUNIZ, Jerônimo O. (2007), "How Puerto Rico Became White: Boundary Dynamics and Intercensus Racial Reclassification". *American Sociological Review*, vol. 72, nº 6, pp. 915-939.
- e BAILEY, Stanley R. (2011), "Brazil in Black and White? Race Categories, the Census, and the Study of Inequality". *Ethnic and Racial Studies*. Disponível em <http://dx.doi.org/10.1080/01419870.2011.607503>.
- MUNIZ, Jerônimo O. (2010), "Sobre o Uso da Variável Raça-Cor em Estudos Quantitativos". *Revista de Sociologia e Política*, vol. 18, nº 36, pp. 277-291.
- NOBLES, Melissa. (2000), *Shades of Citizenship: Race and the Census in Modern Politics*. Stanford, CA, Stanford University Press.
- NOYMER, Andrew, PENNER, Andrew M. e SAPERSTEIN, Aliya. (2011), "Cause of Death Affects Racial Classification on Death Certificates". *PLoS One*, vol. 6, nº 1, e15812.
- OMI, Michael e WINANT, Howard. (1994), *Racial Formation in the United States: From the 1960s to the 1990s*. New York, Routledge.
- PESB (PESQUISA SOCIAL BRASILEIRA). (2002), "Pesquisa Social Brasileira". Universidade Federal Fluminense, dirigida por Alberto Carlos Almeida e Zairo Cheibub.
- RUGGLES, S. *et alii*. (2004), "Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]". Minneapolis, MN, Minnesota Population Center [producer and distributor].
- SANSONE, Livio. (2003), *Blackness without Ethnicity: Constructing Race in Brazil*. New York, Palgrave Macmillan.
- SAPERSTEIN, Aliya. (2006), "Double-Checking the Race Box: Examining Inconsistency between Survey Measures of Observed and Self-Reported Race". *Social Forces*, vol. 85, nº 1, pp. 57-74.
- . (2008), (Re)Modeling Race: Incorporating Racial Theory into Survey Research on Inequality. Tese de pós-doutorado (Sociology and Demography), University of California, Berkeley.
- . (2011), "Capturing Complexity in the United States: Which Aspects of Race Matter and When?" *Ethnic and Racial Studies*. Disponível em <http://dx.doi.org/10.1080/01419870.2011.607504>.
- e SYKES, Bryan L. (2008), "What You See and What She Gets: Isolating the Effect of Inconsistent Racial Classification on Women's Earnings and Income". Population Association of America Annual Meeting, New Orleans, 17-19 abril, pp. 57-74.

- SCHWARTZMAN, Luisa. (2007), "Does Money Whiten? Intergenerational Changes in Racial Classification in Brazil". *American Sociological Review*, vol. 72, nº6, pp. 940-963.
- SIMÕES, Solange e JERONYMO, Mauro. (2007), "Quem é Negro no Brasil? Identidade Racial e Sistemas de Classificação em uma Sociedade Miscigenada", in N. Aguiar (org.), *Desigualdades Sociais, Redes de Sociabilidade e Participação Política*. Belo Horizonte, Ed. UFMG, pp. 119-138.
- SKIDMORE, Thomas E. (1993), *Black into White: Race and Nationality in Brazilian Thought*. Durham, Duke University Press.
- TELLES, Edward. (2002), "Racial Ambiguity among the Brazilian Population". *Ethnic and Racial Studies*, vol. 25, nº 3, pp. 415-441.
- . (2004), *Race in Another America: The Significance of Skin Color in Brazil*. Princeton, NJ, Princeton University Press.
- e LIM, Nelson. (1998), "Does It Matter Who Answers the Race Question? Racial Classification and Income Inequality in Brazil". *Demography*, vol. 35, nº 4, pp. 465-474.
- ZUBERI, Tukufo. (2008), "Deracializing Social Statistics: Problems in the Quantification of Race", in T. Zuberi e E. Bonilla-Silva (orgs.), *White Logic, White Methods: Racism and Methodology*. Lanham, MD, Rowman & Littlefield, pp. 127-134.
- e BONILLA-SILVA, Eduardo. (2008), "Telling the Real Tale of the Hunt: Towards a Race Conscious Sociology of Racial Stratification", in T. Zuberi e E. Bonilla-Silva (orgs.), *White Logic, White Methods: Racism and Methodology*. Lanham, MD, Rowman & Littlefield, pp. 329-341.

Anexo 1

Fotografias Utilizadas na PESB Representando Diferentes Fenótipos



Fotografia 1



Fotografia 2



Fotografia 3

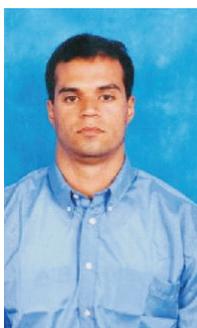


Fotografia 4



Fotografia 5

Fonte: PESB (2002).



Fotografia 6



Fotografia 7



Fotografia 8

Anexo 2

Intervalos de Confiança de 95% para as Probabilidades Preditas Acima

Cor do(a)		Branco			Preto ou Não Branco		
Pai	Mãe	Método 5	Método 6	Método 7	Método 5	Método 6	Método 7
branco	branca	0,95-0,99	0,91-0,97	0,81-0,90	0,01-0,05	0,03-0,08	0,10-0,19
branco	parda	0,65-0,84	0,64-0,83	0,46-0,65	0,16-0,35	0,17-0,36	0,35-0,54
branco	preta	0,25-0,58	0,25-0,58	0,13-0,39	0,42-0,75	0,42-0,74	0,61-0,87
pardo	branca	0,69-0,85	0,63-0,81	0,43-0,61	0,15-0,31	0,19-0,37	0,38-0,56
pardo	parda	0,39-0,62	0,39-0,61	0,13-0,25	0,38-0,61	0,39-0,61	0,75-0,87
pardo	preta	0,06-0,34	0,13-0,43	0,02-0,10	0,66-0,94	0,57-0,87	0,90-0,97
preto	branca	0,34-0,61	0,34-0,60	0,08-0,24	0,39-0,66	0,40-0,66	0,75-0,92
preto	parda	0,12-0,35	0,15-0,38	0,02-0,06	0,65-0,88	0,61-0,85	0,94-0,98
preto	preta	0,02-0,17	0,06-0,24	0,003-0,02	0,82-0,98	0,76-0,94	0,98-1,00

Elaboração do autor com base na PESB 2002.

Anexo 3

Razões de Chance de Ser Preto ou Não Branco Estimadas a Partir de Modelos Logísticos Binomiais segundo o Método de Coleta da Informação Racial

	Método 5	Método 6	Método 7
Mãe parda	15,21***	8,14***	5,11***
Mãe preta	66,71***	34,47***	19,91***
Pai pardo	13,43***	8,92***	5,52***
Pai preto	54,79***	28,36***	43,87***
Mãe e pai pardos	0,23***	0,34***	
Mãe parda e pai preto	0,23**	0,35*	
Mãe preta e pai pardo	0,30	0,24**	
Mãe e pai pretos	0,22*	0,19**	
Índice de consumo:			
2º quartil	1,04	1,19	0,74
3º quartil	0,81	0,96	0,68
4º quartil	0,65	0,88	0,48***
Idades:			
25 e 34 anos	1,54*	1,23	1,17
35 e 44 anos	1,66*	1,15	1,28
45 e 59 anos	1,05	1,44	1,04
60 anos ou mais	1,22	1,29	0,91
Região Nordeste	1,23	0,83	2,22***
Residência na capital	1,34*	1,64***	1,57**
Casado	1,01	0,82	0,76*
Escolaridade:			
até a 4ª série	0,62*	0,57*	1,06
de 5ª a 8ª série	0,81	0,68	0,93
2º grau completo	0,70	0,72	0,85
superior ou mais	0,59	0,51	0,48*
Constante	0,03***	0,06***	0,29***
N	1967	1967	1967
Chi2	891,33***	694,61***	1035,35***
BIC	1.783	1.928	1.813

Elaboração do autor com base na PESB 2002.

* p<0,05; **p<0,01; ***p<0,001.

Nota: O valor exponencial dos coeficientes é apresentado na tabela para representar a razão de chances de ser preto (ou não branco) em função de determinada covariável. O modelo referente ao método 7 não inclui interações entre a raça/cor dos pais em virtude do fato de a variável dependente ser completamente determinada pela cor do ascendente direto de raça/cor mais escura.

ABSTRACT

Black or White? Measurement, Relevance, and Classificatory Concordance in a Country of Racial Uncertainty

This article compares agreement among racial classification schemes in Brazil, considering measurement methodologies with three categories (white, brown, and black) and two (white and black). The racial taxonomic uncertainty is illustrated by the statistically reliable estimates of the probabilities of being white or black, adjusting for individual characteristics, parents' color, and categorization schemes. The findings show that if different data collection methods were used to construct the color variable, only half of the respondents would report their color consistently in all of them. The data come from the Brazilian Social Survey conducted by UFF in 2002. The survey is representative of the country as a whole and is the first to allow the study proposed here.

Key words: uncertainty; race; Brazil

RÉSUMÉ

Noir sur blanc? Mensuration, Importance et Correspondance des Classements au Pays de l'Incertitude Raciale

Dans cet article, on cherche à savoir si, au Brésil, il y a similitude entre des modèles de classement racial selon quatre méthodologies de mensuration avec trois catégories (le blanc, le métis ou le noir) ou deux catégories (le blanc ou le noir). L'incertitude taxinomique raciale est illustrée à partir d'estimations statistiquement sûres des probabilités d'être blanc ou noir, dont le contrôle se fait selon des traits individuels, la couleur de la peau des parents et divers schèmes de catégories. Selon les résultats, si différentes méthodes de prélèvement de données sont utilisées dans la construction de la variable couleur de la peau, seuls 50% des sujets affirment leur propre couleur de façon tranchée dans toutes les méthodes. Ces données sont tirées de la Recherche Sociale Brésilienne menée par l'Université Fédérale Fluminense (UFF) en 2002, à partir d'un échantillon représentatif de la population nationale, le premier à permettre l'étude en question.

Mots-clés: incertitude; race; Brésil