

O teste Stanford-Binet no Brasil: uma análise das revisões de Pernambuco e do Rio de Janeiro (1925-1935)

The Stanford-Binet test in Brazil: the experiments in Pernambuco and Rio de Janeiro (1925-1935)

Ana Cristina S. M. Rocha | Fundação Oswaldo Cruz

anasmrocha@gmail.com

<https://orcid.org/0000-0001-5718-1293>

RESUMO Em 1916, Lewis Terman publicou uma das versões mais populares do teste de inteligência de Binet-Simon nos Estados Unidos: a Stanford-Binet. Ela foi considerada um marco na história dos testes porque inseriu o conceito de quociente de inteligência (QI) nesse tipo de avaliação. Este artigo procura investigar o processo de elaboração de duas versões do Stanford-Binet no Brasil, cujo trabalho foi liderado por Ernani Lopes no Rio de Janeiro e Ulysses Pernambucano no Recife. Nosso objetivo é o de demonstrar que esses pesquisadores tiveram um papel ativo na reelaboração desses testes, que muitas vezes justificaram a vinculação que foi estabelecida também por eles entre raça, classe e inteligência.

Palavras-chave: história da psicologia – Lewis Terman (1877-1956) – testes de inteligência – Ulysses Pernambucano (1892-1943) – Ernani Lopes (1885-1969).

ABSTRACT In 1916, Lewis Terman published one of the most popular versions of the Binet-Simon intelligence test in the United States: the Stanford-Binet. It was considered a milestone in the history of tests because it introduced the concept of intelligence quotient (IQ) into this type of assessment. This article seeks to investigate the process in which two versions of the Stanford-Binet were devised in Brazil, in a collective effort led by Ernani Lopes in Rio de Janeiro and Ulysses Pernambucano in Recife. Our goal is to demonstrate that these researchers played an active role in the re-elaboration of these tests, which often justified the link that was also established by them between race, class, and intelligence.

Keywords: history of psychology – Lewis Terman (1877-1956) – IQ tests – Ulysses Pernambucano (1892-1943) – Ernani Lopes (1885-1969).

Introdução¹

Entre o final da década de 1920 e os anos de 1930, diversos educadores trabalharam com versões de testes de inteligência estrangeiros em escolas brasileiras. Diante da presumida heterogeneidade da população escolar decorrente da expansão e estruturação do sistema público de ensino, estes intelectuais defendiam que classes escolares mais homogêneas do ponto de vista da inteligência seriam uma das chaves da eficiência do ensino, marcado por altas taxas de reprovação e de abandono. As consequências da relação que se estabeleceu entre educação, eficiência e ciência neste período foram exploradas nos estudos sobre o fenômeno do “fracasso escolar”, a exemplo dos trabalhos de Maria Helena Souza Patto (1999, 2004, 2015), referência no campo. Dentre elas está a existência de um perfil que se estabeleceu entre alunos com ou sem potencial de sucesso, que explora marcadores de raça e classe sem, necessariamente, explicitá-los.

Este mesmo fenômeno pode ser analisado do ponto de vista da emergência do que Carlos Monarcha (2008) vai chamar de “movimento dos testes”. Perpassado pelo ideal de modernidade do começo do século XX, este movimento está associado ao esforço dos educadores em estabelecer parâmetros científicos para orientar suas práticas educacionais. Além disso, ele também se relaciona aos debates escolanovistas de centrar a escola no aluno, reconhecendo e respeitando a criança enquanto indivíduo. No Brasil, todas estas transformações foram permeadas por discussões sobre a obrigatoriedade escolar e, assim como em outros países, pelo fenômeno da “escola de massas”, que tem nos grupos escolares uma de suas marcas (Fass, 1980; Souza-Chaloba, 2019).

Muitos dos princípios adotados pelos pesquisadores envolvidos com o uso de testes de inteligência neste período se conectam a mudanças cujos marcos se localizam no século XIX. Um destes marcos se relaciona com as pesquisas de Francis Galton sobre a relação entre hereditariedade e genialidade, cuja importância não se resumiu a sua defesa de que era possível medir a inteligência. Como exploram Theodore Porter (2020) e Ian Hacking (2013), Galton foi fundamental para estabelecer os cálculos estatísticos necessários às medições e comparações que serão feitas posteriormente por autores como Lewis Terman, que exploramos neste artigo.

Outra mudança se relaciona à “invenção da normalidade”, uma noção que transita do par normal/patológico da esfera médica para o par normal/anormal do senso comum (Hacking, 2013, p. 160-161). Quando explora essa questão, Ian Hacking adiciona mais uma camada que é importante ter em mente aqui: a de que “nós temos usado regularmente ‘normal’ para preencher a lacuna entre o ‘ser’ e o ‘dever’” (p. 163). Em seu livro, Hacking estabelece uma tensão entre preservação e melhoramento, o normal durkheimiano – associado ao “correto” e que resguarda o *status quo* – e o normal galtoniano – associado à média e que, portanto, pode ser melhorado (p. 168-169). Esta dualidade é parte fundamental do trabalho dos pesquisadores que vamos explorar neste artigo. De certo modo, eles analisam as crianças enquanto procuram conciliar as expectativas que têm sobre elas com a norma, as médias, e o que é desejável obter com suas experiências.

1 A ideia deste artigo foi apresentada no 32º Simpósio Nacional de História – Anpuh Nacional, em São Luís (MA), cujo resumo foi parcialmente aproveitado nesta introdução. Aos colegas do ST57, “Eugenia, genética e raça” e do projeto “História da genética e de seus impactos: raça, racismo e população no Brasil”, agradeço as sugestões e dúvidas suscitadas no debate, em parte incorporadas na escrita deste trabalho. Aos pareceristas também agradeço pelas sugestões apresentadas.

Em seu livro sobre *A falsa medida do homem*, Stephen Gould (2018) explorou como a criação de testes de inteligência estava atrelada a uma série de pressupostos relacionados ao determinismo biológico. De acordo com o autor, a “abstração da inteligência” foi vista como “entidade única, localizada no cérebro, quantificada na forma de um número único para cada indivíduo” (p. 9). Situada dentro dessa lógica, essa história é também a história do “uso desses números na hierarquização das pessoas numa escala única de méritos, que indica invariavelmente que os grupos oprimidos e em desvantagem – raças, classes ou sexos – são inatamente inferiores e merecem ocupar essa posição” (p. 9).

Neste artigo, nos concentramos nos anos de 1920 e 1930 para analisar este fenômeno a partir de um recorte bastante específico: o trabalho de adaptação e tradução dos testes de Stanford-Binet realizados no Rio de Janeiro e em Pernambuco por pesquisadores brasileiros. Os resultados destes trabalhos foram publicados por Ernani Lopes² e Maria Brasília Leme Lopes³ nos *Archivos Brasileiros de Hygiene Mental (ABHM)* e por Ulysses Pernambucano⁴, Anita Paes Barreto⁵ e outros autores nos *Arquivos da Assistência a Psicopatas de Pernambuco*.

O Stanford-Binet foi uma das mais famosas adaptações do teste de Binet-Simon desenvolvida por Lewis Terman (1877-1956) na Universidade de Stanford, nos Estados Unidos. Embora não seja a única adaptação do Binet-Simon, ela marca uma mudança importante no contexto das mensurações psicológicas. Partindo de um teste originalmente desenhado para auxiliar a identificar crianças com dificuldades de aprendizagem, ele desenvolveu sua adaptação para que

- 2 Ernani Lopes (1885-1969) formou-se na Faculdade de Medicina do Rio de Janeiro em 1910. Como médico psiquiatra, trabalhou no Hospital Nacional de Alienados e na Colônia de Psicopatas do Engenho de Dentro, instituição que dirigiu. Foi livre docente das cadeiras de Clínica Psiquiátrica e Moléstias Nervosas e de Clínica Neurológica na Faculdade de Medicina do Rio de Janeiro. Foi presidente da Liga Brasileira de Higiene Mental (LBHM) além de editor da revista da Liga, os *Archivos Brasileiros de Higiene Mental* (Cerqueira, 2014; Accorsi, 2018).
- 3 Maria Brasília Leme Lopes (1909-1996) formou-se como normalista em 1925 na Escola Normal do Rio de Janeiro. No ano seguinte passou a trabalhar como amanuense na Diretoria-geral de Assistência Municipal. A inserção de Leme Lopes no laboratório de psicologia da LBHM se dá a partir de sua participação no curso de “técnica de psicologia aplicada” ministrado por Alfred Fessard e Annette Baron (conhecidos como o casal Fessard) em 1926. Por sua vez, este curso foi parte das atividades desenvolvidas por Fessard no Brasil, contratado pela Liga para fundar o seu laboratório de psicologia. Embora seja reconhecida como pioneira no campo da psicologia, Leme Lopes tornou-se referência no campo da hemoterapia, área em que passou a atuar a partir de 1944, depois de se formar em medicina (Portavales e Jacó-Villela, 2016; Facchinetti, Mota, Muñoz, 2022).
- 4 Ulysses Pernambucano (1892-1943) nasceu em Recife e formou-se na Faculdade de Medicina do Rio de Janeiro em 1912, período em que fez estágio no Hospital Nacional de Alienados. De volta à Recife, trabalhou no Hospital de Tamarineira e ocupou a cadeira de Psicologia da Escola Normal de Pernambuco, instituição que também dirigiu. Na Faculdade de Medicina do Recife, ocupou a cadeira de Neuro-psiquiatria Infantil e foi catedrático de Clínica Neurológica. As experiências de Pernambucano e sua equipe com testes foram desenvolvidas no Instituto de Psicologia de Pernambuco, que depois passou a fazer parte da Divisão de Assistência a Psicopatas de Pernambuco em 1931, instituição que dirigiu até 1935 (Vincentini, 2018; Prof. Ulysses..., 1944).
- 5 Anita Paes Barreto (1907-2003) formou-se normalista na Escola Normal do Recife em 1924. No ano seguinte passou a trabalhar com Ulysses Pernambucano no recém-criado Instituto de Psicologia de Pernambuco. Mesmo com o fechamento do Instituto de Psicologia, em 1935, Anita Barreto continuou atuando no campo da psicologia e da educação. Foi professora da Faculdade de Filosofia do Recife, e assessora e depois secretária de Educação do Recife durante o governo de Miguel Arraes. É autora de estudos pioneiros no campo da psicologia aplicada, alguns dos quais realizados em conjunto com Ulysses Pernambucano. Pela importância de seu trabalho para o campo da psicologia e sua profissionalização, foi homenageada pelo Conselho Federal de Psicologia em 1997. Em 2020, foi declarada Patrona da Psicologia em Pernambuco pela Assembleia Legislativa do estado (Anita..., 2000; Recife, 2020).

fosse possível comparar o desempenho dessas crianças. Ao utilizar o conceito de Quociente de Inteligência (QI) de Stern, Terman desenhou uma tabela que hierarquizava esse desempenho e classificava essas crianças numa graduação que ia desde a “genialidade” e a “inteligência muito superior” até o “*feeble-minded*” ou “débil mental” (Terman, 1916, p. 79; Gould, 2018).

A popularidade do Stanford-Binet nos Estados Unidos foi frequentemente abordada no debate sobre o uso dos testes de inteligência no Brasil, como vamos explorar na próxima seção. Embora muito citado como exemplo nestes trabalhos, é importante notar que o Stanford-Binet não foi o único teste utilizado no país. Muitas pesquisas realizadas sobre o tema indicam a multiplicidade de autores e testes aqui adotados, a exemplo do Dearborn, do Pinter-Cunningham, das cem questões de Ballard e do labirinto de Porteus (Rocha, 2020; Rota Jr., 2016; Jacó-Vilella e Silva, 2016; Campos e Borges, 2014). Por outro lado, em nossa pesquisa, o trabalho de Terman tem se mostrado central para os esforços que se desenvolveram em paralelo para identificar as crianças que na época eram chamadas de “supranormais”, como indicam as experiências de Isaías Alves (1932b, 1934) no Distrito Federal e as publicações decorrentes do trabalho desenvolvido na Escola de Aperfeiçoamento de Belo Horizonte (Frota, 1933; Antipoff, 1946).

A análise do modo como educadores brasileiros se apropriaram de ferramentas de mensuração psicológica, como o teste Stanford-Binet, está relacionada também a uma questão mais ampla: a de como estes intelectuais interpretaram as concepções sobre hereditariedade, raça, gênero e classe de autores como Lewis Terman, cujas convicções estavam explicitamente colocadas ao lado da discussão que ele desenvolve sobre os resultados obtidos com suas experiências. Neste sentido, Terman é especialmente importante não só pelo alcance de sua ferramenta, mas também pelo seu papel no movimento eugênico estadunidense (Stern, 2005; Carson, 2007, Levine, 2017).

Quando explora a eugenia no contexto latino-americano, Stepan (2005, p. 46) sintetiza parte das mudanças pelas quais o Brasil passou no fim do século XIX: “o colapso final de uma sociedade escravocrata, a abertura do país à imigração europeia em grande escala [...], a queda da monarquia e a criação da República”. As primeiras décadas do século XX são de intenso debate sobre o caráter nacional brasileiro, que tenta conciliar o diagnóstico de uma população predominantemente mestiça com as ideias de modernidade e progresso que têm a Europa como parâmetro.

Neste período, saúde e educação viraram as bandeiras de intelectuais comprometidos com a questão nacional, seja a partir de campanhas contra o analfabetismo e pela defesa de uma ampla reforma educacional, seja a partir de campanhas por melhores condições de saúde da população – como fez o movimento sanitário. Para Stepan (2005, p. 47), a questão nacional torna-se sinônimo de uma “questão social”. Ainda de acordo com a autora,

O grupo que mais inquietava os médicos, os especialistas em saneamento e os reformadores brasileiros era, em sua maior parte, constituído por negros e mulatos. Esses profissionais presumiam que doenças sociais se acumulavam na base da hierarquia sócio racial – que os pobres eram pobres porque eram anti-higiênicos, sujos, ignorantes e hereditariamente inadequados. Os preconceitos de raça e classe fundiram-se, por conseguinte, na linguagem da hereditariedade (Stepan, 2005, p. 47).

É também neste entrecruzamento entre raça e classe, bem como entre pobreza, doença e degeneração que se situam os diagnósticos sobre a inteligência das crianças que frequentavam as escolas brasileiras. Por isso, este artigo também dialoga com o argumento de Jerry Dávila (2006) sobre o modo como os educadores brasileiros se apropriaram das afirmações de Lewis Terman sobre o peso da hereditariedade no desenvolvimento da inteligência de cada indivíduo. De acordo com ele,

Embora Lewis Terman acreditasse na superioridade intrínseca da herança genética de alguns intelectos, seu teste foi facilmente adaptado à versão brasileira de degeneração porque Alves e os outros liam os resultados dos testes como indicativos das condições culturais e sociais pobres dos brasileiros testados. Ainda assim, os testes separavam os alunos brancos dos de cor e os alunos ricos dos pobres (Dávila, 2006, p. 68).

Assim como sinaliza Dávila em seu livro, é possível identificar uma interseção entre as aplicações experimentais dos testes de inteligência coordenadas por educadores como Isaías Alves com o debate racial da época (Jacó-Vilela, 2014; Rocha, 2011, 2023).

O que este artigo procura demonstrar é que, mais do que uma interpretação enviesada dos resultados de aplicações de testes estrangeiros em crianças brasileiras, os pesquisadores aqui explorados tiveram um papel ativo na própria elaboração de ferramentas que justificaram a vinculação estabelecida também por eles entre raça, classe e inteligência. Para isso, é preciso olhar para o modo como eles redesenharam os testes originais nos quais se basearam, num processo que era considerado por todos como parte essencial do trabalho de adequação dos testes de inteligência às realidades locais.

Essa lógica parte também de um princípio discutido por Maria Souza Patto (2015) em seu livro sobre *A produção do fracasso escolar*. Ele trata do modo como, durante o trabalho de elaboração dos testes psicológicos, as nossas diversas habilidades foram vinculadas ou não a um indicativo do que chamamos de inteligência. Ainda no século XIX, Francis Galton acreditava que “medir processos sensoriais (como a discriminação visual, auditiva e cinestésica) e motores (como a velocidade do tempo de reação)” era um caminho para “estimar o nível intelectual” (Patto, 2015, p. 61-62). Mas essa premissa foi se modificando ao longo do tempo. De acordo com a autora,

É importante registrar que tão logo a psicologia começou a acumular dados que mostravam diferenças na realização de testes sensório-motores favoráveis a grupos humanos tidos como inferiores pela ideologia racista, o bom rendimento em processos sensoriais e motores, tido por Galton pouco mais de uma década antes como indicador de alta capacidade intelectual, passa a receber interpretação oposta, isto é, passa a ser tomado como indicativo de inferioridade intelectual (Patto, 2015, p. 62, nota 26).

Assim como Patto (2015), quando Andrew Winston (2020, p. 4) analisa os primeiros estudos que buscavam mensurar as habilidades mentais nos Estados Unidos, ele explora essa mesma variação. De acordo com ele, o “tempo de reação se revelaria um recurso discursivo altamente flexível; tempos mais rápidos poderiam ser interpretados tanto como um sinal de desenvolvimento racial ‘primitivo’ quanto como ‘avançado’” (p. 4).

Entre os anos que separam as primeiras experiências de Galton e os testes que vamos analisar aqui, os pesquisadores discutiram exaustivamente não só o próprio conceito de inteligência, como os fatores que poderiam ser considerados como significativos para sua mensuração. Como sinaliza Patto (2015), essas discussões não estavam desvinculadas das ideias eugênicas defendidas por Galton e seus sucessores, ainda que muitas vezes tenham se pulverizado em investigações bastante específicas sobre fatores como memória, atenção e fadiga.⁶

No recorte que adotamos, uma parte significativa desse trabalho de seleção das habilidades a serem medidas já havia sido realizado. Testes como os de Binet procuravam mensurar uma série delas, que só em conjunto passaram a explicar o desempenho dos alunos testados. Além disso, reconhecia-se o papel de fatores externos nos resultados obtidos das crianças: barulho, falta de familiaridade com o examinador, duração excessiva do exame, dentre outros. Os pesquisadores também desenharam estratégias para mitigá-los, o que também explica parte da importância que davam às “instruções ao examinador” e às condições de aplicação dos testes. Até mesmo o peso da escolaridade no desempenho de algumas das questões aplicadas era alvo de debates, ainda que este fosse um dos pontos de menos consenso.

De todo modo, não podemos encarar as experiências aqui desenvolvidas como meras aplicações de princípios que lhes eram exteriores. Os defensores do uso de testes de inteligência no Brasil reconheciam a necessidade de não só traduzir como adaptar as questões dos testes às especificidades locais, desde que preservando as condições de comparabilidade entre as experiências. Deste modo, também aqui, as habilidades medidas foram discutidas e os “testes” foram revistados camada a camada: figuras de selos foram trocadas por moedas, diamantes por losangos e traduções literais deram lugar a situações semelhantes (Rocha, 2020). Neste processo, alguns testes foram eliminados e a importância de determinadas questões foi atenuada, como veremos a seguir. Estas modificações são fundamentais para compreender o fenômeno que exploramos neste artigo.

As adaptações foram parte da força argumentativa dos pesquisadores brasileiros, que exploravam seus resultados muitas vezes reafirmando as expectativas de raça e classe que recaíam sobre a população escolar. Por isso, examinar não só o formato final das adaptações, mas também o seu processo de construção, pode contribuir para uma melhor compreensão do modo como os testes de inteligência foram utilizados no Brasil. Para isto, escolhemos comparar duas das adaptações dos testes Stanford-Binet realizadas no Brasil entre as décadas de 1920 e 1930, sinalizando um horizonte de escolhas e expectativas que seus principais autores, Ernani Lopes e Ulysses Pernambucano, tiveram à disposição. Antes, vamos explorar o contexto de produção da versão original deste teste, produzido por Terman, assim como alguns dos usos anteriores desta ferramenta no Brasil.

Lewis Terman e o Stanford-Binet no Brasil

A publicação de *The measurement of intelligence* em 1916 – livro que apresenta a primeira versão do Stanford-Binet desenvolvida por Lewis Terman – é considerado um marco na história da psicologia e dos testes de inteligência. Em seu trabalho sobre o movimento dos testes nos

6 Um bom exemplo são os demais trabalhos de Maria Brasília Leme Lopes desenvolvidos no Laboratório da Liga e publicados nos *ABHM*. Cf. Jacó-Vilela (2014).

Estados Unidos e o papel de Lewis Terman nesse contexto, Paul Chapman (1988, p. 17) afirma que “a publicação do Stanford Binet culminou mais de 25 anos de investigação intensiva na América e na Europa sobre a natureza da inteligência e como medi-la”. Isso implica reconhecer que seu contexto de criação engloba o desenvolvimento da psicologia experimental na Alemanha no final do século XIX, as investigações de Francis Galton sobre as diferenças individuais, e o florescimento das pesquisas sobre aprendizagem na França, que tem a “escala métrica da inteligência” de Alfred Binet como um de seus resultados.

Nos Estados Unidos, Henry Goddard (1866-1957) foi um dos primeiros pesquisadores a traduzir e aplicar experimentalmente o teste de Binet-Simon em 1908. Como diretor de pesquisas do Vineland Training School for Feeble-Minded Girls and Boys em Nova Jersey, ele desenvolveu pesquisas sobre deficiência intelectual utilizando a escola como seu laboratório. Além da adesão ao movimento eugenista, Goddard e Terman têm em comum o doutorado em psicologia realizado na Clark University em Massachusetts, sob a orientação de Stanley Hall – que por sua vez estudou psicologia na Alemanha, sob a orientação de Wundt (Chapman, 1988).

Já Lewis Terman começou sua formação profissional no Central Normal College em Indiana, obtendo o diploma de normalista em 1898. Em 1901 ingressou na Indiana University, onde fez seu *Bachelor of Arts* e iniciou seus estudos em psicologia durante o mestrado. Em sua tese de doutorado – defendida na Clark University em 1905 e intitulada *Genius and stupidity* – Terman investigou os tipos de processos mentais que estavam relacionados à inteligência. A partir de seus estudos, ele defendia que a chave para a inteligência estava na “capacidade de desenvolver o pensamento abstrato” (Terman, 1921, p. 128). Assim, parte das modificações que Terman fez no teste original de Binet-Simon também se relaciona com essa convicção.

Durante seu período de formação, Lewis Terman também trabalhou como professor em diversas escolas e dirigiu a San Bernardino School, na Califórnia. Em 1910 ingressou na Stanford University como professor de Psicologia Educacional no recém-criado Departamento de Educação. Foi nesta universidade que ele retomou suas pesquisas sobre inteligência, aplicando experimentalmente o Binet-Simon em estudantes da região, num trabalho que teve entre seus resultados o desenvolvimento de uma versão ampliada do teste original: a escala Stanford-Binet (Minton, 1988; Chapman, 1988).

De acordo com Terman, além das questões originais de Binet, quarenta questões adicionais foram inseridas nas aplicações experimentais, de modo que os testes menos eficientes pudessem ser eliminados de sua escala. Ele também adicionou uma questão a mais para cada idade, de modo que todas as idades dispunham de seis testes, em lugar dos cinco testes da escala de Binet. Além disso, em alguns casos, a posição dos testes, que expressava sua correspondência com uma idade específica, também foi modificada (tanto para cima quanto para baixo). Terman também ampliou as idades testadas, tornando possível verificar a inteligência não só do “adulto médio”, como também do “adulto de inteligência superior”. Como resultado, a Stanford-Binet tinha noventa questões, em lugar das 54 originalmente criadas por Binet e Simon (Terman, 1916).

Ao estudar o desenvolvimento da eugenia nos Estados Unidos, Alexandra Stern (2005) destaca a importância de Lewis Terman para este movimento. Para a autora, sua relevância não decorre apenas de sua participação em debates como os promovidos pela American Eugenics Society, mas também pelos discursos e políticas que sua escala viabilizou. Referindo-se a Terman, Stern (2005, p. 19) afirma que:

Rastrear sua trajetória ilumina a ascensão meteórica dos testes padronizados e demonstra como as correlações entre raça e inteligência foram incorporadas aos métodos estatísticos. As classificações numéricas para debilidade mental, imbecilidade e idiotice e os quocientes médios de inteligência (QI) esperados de diferentes grupos raciais que ele detalhou em *The measurement of intelligence*, publicado em 1916, dominaram a psicometria por anos.

Além da psicometria em si, Stern (2005) sinaliza como sua escala serviu de base para a defesa do Johnson-Reed Immigration Act de 1924⁷ e da esterilização compulsória dos indivíduos que fossem classificados com QI abaixo de setenta, o que implicava algum tipo de debilidade mental de acordo com os parâmetros de Terman.

Embora nosso foco seja o seu trabalho com o Stanford-Binet, cuja escala foi publicada em *The measurement of intelligence* (Terman, 1916), é preciso mencionar que as posições de Terman relacionadas a contribuição de fatores como raça, classe e gênero para o desenvolvimento da inteligência mudaram ao longo dos anos. A partir da década de 1930, Lewis Terman passou a focalizar menos em questões raciais e mais em questões de gênero para explicar as diferenças de desempenho nos testes (Stern, 2005). Assim, se em 1916 Terman afirmava que a inteligência média de homens e mulheres era semelhante, em 1930 ele passou a considerar a possibilidade de que a inteligência estivesse associada à traços de masculinidade/feminilidade. No terceiro volume da série *Genetic studies of genius*, publicado em 1930, Terman explicou a diferença do desempenho encontrada em seu estudo longitudinal sobre a genialidade poderia estar relacionada ao fato de que “o desenvolvimento mental dos meninos continuasse depois que o das meninas estagnasse” (Vialle, 1994, p. 33).

De todo modo, os trabalhos que aqui analisamos não parecem considerar os trabalhos de Terman pós-1930. Embora lidem com uma série de referências atualizadas do debate estadunidense sobre o uso de testes de inteligência, os artigos que são resultado das pesquisas lideradas por Ernani Lopes e Ulysses Pernambucano publicizaram, no começo dos anos de 1930, o resultado de um trabalho que havia sido iniciado ainda na década anterior, de 1920. Ainda assim, o sexo das crianças testadas costumava ser um dos dados registrados pelos pesquisadores durante as aplicações. Neste sentido, embora não seja objeto deste artigo, a relação que se estabelece entre gênero e inteligência no Brasil é um tema relevante, a ser explorado em trabalhos futuros.

Por outro lado, embora nem todos os intelectuais brasileiros tenham discutido suas traduções e adaptações do Stanford-Binet de forma detalhada – como no caso das experiências lideradas por Ernani Lopes e Ulysses Pernambucano –, é possível afirmar que o uso destes testes antecede o trabalho desenvolvido por eles. Ao analisar o uso de testes psicológicos no Pavilhão de Observações do Hospício Nacional de Alienados (HNA), Jacó-Villela et al. (2022, p. 371) registram o uso do Stanford-Binet em três pacientes: o primeiro em 1917 e os dois outros em 1920. A informação foi obtida a partir dos Cadernos de Observações Clínicas examinados, com datas entre 1907 e 1925. Desse modo, não foi possível entender as condições de uso e de

7 A lei imigratória de 1924, também conhecida como “Johnson-Reed Act”, passou a limitar a entrada de imigrantes nos Estados Unidos estabelecendo cotas por nacionalidade. A lei também estabeleceu a proibição da imigração asiática, resguardados os casos do Japão e das Filipinas, que tiveram legislações próprias. Antes da lei de 1924, os Estados Unidos já proibiam a imigração chinesa através do “Chinese Exclusion Act”. A justificativa da lei era a de preservar a “homogeneidade” da população estadunidense. Além de Stern (2005), cf. <https://history.state.gov/milestones/1921-1936/immigration-act>.

adaptação do teste, mas cabe ressaltar, como fazem os autores, a proximidade entre a publicação do Stanford-Binet em 1916 e seu primeiro uso no HNA, em 1917.

Em artigo de 1930, Mário Saturnino,⁸ afirmou que o uso do Stanford-Binet na seção do HNA coordenada por Murillo Campos era “sistemático” (Saturnino, 1930, p. 13). Como tenente-médico do Exército, Campos ficou responsável pelo atendimento dos militares ali internados (Jacó-Vilella et al., 2022).⁹ A partir do trabalho ali realizado, Saturnino analisou uma amostra de cem soldados, escolhidos de forma aleatória, por sorteio. Destes, pelo menos 15 foram diagnosticados com alguma “debilidade mental”. Apenas oito deles foram classificados igual ou acima da idade mental média de um adulto nos padrões de Terman, que variava entre 15 e 17 anos (Saturnino, 1930, p. 14; Terman, 1916, p. 55).

Além da experiência do HNA, o Stanford-Binet foi explorado por Albuquerque no livro *Tests*, que popularizou o debate sobre o uso de testes de inteligência no Brasil e foi originalmente publicado em 1924. O livro contém uma tradução do Stanford-Binet brevemente comentada pelo autor, mas que não inclui os testes voltados para adultos (Albuquerque, 1928, p. 88-100). Como sinaliza Albuquerque, seu objetivo era o de divulgar tanto a concepção vinculada aos testes de inteligência quanto alguma de suas versões. Assim, o esforço do autor estava mais concentrado na síntese e na discussão do tema em si.

No ano seguinte, em 1925, C. A. Baker publicou um trabalho sobre o uso de testes de inteligência em geral que continha também um “guia para realização do teste Binet-Simon-Terman” (Monarcha, 2008, p. 8). Em 1926, Isaías Alves comentou a versão do Stanford-Binet em seu livro, *Teste individual de inteligência*, mas optou por traduzir a escala Binet-Burt, uma adaptação londrina do Binet-Simon, que tinha a vantagem de ser mais curta que a versão de Terman.

Além dessas publicações, é possível inferir que outros educadores conheciam o Stanford-Binet em sua versão original, a partir das pistas que encontramos em seus arquivos pessoais. Este é o caso de Lourenço Filho, cujo arquivo possui um cartão de registro das respostas dos alunos que era parte do material de aplicação do teste comercializado pela Riverside Press ([Documentos...], 1927-1964). Já no arquivo pessoal de Alda Lodi, a versão que encontramos é a de Ernani Lopes, já com material de aplicação e registrado como parte do material da Escola de Aperfeiçoamento (EA) de Belo Horizonte.¹⁰ A pesquisa de Rota Jr. (2016) também atesta o uso do Stanford-Binet como parte das atividades da EA.

As versões

As versões do Stanford-Binet analisadas para este artigo foram fruto de um trabalho coletivo, liderado por Ernani Lopes no Laboratório de Psicologia da Liga Brasileira de Higiene Mental¹¹ e por Ulysses Pernambucano, primeiro no Instituto de Psicologia da Escola Normal e

8 “Médico do Corpo de Saúde do Exército e titular da Seção de Higiene Militar da LBHM” (Saturnino, 1930, p. 12).

9 Murilo de Campos era também membro da LBHM.

10 O Arquivo Pessoal de Alda Lodi está depositado na Biblioteca Bartolomeu de Queiroz do Museu da Escola de Minas Gerais, em Belo Horizonte. Embora disponível para consulta, a documentação não foi catalogada.

11 A LBHM foi fundada por Gustavo Riedel no Rio de Janeiro, em 1923. De acordo com José Reis (1994, p. 50), os objetivos da Liga associavam questões relacionadas à “assistência psiquiátrica aos doentes ‘nervosos e mentaes’, dentro e fora do asilo” com a “atenção à higiene mental propriamente dita, definitivamente alinhada a

depois como parte das atividades da Divisão de Assistência a Psicopatas de Pernambuco.¹² Elas foram divulgadas de forma semelhante, em artigos publicados a partir de 1931 nas revistas em que eles eram editores: os *Archivos Brasileiros de Higiene Mental* e os *Arquivos da Assistência a Psicopatas de Pernambuco (AAPP)*, respectivamente. Nos dois casos, eles contaram com a participação de professoras que já atuavam no ensino primário, o que pode ter facilitado a circulação das avaliações no interior das escolas.

O trabalho de adaptação coordenado por Ulysses Pernambucano foi iniciado em 1925 (Pernambucano e Campos, 1931). O primeiro número dos *Archivos*, publicado em outubro de 1931, tem um artigo sobre a adaptação intitulado "O teste 'a bola e o campo' em crianças de 12 e 13 anos", que discute os resultados da aplicação deste teste específico em alunos das escolas de Recife. A "escala pernambucana" foi publicada em seu formato final em 1935 nos *AAPP* por Anita Barreto (1935). Durante esse período, a equipe também desenvolveu experiências paralelas, que ajudavam na adaptação. Um exemplo foi o estudo experimental sobre "O vocabulário das crianças das escolas primárias do Recife" (Pernambucano e Barreto, 1931), que contribuía para refletir sobre a adequação do vocabulário a ser utilizado na tradução dos testes.

Já Ernani Lopes vinculou o começo do trabalho à criação do laboratório de Psicologia da Liga Brasileira de Higiene Mental em 1926. Em 1927, realizou uma conferência na sede da Liga explorando as traduções do Stanford-Binet realizadas tanto no Brasil quanto em outros países (Lopes, 1931a). Ao contrário das publicações de Pernambuco, Lopes nomeou todos os seus artigos com um mesmo título: "Subsídios à adaptação brasileira dos testes de Binet-Terman". Durante o processo de adaptação, Idalina Abreu Fialho Nascimento Gurgel¹³ aplicou os testes na escola primária da Pequena Cruzada¹⁴ e na Escola Deodoro. Nesta última, trabalhou em conjunto com Maria Brasília Leme Lopes (*O Laboratório...*, 1932, p. 100-101).

Aqui, cabe também dizer que o trabalho publicado por Ernani Lopes era considerado por ele mesmo uma versão preliminar, o que justifica a escolha do título de seus artigos como "subsídios à adaptação". Grande parte do trabalho de estalonamento posterior ficou a cargo de M. B. Leme Lopes, que em 1932 tornou-se "psychologista-chefe" da recém-fundada Clínica de Eufrenia. Trabalhando com os assistentes da clínica, cabia a ela por exemplo "a organização completa do material de Binet-Terman"; o "aproveitamento dos dados já coligidos para a estalonagem da série Binet-Terman, adaptação de Ernani Lopes"; e a equivalência de duas séries de palavras que faziam parte de um dos testes desta escala (*O Laboratório...*, 1932, p. 103-104).

um tipo de intervenção de caráter médico-social, com indicações preventivas e eugênicas". Sobre a Liga, suas conexões com as ideias de "higiene mental" e de eugenia ver: Reis (1994), Stepan (2005) e Souza (2008).

- 12 O Instituto de Psicologia da Divisão estava subordinado à Seção de Higiene Mental. Para entender o funcionamento e a atuação de Ulysses Pernambucano neste espaço, cf. Vincentini (2018).
- 13 Além de membro da Liga, Idalina de Abreu Fialho Nascimento Gurgel era professora do Instituto Benjamin Constant e parte da equipe do Laboratório de Psicologia. De acordo com Castro (2014, p. 92), Idalina Gurgel também atuou como secretária do Conselho de Fiscalização das Expedições Artísticas e Científicas no Brasil (1933-1968).
- 14 Fundada em 1921 por "Laurita", filha de Epitácio Pessoa, a associação civil Pequena Cruzada de Santa Therezinha do Menino Jesus tinha como "fim principal de doutrinar a religião católica às crianças pobres e de promover-lhes por todos os meios bem-estar moral e material" (*Pequena Cruzada...*, p. 3, 12, s.d.). Na década de 1920, a escola primária da Pequena Cruzada funcionava no bairro do Catete, onde eram também desenvolvidas outras atividades assistenciais da associação, como as oficinas de costura e bordado e a distribuição de roupas e alimentos.

A adaptação de Ulysses Pernambucano (1930, p. 87) contou com a ajuda de “quatro auxiliares conhecedoras das técnicas, pacientes e habituadas – como professoras que são – a lidar com as crianças”. Além de Anita Paes Barreto, citada por Pernambucano em seu artigo, as outras professoras foram Cirene Coutinho e Alda Campos, que publicaram artigos discutindo as aplicações do Stanford-Binet nos *Arquivos da Assistência a Psicopatas de Pernambuco* (Coutinho, 1933; Pernambucano e Campos, 1931; Barreto e Campos, 1934). Além delas, autoras como Annita Pereira da Costa e Maria Leopoldina de Oliveira aparecem registradas como auxiliares do Instituto de Psicologia nos *Arquivos*, de modo que é possível inferir que, nos anos seguintes, mais de quatro professoras tenham trabalhado com as aplicações.

Antes de passar aos testes em si, é importante destacar a diferença entre as possibilidades de análise das fontes relacionadas aos trabalhos que vamos comparar aqui. Como um trabalho preliminar, Ernani Lopes fez menos modificações que Ulysses Pernambucano em relação aos critérios de seleção dos testes e ao reordenamento das questões. A última parte da adaptação de Ernani Lopes foi publicada em 1932, com as perguntas correspondentes ao nível de um “adulto médio”. Embora encerre a publicação com a promessa de que “concluirá proximamente” (Lopes, 1932, p. 70), não encontramos mais artigos sobre o tema nos *ABHM*. Ainda assim, é possível relacionar a versão publicada por Ernani Lopes nos *Arquivos* com a que foi usada na Clínica de Eufrenia (*A clínica...*, 1932, p. 74) e a que foi encontrada no arquivo pessoal de Alda Lodi, em Belo Horizonte.

De todo modo, já nos seus “subsídios”, Ernani Lopes flexibilizou os critérios de correção originais, uma lógica que também foi seguida por outros pesquisadores, como Isaías Alves e Noemy Rudolfer em suas experiências (Rocha, 2020). Em contraste, a adaptação coordenada por Ulysses Pernambucano modificou a ordem das questões e eliminou testes específicos em virtude dos resultados obtidos no processo de estalonamento. A diferença que nos interessa aqui não é a opção pela mudança ou manutenção das questões, mas o modo como estes pesquisadores interpretaram os testes, explicaram suas decisões e o desempenho das crianças nessas avaliações.

Para isso, a ideia de “estalonamento” é chave para entender o trabalho que estava sendo empreendido por esses pesquisadores. De forma geral, cada teste era desenvolvido para ser resolvido por um sujeito de inteligência mediana de uma idade específica. Para estabelecer a adequação entre resolução positiva do teste e idade, os pesquisadores aplicavam experimentalmente essas questões em uma amostra não selecionada. Assim, a taxa de sucesso da amostra em um determinado teste determinava a adequação desse mesmo teste para aplicação em indivíduos dessa mesma idade. Por consequência, esse era também o processo que determinava o padrão a partir do qual as crianças testadas posteriormente seriam comparadas. Esse processo de adaptação dos testes a uma população local, mais complexo que a tradução das perguntas, é chamado de estalonamento.

Um dos problemas enfrentados pelos pesquisadores brasileiros derivava justamente dessa premissa. Os testes de Binet foram desenhados pressupondo que essa “taxa de sucesso” variasse entre cerca de 2/3 a 3/4 das crianças avaliadas durante o processo de estalonamento. Essa variação tornou-se alvo de debate nas adaptações do Binet-Simon, num processo descrito por Maria Brasília Lopes e Idalina Fialho (1929, p. 70) como “nacionalização” das escalas de inteligência. Segundo elas:

Por nacionalizar significamos a necessidade de rever a distribuição dos tests na escala, para colocá-los em correspondência real com a inteligência brasileira. Com provas meramente traduzidas, é impossível afirmar a superioridade, igualdade ou inferioridade mental da inteligência de nossos discípulos relativamente a dos países em que foram *estalonadas* tais provas. Nenhuma conclusão particular, nenhum prognóstico individual poderá ser fornecido por escalas dessa natureza (Lopes e Fialho, 1929, p. 70, grifo nosso).

A necessidade de “nacionalizar as escalas” não era, portanto, exclusiva do Brasil. As diversas adaptações dos testes de Binet passaram por processos de padronização e aferição estatística a partir de experiências locais. Embora as experiências de Binet tenham explicitado a taxa que considerava adequada para estabelecer que um teste pertencia a uma determinada idade, outros autores estabeleceram seus próprios parâmetros. Eles se aproximavam do recorte de Binet, mas procuravam ser mais exatos, variando entre 60% (Kuhlman) e 75% de aproveitamento (Goddard) como padrão. Quanto a isto, Lopes e Fialho (1929) não deixaram espaço para dúvidas, estabelecendo 75% de aproveitamento como padrão a ser seguido. As autoras também destacavam que este processo de adaptação foi também realizado por Terman, que concluiu “baseado em estatísticas levantadas, não haver correspondência exata entre o nível mental das crianças francesas e norte-americanas” (p. 71).

No entanto, essa diferença não foi tomada por Terman como um sinal de inferioridade ou superioridade das primeiras em relação às segundas. Para ele, a escala de Binet como um todo era “muito fácil nas faixas mais baixas e muito difícil nas mais altas” – um problema que já tinha sido apontado em outras investigações (Terman, 1916, p. 48). Portanto, sua versão da escala procurava “corrigir” um problema da escala original, e não apenas adaptá-la às crianças estadunidenses. Esta nem sempre foi a conclusão a que chegaram os pesquisadores brasileiros.

Em 1933, Cirene Coutinho, auxiliar do Instituto de Psicologia de Pernambuco, explorou os resultados obtidos no trabalho de adaptação dos testes “Binet-Simon-Terman” que foram originalmente classificados como adequados às idades de 3 e 4 anos. Em seu artigo, Coutinho (1933, p. 15) destacou a dificuldade do Instituto em conseguir reunir o número necessário de crianças para a experiência (“cinquenta de cada sexo, por idade”), afirmando que a solução encontrada foi a de realizar visitas domiciliares. Além disso, afirmou que as experiências foram “realizadas em todos os meios sociais, sem preocupação de seleção” (p. 16, grifo nosso).

Ao contrário de Terman, Coutinho não interpretou a necessidade de reordenamento dos testes da Stanford-Binet como um problema da escala original. Assim como Lopes e Fialho (1929), ela explicou esta mudança como um sinal da diferença encontrada entre as crianças brasileiras e estadunidenses, mas foi além. De acordo com a autora:

Por mais aperfeiçoada que nos parecesse a revisão de Terman, era materialmente impossível permanecerem os testes nas mesmas idades. A mudança nas *condições do meio* seria suficiente para justificar a variação da distribuição dos testes nas idades acrescida entre nós pelo fato de estendermos as experiências às *classes de nível social inferior* (Coutinho, 1933, p. 16, grifo nosso).

De forma específica, a associação que foi feita entre desempenho nos testes e classe no Brasil, que aqui se expressa a partir das ideias de “meio social” ou “nível social”, nem sempre implicava inferioridade intelectual dos mais pobres, mas estabelecia uma diferença. O teste cinco

da idade de 3 anos desenvolvido por Binet e mantido na escala de Terman é um exemplo. Nele, cabia à criança informar seu sobrenome.

Embora fosse um teste simples, todos os autores admitiam que existia a possibilidade de que a criança não respondesse ao teste simplesmente por ignorar esta informação aos 3 anos. Ainda assim, Terman (1916, p. 148) considerava que essa possibilidade era negligenciável, porque eram “extremamente raras” as crianças que “nunca escutaram seu nome de família”. Além disso, para ele, a pontuação do teste afetava pouco o desempenho global da criança em sua avaliação.

A associação entre classe e desempenho nos testes não era exclusiva das pesquisas aqui analisadas. Em *Tests* (1928), Albuquerque afirmava que

Binet tinha experimentado em escolas públicas francesas com *crianças que, pela sua maior parte, provinham de meios operários*. O resultado foi que, fazendo-se iguais experiências nos Estados Unidos *com alunos de todas as classes*, verificou-se que os *tests* de Binet para os primeiros anos eram muito fáceis (Albuquerque, 1928, p. 79, grifo nosso).

Se por um lado as crianças mais novas tinham maior facilidade de responder as questões de Binet, as mais velhas encontraram maior dificuldade em questões desenhadas para idades mais avançadas, como já mencionamos. Neste caso, o critério de classe não foi acionado por Albuquerque como explicação. Ele afirmava que era justamente para contemplar essas diferenças que servia o processo de estalonamento dos testes, e sua adequação para o público local.

Nas adaptações que examinamos, tanto Ernani Lopes quanto Cirene Coutinho mantiveram o teste na mesma ordem apresentada por Binet e Terman, mas fizeram ressalvas. A de Coutinho estava diretamente ligada aos recortes socioeconômicos, usados como explicação para o baixo desempenho das crianças testadas, que ficaram no limite da adequação (60,8% de aproveitamento). De acordo com ela:

Não atribuímos este resultado, senão à falta de oportunidade de aprender seu nome de família, ao pouco caso ou desinteresse dos pais e *principalmente porque levamos nossas pesquisas às classes humildes, onde o fracasso foi extremo*. Apoiamos Terman quanto à opinião de que embora a criança não tenha a oportunidade de aprender o seu nome de família como o seu prenome, ouvindo algumas vezes, se é de inteligência normal, o interesse que ele desperta, provoca a sua lembrança. Observamos este fato em criança de certo meio abastado, ao passo que era em geral completa a ignorância entre as de nível inferior (Coutinho, 1933, p. 19, grifo nosso).

De certo modo, o teste e a interpretação que faziam de seus resultados estavam conectados ao valor que o sobrenome de cada criança tinha, não só para elas, mas para sociedade de uma forma geral. É assim que explicavam que uma “criança de certo meio abastado” pudesse lembrar do seu sobrenome, enquanto uma criança “de nível inferior” ignorasse completamente esta informação. Estes últimos, embora com família, não tinham, simbolicamente, sobrenome.

Já no caso da adaptação de Ernani Lopes, suas observações estão registradas no item “avaliação”. Neste teste, Lopes adiciona à afirmação de Terman de que a criança não deveria ser penalizada por “defeitos de pronúncia” a seguinte observação: “considere-se correto o sobrenome do padrasto, da mãe ilegítima ou da pessoa de que a criança dependa como se fora filho

(Izaías Alves)” (Lopes, 1931a, p. 28). Neste trecho, Ernani Lopes nos remete à adaptação de Isaiás Alves (1932a) do teste de Binet, que partiu da adaptação de Cyril Burt (1922). Embora Terman não entre no mérito da questão familiar, o próprio Binet já observava em 1908 a possibilidade de existir uma divergência entre a resposta dada pela criança e a resposta esperada (que teria como base o seu registro de matrícula). De acordo com ele:

Às vezes acontece de a criança dar outro nome, diferente do que ela foi matriculada. Para explicar esses erros é necessário lembrar que existem muitas crianças ilegítimas e, o que é mais lamentável, de que a mãe da criança pode ter tido maridos diferentes, e o nome que a criança carrega nos anos seguintes nem sempre é o mesmo (Binet e Simon, 1916, p. 195).

Desta forma, Ernani Lopes não conecta a adequação da questão às expectativas relacionadas à recortes de classe, mas adiciona uma observação que sinaliza como nem todas as configurações familiares favoreciam o bom desempenho da criança no teste – mais um fator que não está relacionado à inteligência. Se considerassem a diferença cultural de tratamento usual no Brasil, onde raramente se chama as pessoas pelo sobrenome, o teste pareceria ainda mais inadequado. Ainda assim, os pesquisadores brasileiros que mapeamos optaram por manter a pergunta em suas adaptações. Seja Isaiás Alves (1932a) em 1926 a partir da escala Binet-Burt, seja Ernani Lopes¹⁵ e Cirene Coutinho a partir da Stanford-Binet.

A opção pela manutenção deste teste é também expressão dos critérios escolhidos para interpretar os resultados e apontar as crianças mais pobres como responsáveis pelo baixo desempenho das crianças brasileiras, se comparadas às crianças de outros países. Em contrapartida, a opção pela eliminação de outras questões nos mostra como era possível adotar outra postura, desde que considerassem o motivo relevante. Exemplo disso encontra-se no artigo de Ulysses Pernambucano e Alda Campos (1931) sobre os resultados obtidos no teste por eles traduzidos como “A bola e o campo”.¹⁶ De forma resumida, o teste consistia em pedir à criança que ela desenhasse o caminho que faria para procurar uma bola perdida em um campo circular, escondida pela grama, usando uma figura de apoio (Figura 1).

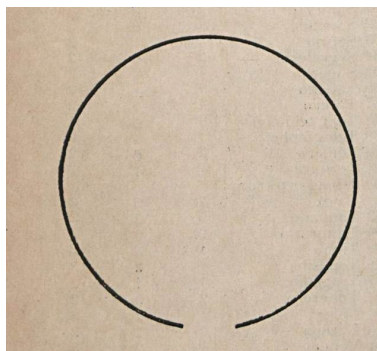


Figura 1: Teste “A bola e o campo”

Fonte: Pernambucano e Campos (1931, p. 34).

15 Na versão usada na Clínica de Eufrenia, o teste foi mantido na escala, na idade de 8 anos. Cf. A clínica..., 1932.

16 Teste VIII.1 na escala original de Terman. Nesta versão, uma bola de *baseball* estaria perdida em um campo circular. Ver: Terman (1916, p. 210-212). A ideia de uma bola inespecífica um campo gramado já é uma adaptação das traduções brasileiras, geralmente justificadas pela falta de familiaridade dos brasileiros com o esporte.

O exame foi uma adição de Terman ao Binet-Simon e contava com critérios diferentes de avaliação para as idades de 8 e 12 anos. As respostas eram qualificadas gradualmente e tipificadas em quatro possibilidades,¹⁷ que sinalizavam maturidades diferentes. Terman argumentava que o resultado deste teste expressava habilidades do indivíduo que estariam ligadas ao “senso prático” (Terman, 1916, p. 213; Lopes, 1931b, p. 100). Ele justificava sua adição na escala como um dos recursos utilizados para compensar a ênfase excessiva que o Binet-Simon daria ao “pensamento abstrato e à compreensão verbal” (Terman, 1916, p. 212-213).

Em sua adaptação, Ernani Lopes (1931b) relativizou as críticas posteriores feitas ao teste, que se relacionavam à clareza das instruções fornecidas, aos critérios de correção e ao resultado de pesquisas desenvolvidas posteriormente sobre retentividade e memória. Também registrou que, ao adaptar a ideia original, “já houve quem sugerisse a substituição de uma bola por uma ‘bolsa com dinheiro’” (Lopes, 1931b, p. 101), de modo que o incentivo à resolução fosse maior.

Ulysses Pernambucano e Alda Campos (1931) aplicaram este teste em 422 estudantes de 12 e 13 anos, usando como critério inicial de sucesso o tipo D de resolução, estabelecido por Terman para crianças de 12 anos. Como resultado, apenas 33% dos estudantes de 12 anos foram considerados aptos a resolvê-lo. A taxa subia um pouco mais aos 13 anos, com 44% de aprovação. Os autores ressaltaram que, mesmo quando consideraram como chave de correção o tipo C de resolução, estabelecido originalmente como parâmetro para crianças de 8 anos, não conseguiram obter a porcentagem mínima para considerar o teste adequado.¹⁸ Nos artigos publicados posteriormente (Barreto e Campos, 1934; Barreto, 1935) é possível perceber que este foi um dos dez testes eliminados pela versão pernambucana, de um total de noventa da escala original (Barreto, 1935, p. 185).

Na opinião de Pernambucano e Campos (1931, p. 36), a explicação para os índices obtidos não recaía sobre os indivíduos testados, tampouco sobre a qualidade da questão. Como um teste que pretendia medir o “juízo prático”, o que os números expressavam estava relacionado à “defeituosa educação de nossas crianças, quer no meio familiar, quer no escolar”. Eram um sinal do nosso “verbalismo”. Na opinião dos autores, os resultados deveriam, portanto, servir como indicativo de que os princípios dos “métodos ativos de educação” deveriam ser perseguidos. Por fim, repetiam o jargão de que “a escola deve preparar para a vida” (p. 36).

Aqui cabe lembrar que os autores discutiam uma questão de um teste de inteligência, ferramenta de avaliação que tinha como premissa a ideia de que o peso da escolarização não deveria definir o desempenho do sujeito avaliado. Assim como no caso do teste do sobrenome ou “nome de família”, os resultados obtidos com o teste “A bola e o campo” não estavam conectados à inteligência.

Embora fossem bastante inferiores ao esperado, o percentual de sucesso estabelecido inicialmente para este teste em 1931 foi o de 75% de aprovação. Em 1933, se tivessem usado o mesmo parâmetro, o teste do “nome de família” também não seria mantido na escala pernambucana, já que apenas 60,8% das crianças responderam à questão corretamente. Como vimos, o percentual que determinava a adequação de um teste não era um consenso entre os diversos

17 Os tipos A e B eram considerados como um fracasso da criança em resolver a questão. Já o tipo C era considerado satisfatório para a idade de 8 anos e o tipo D era considerado satisfatório para a idade de 12 anos.

18 Neste caso, o resultado foi melhor, com 54,9% de aprovação dos alunos com 12 anos e 58,25% de aprovação dos alunos de 13 anos.

autores envolvidos no debate. No entanto, a variação no critério dos pesquisadores pernambucanos é digna de nota.

Em 1932, Alda Campos publicou um artigo sobre a "Revisão da Escala Binet-Simon-Terman" para as idades de 7 e 8 anos nos *Archivos*. Ali, discutiu-se novamente o critério a ser adotado na seleção das questões. De acordo com ela, o percentual estabelecido "para que um *test* seja devidamente colocado num grupo de idade" variava de acordo com os autores. Além disso, afirmava que "na última revisão de Terman, nenhuma percentagem definida é proposta", embora em tentativa anterior Terman e Childs tenham sugerido um percentual de 66%. Em seguida, Alda Campos (1932) explicita o critério adotado "provisoriamente" para a revisão pernambucana: "65% a 75%, só admitindo percentagens mais baixas quando as encontradas na idade imediatamente superior fossem consideravelmente elevadas", mas ressaltou que nem sempre esse critério era "rigorosamente observado" (p. 85).

Em 1933, Cirene Coutinho (1933, p. 17) retomou a discussão sobre o percentual de aproveitamento a ser considerado, mas dessa vez mencionando que "o próprio Terman em seu livro *The measurement of intelligence*, p. 226, declara que adotou a percentagem de 60% a 65% para estabelecer os níveis mentais das idades de 8, 10, 12 e 14 anos no teste de vocabulário". Em seguida, afirmou que, estatisticamente, os percentuais de 60% a 75% foram considerados pelo Instituto como os mais adequados para a realidade local, sinalizando uma nova flexibilização do critério de adequação.

Quando explorou os testes estabelecidos para a idade de 3 anos, Coutinho (1933, p. 19) reconheceu que a porcentagem de sucesso no teste do nome de família era "relativamente fraca, uma vez que marca quase o nível inferior do limite estipulado". Ainda assim, justificou a sua manutenção na escala, descartando a importância da diferença encontrada, ainda que tenha notado que a questão prejudicava majoritariamente um grupo específico: o de alunos mais pobres. Para isso, retomava Terman, que minimizava o prejuízo do "examinando" que fosse "injustamente penalizado por lhe faltar a oportunidade de adquirir o conhecimento que o teste requer" (p. 20). Nos anos seguintes, é possível notar que a taxa de 60% passou a ser aceita como padrão mínimo para definir a permanência de outros testes na versão pernambucana.

De certo modo, eram justamente essas diferenças que justificavam a avaliação final de que eram as crianças pobres que diminuía os índices gerais de desempenho das crianças testadas. Entre o processo de estalonamento e o estabelecimento da ferramenta de avaliação, a inadequação do teste passava a ser desconsiderada, enquanto o baixo desempenho de perfis específicos era naturalizado como parte dos "efeitos do meio". Depois, os professores, tendo em mãos um instrumento de avaliação já considerado adaptado à população local, dificilmente questionariam a adequação da ferramenta, mesmo diante dos resultados obtidos. Foi neste sentido que o trabalho de adaptação destes pesquisadores de testes como o Stanford-Binet reforçou expectativas, ao invés de subvertê-las.

Considerações finais

Neste artigo, demonstramos o papel ativo dos pesquisadores brasileiros na adaptação e construção de testes de inteligência que considerassem as particularidades locais. Neste sentido, exploramos o modo pelo qual pesquisadores como Ernani Lopes e Ulysses Pernambucano não

apenas traduziram testes estrangeiros como o Stanford-Binet, mas também refletiram sobre seu formato e adequação a partir de aplicações experimentais e de uma participação efetiva no debate que se desenrolava sobre as versões já existentes.

Ao longo da pesquisa, foi fundamental examinar o processo de estalonamento desses testes. Essa análise nos possibilitou explorar em que medida concepções relacionadas a um debate mais amplo sobre a “questão nacional/social” (Stepan, 2005) se refletiu nas decisões tomadas por estes autores. Os artigos que discutem essas adaptações – publicados nos *Archivos Brasileiros de Higiene Mental* e nos *Arquivos da Assistência a Psicopatas de Pernambuco* – sinalizam que, diante da dificuldade que encontraram em adaptar questões específicas, havia um horizonte de escolhas possíveis. São justamente essas possibilidades que nos mostram como os pesquisadores fizeram escolhas situadas e conectadas a convicções da época, como a ideia de que crianças mais pobres eram naturalmente menos inteligentes do que crianças mais ricas. Sem esta convicção, não seria possível manter na escala testes que claramente prejudicavam grupos específicos, como é o caso do teste do “nome de família”.

De certo modo, é também por isso que alguns dos resultados das experiências brasileiras confirmaram as afirmações de autores estadunidenses como Lewis Terman a respeito da interseção entre inteligência e recortes socioeconômicos como o que mostramos aqui, ou ainda recortes de raça e gênero que vemos em outras experiências. Mais do que uma adequação de um discurso à realidade brasileira, esses pesquisadores conheciam os efeitos de suas aplicações, ainda que nem sempre enfatizem os dilemas que enfrentaram enquanto adaptavam os testes. Assim, elaboravam versões que guardariam as marcas de suas escolhas que, resguardadas pela devida aferição estatística, tornavam-se o padrão pelo qual se definiam os mais e os menos inteligentes.

Agradecimentos

Esta pesquisa foi financiada pela Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro - Faperj, com recursos da bolsa de Pós-doutorado Faperj Nota 10 (processo n. E-26/204.303/2021).

Referências bibliográficas

A CLÍNICA de Eufrenia. *Arquivos Brasileiros de Higiene Mental*, v. 5, n. 2, p. 65-98, 1932.

ACCORSI, G.E. Ernani Lopes. Médicos que atuaram no Hospital Nacional de Alienados (blog). In: Biblioteca Virtual em História do Patrimônio Cultural da Saúde, 2018. Disponível em: <https://hpcs.bvsalud.org/wp-content/uploads/2018/07/Lopes-Ernani.pdf> Acesso em: 9 abr. 2024.

ALBUQUERQUE, J.J.C.C. *Tests*. Rio de Janeiro: Francisco Alves, 1928.

ALVES, I. *Teste individual de inteligência*. 3. ed. Rio de Janeiro: Typographia d'A Encadernadora, [1926] 1932a.

ALVES, I. Teste collectivos de intelligencia (Terman group test) e a sua aplicação nas escolas públicas. *Boletim de Educação Pública*, v. 2, n. 3 e 4, p. 397-435, 1932b.

- ALVES, I. A seleção dos super-normais. *Revista do Ensino*, n. 101, p. 66-72, 1934.
- ANITA Paes Barreto. *Psicologia: Ciência e Profissão*, v. 20, n. 3, p. 49, 2000. Disponível em: <https://doi.org/10.1590/S1414-98932000000300009>. Acesso em: 27 mar. 2024.
- ANTIPOFF, H. A criança bem dotada. *Revista do Ensino*, n. 176, p. 21-25, 1946.
- BARRETO, A. Revisão pernambucana da escala Binet-Simon-Terman. *Arquivos da Assistência a Psicopatas de Pernambuco*, v. 5, n. 1 e 2, p. 164-188, 1935.
- BARRETO, A.; CAMPOS, A. Revisão pernambucana da escala Binet-Simon-Terman. *Arquivos da Assistência a Psicopatas de Pernambuco*, v. 4, n. 1, p. 1-18, 1934.
- BINET, A.; SIMON, T. *The development of intelligence in children (The Binet-Simon scale)*. Baltimore: Williams & Wilkins, 1916.
- CAMPOS, A. Revisão da escala Binet-Simon-Terman. *Arquivos da Assistência a Psicopatas de Pernambuco*, v. 2, n. 1, p. 84-87, 1932.
- CAMPOS, R.H.; BORGES, A.A.P. De Genève à Belo Horizonte, une histoire croisée: circulation, réception et réinterprétation d'un modèle européen des classes spéciales au Brésil des années 1930. *Paedagogica Historica*, v. 50, p. 195-212, 2014.
- CARSON, J. *Measure of merit: talents, intelligence, and inequality in the French and American Republics, 1750-1940*. New Jersey: Princeton University Press, 2007.
- CASTRO, R.D. *A sublimação do id primitivo em ego civilizado: o projeto dos psiquiatras-psicanalistas para civilizar o país (1926-1944)*. Tese (Doutorado em História das Ciências e da Saúde) – Casa de Oswaldo Cruz/Fiocruz, Rio de Janeiro, 2014.
- CERQUEIRA, E.C.B. *A Sociedade Brasileira de Neurologia, Psiquiatria e Medicina Legal: debates sobre ciência e assistência psiquiátrica (1907-1933)*. Dissertação (Mestrado em História das Ciências e da Saúde) – Casa de Oswaldo Cruz/Fiocruz, Rio de Janeiro, 2014.
- CHAPMAN, P.D. *Schools as sorters: Lewis M. Terman, applied psychology, and the intelligence testing movement, 1890-1930*. New York: New York University Press, 1988.
- COUTINHO, C. Contribuição à revisão pernambucana da escala de Binet-Simon-Terman. *Arquivos da Assistência a Psicopatas de Pernambuco*, v. 3, n. 1, p. 15-26, 1933.
- DÁVILA, J. *Diploma de brancura*. São Paulo: Editora Unesp, 2006.
- [DOCUMENTOS referentes ao Instituto de Educação]. Arquivo pessoal de Lourenço Filho, série temática, classificação LF t Inst. Educ. Rio de Janeiro: CPDOC/Fundação Getúlio Vargas, 1927-1964.
- FACCHINETTI, C.; MOTA, A.; MUÑOZ, P. A higiene mental no Brasil: racismo, eugenia e infância no Rio de Janeiro (anos 1920 a 1960). In: CAMPOS, R.; RUPERTHUZ, M. (org.). *Higiene mental, psiquiatria y sociedad en Iberoamérica (1920-1960)*. Madrid: La Catarata, 2022. p. 20-30.
- FASS, P. The IQ: a cultural and historical framework. *American Journal of Education*, v. 88, n. 4, p. 431-458, 1980.
- FROTA, Z. Crianças super-normais através do estudo de L. Terman. *Revista do Ensino*, v. 7, n. 97, p. 7-25, 1933.
- GOULD, S.J. *A falsa medida do homem*. São Paulo: Martins Fontes, 2018.
- HACKING, I. *The taming of chance*. Cambridge: Cambridge University Press, 2013.
- JACÓ-VILELA, A.M. Psychological measurement in Brazil in the 1920s and 1930s. *History of Psychology*, v. 17, n. 3, p. 237-248, 2014. Disponível em: <https://doi.org/10.1037/a0035333>. Acesso em: 27 mar. 2024.
- JACÓ-VILELA, A.M.; SILVA, V. Medición psicológica en Brasil: la recepción particular de las técnicas de evaluación. In: BARRERA, R. *História local de la psicología*. Santiago: Ediciones Universidad Santo Tomás, 2016. p.125-144.
- JACÓ-VILELLA, A.M. et al. Psicologia na psiquiatria: testes psicológicos no pavilhão de observação do Hospício Nacional de Alienados (1907-1925). In: VENANCIO, A.T.; DIAS, A.T. (org.). *O Hospício da Praia Vermelha: do*

- Império à República (Rio de Janeiro, 1852-1944). Rio de Janeiro: Editora Fiocruz, 2022.
- LEVINE, P. *Eugenics: a very short introduction*. New York: Oxford University Press, 2017.
- LOPES, E. Subsídios à adaptação brasileira dos testes de Binet-Terman. *Archivos Brasileiros de Higiene Mental*, v. 4, n. 1, p. 17-52, 1931a.
- LOPES, E. Subsídios para a adaptação brasileira dos testes de Binet-Terman (continuação). *Archivos Brasileiros de Higiene Mental*, v. 4, n. 2, p. 97-132, 1931b.
- LOPES, E. Subsídios à adaptação brasileira dos testes de Binet-Terman (continuação). *Archivos Brasileiros de Higiene Mental*, v. 5, n. 1, p. 39-70, 1932.
- LOPES, M.B.L.; FIALHO, I.A. Sugestões para o emprego dos tests. *Archivos Brasileiros de Higiene Mental*, v. 2, n. 2, p. 68-77, 1929.
- MINTON, H.L. *Lewis M. Terman: pioneer in psychological testing*. New York: New York University Press, 1988.
- MONARCHA, C. Testes ABC: origem e desenvolvimento. *Boletim: Academia Paulista de Psicologia*, v. 28, n. 1, p. 7-17, 2008.
- O LABORATÓRIO de psicologia. *Archivos Brasileiros de Higiene Mental*, p. 99-107, 1932.
- PATTO, M.H.S. *A produção do fracasso escolar: histórias de submissão e rebeldia*. 4. ed. São Paulo: Casa do Psicólogo, 2015.
- PATTO, M.H.S. Ciência e política na Primeira República: origens da psicologia escolar. *Mnemosine*, v. 1, n. especial, p. 203-225, 2004.
- PATTO, M.H.S. Estado, ciência e política na Primeira República: a desqualificação dos pobres. *Estudos Avançados*, v. 13, n. 35, p. 167-198, 1999.
- PEQUENA CRUZADA. Pequena Cruzada: 100 anos de história, s.d. p. 2-29. Disponível em: https://www.pequenacruzada.org.br/_files/ugd/47aff8_38063583cf6341d9b345a0546f398e34.pdf. Acesso em: 27 mar. 2024.
- PERNAMBUCANO, U. A psicologia em Pernambuco. *Archivos Brasileiros de Higiene Mental*, v. 3, n. 3, p. 85-90, 1930.
- PERNAMBUCANO, U.; BARRETO, A. P. O vocabulário das crianças das escolas primárias do Recife. *Arquivos da Assistência a Psicopatas de Pernambuco*, v. 1, n. 1, p. 3-33, 1931.
- PERNAMBUCANO, U.; CAMPOS, A. O teste "A bola e o campo" em crianças de 12 e 13 anos. *Arquivos da Assistência a Psicopatas de Pernambuco*, v. 1, n. 1, p. 34-36, 1931.
- PORTAVALES, V.; JACÓ-VILLELA, A.M. Maria Brasília Leme Lopes (1909-1996). In: CNPq. *Pioneiras da ciência no Brasil*: 6. ed. 2016. Disponível em: <https://www.gov.br/cnpq/pt-br/aceso-a-informacao/acoes-e-programas/programas/mulher-e-ciencia/pioneiras-da-ciencia-1/pioneiras-6a-edicao>. Acesso em: 3 abr. 2024.
- PORTER, T. *The rise of statistical thinking, 1820-1900*. New Jersey: Princeton University Press, 2020.
- PROF. ULYSSES Pernambucano. *Arquivos de Neuro-Psiquiatria*, v. 2, n. 1, p. 3-4, 1944. Disponível em: <https://doi.org/10.1590/S0004-282X1944000100001>. Acesso em: 27 mar. 2024.
- RECIFE. Projeto de lei ordinária n. 1.335/2020. Declara Anita Paes Barreto como patrona da Psicologia em Pernambuco. Recife: Assembleia Legislativa de Pernambuco, 2020. Disponível em: <https://www.alepe.pe.gov.br/proposicao-texto-completo/?docid=6195&tipoprop=p>. Acesso em: 27 mar. 2024.
- REIS, J.R.F. *Higiene mental e eugenia: o projeto de regeneração nacional da Liga Brasileira de Higiene Mental (1920-30)*. Dissertação (Mestrado em História) – Universidade Estadual de Campinas, Campinas, 1994.
- ROCHA, A.C.S.M. *O que fazer com os rudes? Isaías Alves e as divergências sobre o papel da inteligência na organização escolar (1930-1942)*. Dissertação (Mestrado em História, Política e Bens Culturais) – Fundação Getúlio Vargas, Rio de Janeiro, 2011.
- ROCHA, A.C.S.M. *Entre o Brasil e os Estados Unidos: intelectuais, ideias e projetos de educação (1927-1935)*. Rio de Janeiro: Editora Fiocruz, 2020.

- ROCHA, A.C.S.M. Isaías Alves. In: JACÓ-VILELA, A.M.; KLAPPENBACH, H.; ARDILA, R. (ed.). *The Palgrave biographical encyclopedia of psychology in Latin America*. Cham: Palgrave Macmillan, 2023. p. 38-40. Disponível em: https://doi.org/10.1007/978-3-030-56781-1_562. Acesso em: 27 mar. 2024.
- ROTA JR., C. *Recepção e circulação de testes de inteligência na escola de aperfeiçoamento de professores de Belo Horizonte (1929-1946)*. Tese (Doutorado em Educação) – Universidade Federal de Minas Gerais, Belo Horizonte, 2016.
- SATURNINO, M. Psychometria de 100 soldados pelos testes de Binet. *Archivos Brasileiros de Higiene Mental*, v. 3, n. 1, p. 12-17, 1930.
- SOUZA, V.S. Por uma nação eugênica: higiene, raça e identidade nacional no movimento eugênico brasileiro dos anos 1910 e 1920. *Revista Brasileira de História da Ciência*, v. 1, n. 2, 2008. Disponível em: <https://doi.org/10.53727/rbhc.v1i2.397>. Acesso em: 27 mar. 2024.
- SOUZA-CHALOPA, R.F. A contribuição dos estudos sobre grupos escolares para a historiografia da educação brasileira: reflexões para debate. *Revista Brasileira de História da Educação*, v. 19, p. 1-24, 2019. Disponível em: <https://doi.org/10.4025/rbhe.v19.2019.e063>. Acesso em: 27 mar. 2024.
- STEPAN, N. *A hora da eugenia: raça, gênero e nação na América Latina*. Rio de Janeiro: Editora Fiocruz, 2005.
- STERN, A. *Eugenic nation: faults and frontiers of better breeding in modern America*. Berkeley: University of California Press, 2005.
- TERMAN, L.M. *The measurement of intelligence*. Cambridge: Riverside Press, 1916.
- TERMAN, L.M. Intelligence and its measurement: a symposium – II. *Journal of Educational Psychology*, v. 12, n. 3, p. 127-133, 1921.
- VIALLE, W. Terminal science? The work of Lewis Terman revisited. *Roepers Review*, v. 17, n. 1, p. 32-38, 1994.
- VINCENTINI, R. *Entre sanatórios e terreiros: Ulysses Pernambuco, René Ribeiro e o projeto reformista da psiquiatria social de Recife (1910-1940)*. Dissertação (Mestrado em História das Ciências e da Saúde) – Casa de Oswaldo Cruz/Fiocruz, Rio de Janeiro, 2018.
- WINSTON, A.S. Scientific racism and North American psychology. In: *The Oxford encyclopedia of psychology*, 2020. Oxford: Oxford University Press, 2020. Disponível em: <https://doi.org/10.1093/acrefore/9780190236557.013.516>. Acesso em: 27 mar. 2024.

Recebido em setembro de 2023

Aceito em novembro de 2023