UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

**Avaliação dos instrumentos de mensuração de
competências socioemocionais no contexto
escolar**

Luan Pires Paciência

Orientador: Prof. Dr. Ricardo de Abreu Madeira

São Paulo - Brasil
2016

Prof. Dr. Marco Antonio Zago
Reitor da Universidade de São Paulo

Prof. Dr. Adalberto Américo Fischmann
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Hélio Nogueira da Cruz
Chefe do Departamento de Economia

Prof. Dr. Márcio Issao Nakane
Coordenador do Programa de Pós-Graduação em Economia

LUAN PIRES PACIÊNCIA

**Avaliação dos instrumentos de mensuração de
competências socioemocionais no contexto
escolar**

Dissertação apresentada ao Departamento de
Economia da Faculdade de Economia, Admini-
stração e Contabilidade da Universidade
de São Paulo como requisito parcial para a
obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. Ricardo de Abreu Madeira

Versão Corrigida

(versão original disponível na Biblioteca da Faculdade de Economia, Administração e
Contabilidade)

São Paulo - Brasil

2016

FICHA CATALOGRÁFICA

Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Paciência, Luan Pires

Avaliação dos instrumentos de mensuração de competências socioemocionais no contexto escolar / Luan Pires Paciência. -- São Paulo, 2016.

104 p.

Dissertação (Mestrado) – Universidade de São Paulo, 2016.
Orientador: Ricardo de Abreu Madeira.

1. Economia – Aspectos psicológicos 2. Questionário de personalidade 3. Aplicação de testes em grupos 4. Competências socioeconômicas 5. Habilidades não cognitivas 6. Experimento aleatório I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

CDD – 330.1915

Agradecimentos

Agradeço aos meus pais por me apoiarem em todos os sentidos e me encorajarem nas minhas decisões. Agradeço ao Guilherme por também me apoiar, dar suporte nesses anos de mestrado e ter aturado meus não raros momentos de estresse.

Agradeço às pessoas que possibilitaram a realização do experimento: à ONG Parceiros da Educação, especialmente à Lúcia Favero por fazer o contato com as escolas; à equipe de coordenação das 10 escolas parceiras; aos alunos que participaram do experimento; à FIPE que financiou o experimento, ao Instituto Ayrton Senna que permitiu a utilização das versões traduzidas dos instrumentos; e à equipe de aplicadores que topou desde o começo encarar o desafio de entrar nas salas de aula e conquistar a atenção dos alunos.

Agradeço ao Alex e à Deborah que me ajudaram muito na semana crítica do experimento e à Luíza que me deu acesso a um dos trabalhos mais relevantes para o conteúdo dessa pesquisa. Agradeço também à Maína e à Priscilla que me ajudaram a manter o foco e por serem conselheiras nos momentos chaves da dissertação.

Às pessoas que me ajudaram a tomar a decisão pelo mestrado lá em 2013: Marina, Felipe, Neide, Marcela e Paula.

Ao meu orientador Ricardo Madeira, que me deu autonomia no trabalho, me fazendo ir mais longe com minhas próprias pernas do que achei que fosse possível, além das sugestões de melhorias ao longo do trabalho juntamente com os professores Daniel Santos e Marcos Rangel. À CAPES pelo apoio financeiro ao longo dos dois anos de mestrado.

Agradeço também aos meus colegas de mestrado e aos amigos em geral por terem tornado esses anos de mestrado muito mais leves e agradáveis com suas companhias.

Resumo

Essa dissertação tem como objetivo analisar o comportamento manipulador relacionado à aplicação de instrumentos de mensuração de competências socioemocionais no ambiente escolar estimulado por duas abordagens que configuram contextos *low-stake*: o efeito da disseminação da informação sobre essas competências e suas relações com bem-estar e progresso social e da possibilidade dos professores receberem as respostas individuais dos alunos. Para isso, realizou-se um experimento com 2853 alunos de 10 escolas públicas da rede estadual de São Paulo. Os resultados indicam que os alunos são capazes de inflar suas respostas no sentido da desejabilidade social e que as duas abordagens fazem com que os alunos se sintam mais motivados em manipular suas respostas. Além disso, os resultados mostram que a inclusão de uma notificação sobre um mecanismo de detecção de respostas falsas no questionário cria uma barreira e desestimula o comportamento manipulador.

Palavras-chaves: questionário de personalidade, aplicação de testes em grupos, competências socioemocionais, habilidades não cognitivas, experimento aleatório, economia da personalidade

Abstract

This dissertation aims to examine faking behaviour in personality assessments at schools. Two approaches that create low-stakes environments are investigated: information spreading about non-cognitive skills and their impact on well-being and social progress and the fact that teachers would receive students' individual answers. Using a CRT technique with 2853 students from 10 São Paulo state public schools, we find that students are capable to inflate their responses and these two approaches turn them more motivated to engage in faking. Moreover, we show that a warning about untruthful answers attached to the questionnaire imposes a cost, discouraging faking behaviour.

Key-words: personality assessments, randomized controlled trials, non-cognitive skills, economics of personality

Sumário

1	Introdução	11
2	O experimento	17
2.1	Construção do instrumento	17
2.1.1	Big Five Inventory	19
2.1.2	Strengths and Difficulties Questionnaire	20
2.1.3	Self-Efficacy Questionnaire for Children	20
2.1.4	Nowicki-Schickland Locus of Control Scale	21
2.1.5	Rosenberg Self-Esteem Scale	22
2.1.6	Core Self Evaluation Scale	22
2.1.7	Short Grit Scale	23
2.1.8	Escala de Desejabilidade Social	23
2.1.9	Cognitive Reflection Test	24
2.1.10	Raven's Standard Progressive Matrices	24
2.1.11	Questionário socioeconômico	25
2.2	Escolha das abordagens	25
2.2.1	Disseminação da informação	26
2.2.2	Professores receberem as respostas individuais	27
2.2.3	Notificação	28
2.3	Estratégia de identificação	29
2.3.1	Análise dos itens	30
2.3.2	Análise dos constructos	33
2.3.3	Análise da consistência interna	33
2.3.4	Análise das correlações	34
3	Amostra	37
4	Resultados	47
4.1	Efeito da informação	47
4.2	Efeito da notificação sobre a informação	50
4.3	Efeito dos professores receberem os resultados	56

4.4	Efeito da Notificação sobre os professores receberem os resultados	60
5	Conclusão	65
	Referências	69
	Apêndice A - Detalhes do experimento	78
	Apêndice B - Distribuição dos alunos entre grupos e questionários por escola .	79
	Apêndice C - Testes de balanceamento	89
	Apêndice D - Resultados via análise de regressão	89
	Apêndice E - <i>Heat maps</i> - Itens	91
	Apêndice F - Questionários	97

1 Introdução

Ao longo das últimas décadas, evidências acerca da importância das competências socioemocionais vêm se acumulando, deixando cada vez mais claro o quanto elas podem ser relevantes para explicar o bem-estar individual e progresso social das economias. Inúmeros trabalhos documentaram a importância das competências socioemocionais em áreas como educação, mercado de trabalho, crime, saúde, entre outros (Hampson et al. (2007); John et al. (1994); Vazsonyi et al. (2001); Almlund et al. (2011); Lounsbury et al. (2004); Santos (2014); Roberts et al. (2007); Duncan e Magnuson (2011); Störmer e Fahr (2013); Kern e Friedman (2008)). Além disso, trabalhos como Roberts, Walton e Viechtbauer (2006) e Almlund et al. (2011) apontam que essas competências são maleáveis, ou seja, podem ser estimuladas e em um nível maior, inclusive, que as habilidades cognitivas dependendo da faixa etária dos indivíduos.

Diante desses fatos, é natural que esse tema tenha ganhado espaço e importância nos debates sobre educação, e como consequência, que surjam políticas públicas no sentido de estimulá-las. Um exemplo notável nesse sentido teve início na Califórnia nos anos de 2014 e 2015 em que nove distritos iniciaram a mensuração dessas competências em alunos para utilizá-las como métricas de avaliação de performance das escolas em um contexto de *accountability*. Contudo, Duckworth e Yeager (2015) questionam a adequação dessas medidas para esses propósitos, tendo como um dos motivos preocupantes o fato de os alunos poderem inflar suas respostas nos questionários de autorelato por razões próprias ou serem coagidos a isso para melhorar sua avaliação como aluno ou das escolas.

Chama-se de comportamento manipulador essa alteração sistemática nas respostas e viesadas no sentido que o indivíduo considera mais adequado para atingir algum objetivo por ele desejado (Ziegler, MacCann e Roberts (2011)). Esse tipo de comportamento é um problema para políticas de *accountability* uma vez que essas medidas podem capturar a capacidade em inflar as competências e não propriamente as competências reais, criando sérios problemas para a avaliação dessas escolas. Porém, esse comportamento mostra-se problemático não apenas em contextos de *accountability*. Em outros muito menos

pretensiosos, como por exemplo de simples mapeamento dessas competências em qualquer que seja a unidade de observação, é necessário que os instrumentos sejam boas medidas das competências reais dos indivíduos para que se tenha um panorama fiel e justique o emprego de recursos nesse sentido.

É possível observar que a preocupação com esse comportamento é bastante antiga. Já na década de 40 o *Minnesota Multiphasic Personality Inventory* (MMPI) (McKinley e Hathaway (1943)) contava com um protocolo para detectar respostas distorcidas dos indivíduos. Obviamente esse problema é presente quando são os próprios alunos que reportam suas competências nos chamados questionários de autorelato, mas essa não é a única forma de se mensurar essas competências. Existem meios alternativos como entrevistas, dinâmicas, questionários de hetero-relato, onde agentes próximos aos indivíduos respondem sobre eles, entre outros. Contudo, os questionários de autorelato são a forma mais factível de mensuração em larga escala dessas competências, tornando-se o mais utilizado para esse objetivo, o que explica todo o esforço em tentar entender as motivações, impactos e formas de contornar esse comportamento manipulador.

O comportamento manipulador, como o nome mesmo deixa claro, refere-se a um comportamento e não um constructo psicológico, isso é, ocorre a depender do contexto específico em que interagem fatores específicos do ambiente, da estrutura do questionário e de características pessoais do indivíduo (Heggstad, George e Reeve (2006); Heggstad (2011)).

O modelo de McFarland (McFarland e Ryan (2000)), por exemplo, aponta alguns fatores que levam os indivíduos a esse comportamento. Em primeiro lugar, crenças ou valores pessoais relacionadas à manipulação ou à mentira de maneira geral. De acordo com os autores, há um conjunto de variáveis, entre elas crenças, moral, religião e os próprios traços de personalidade, que ponderam o julgamento pessoal sobre esse comportamento. Assim, indivíduos que acreditam que a mentira é errada em qualquer circunstância são menos prováveis a manipular esses questionários.

Além disso, é necessário que o indivíduo veja no comportamento manipulador

uma oportunidade em maximizar suas chances de atingir algum resultado desejado, ou seja, deve haver alguma motivação para que ele procure inflar suas respostas. Entretanto, algumas características específicas da situação podem reduzir esse comportamento ao criar barreiras para tal, como mecanismos incorporados nos questionários que alertam a possibilidade do comportamento manipulador ser detectado. A preocupação em ter suas respostas identificadas como falsas torna a situação menos oportuna para o indivíduo agir de acordo com esse comportamento (Doll (1971); Schrader e Osburn (1977)).

Ademais, é fator relevante a capacidade que o indivíduo tem em manipular suas respostas de maneira eficaz. Agem nesse sentido, por exemplo, o conhecimento da escala, a transparência dos itens e a habilidade cognitiva. Indivíduos com mais informação sobre o que o instrumento está medindo são mais capacitados a inflar suas respostas (Cunningham, Wong e Barbee (1994)). Já transparência dos itens se refere à facilidade com que a resposta mais desejada daquele item é descoberta. Itens mais transparentes nesse sentido são mais facilmente manipulados (Lautenschlager (1994)). Além disso, a habilidade cognitiva ajuda os indivíduos a perceberem quais as respostas mais desejáveis em determinados contextos (Heggstad (2011)).

Outrossim, a real competência do indivíduo representa um potencializador ou limitador desse comportamento, isso porque indivíduos com alto nível de determinada competência tem pouco espaço para inflar suas respostas, ao ponto que indivíduos com baixo nível tem muito mais liberdade para superestimar suas competências.

Por fim, autores sugerem que outras características do indivíduo afetam a prevalência do comportamento manipulador, tais como gênero e idade, sendo o comportamento manipulador mais recorrente em homens e indivíduos mais novos (Ziegler (2007)).

Entendido os fatores que afetam o comportamento manipulador, é importante analisar qual seu impacto nas medidas de competências socioemocionais. Como supracitado, esse comportamento depende da interação de vários aspectos, de forma que a prevalência desse comportamento e seu viés causado não podem ser generalizados para todas as situações.

Vale pontuar, contudo, que há pouca dúvida sobre a capacidade dos indivíduos em manipular esses questionários quando são instruídos para tal. Há vários trabalhos que investigam o efeito da manipulação em situações *high-stake* (aquelas em que os indivíduos têm motivação para inflar suas respostas, como por exemplo seleção de emprego). Para tanto, comparam as respostas dos indivíduos em um contexto em que são instruídos a responderem às questões como se estivessem sendo avaliados com as respostas desses mesmos indivíduos sem instrução para tal (Ziegler (2007); Martin, Bowen e Hunt (2002)). Em uma meta-análise, Viswesvaran e Ones (1999) mostraram que essas instruções podem ter impacto de mais de meio desvio padrão nas medidas.

Entretanto, Smith e Ellingson (2002) argumentam que os efeitos obtidos nos moldes descritos acima não são suficientes para inferir sobre o efeito desse comportamento no mundo real. Captam, no máximo, o limite superior desse impacto. Além disso, há poucos trabalhos que investigam o efeito do comportamento manipulador em situações *low-stake*.

Nesse sentido, é importante notar que o objetivo motivacional para o comportamento manipulador (utilizando a definição de Ziegler (2007) acima) não precisa ser algo óbvio, de grande magnitude ou que faça com que o indivíduo preencha o questionário deliberadamente superestimando suas respostas. A definição também contempla objetivos muito mais sutis como, por exemplo, desejo de adequação social que podem afetar a forma com que os indivíduos respondem esses questionários de maneira não intencional.

Diante disso, o presente trabalho procura preencher essa lacuna, investigando a prevalência do comportamento manipulador e seus impactos nas medidas de competências socioemocionais em situações *low-stake* através de dois canais sutis que podem estimular esse comportamento. Além disso, por entender que há um esforço por parte de diversos agentes em desenvolver e medir essas competências no ambiente escolar, uma vez que há fortes evidências do impacto delas no bem-estar e progresso social, o trabalho procura contribuir para o acúmulo de informação sobre quais dificuldades relacionadas ao comportamento manipulador podem ocorrer nesse ambiente e propor alternativas, caso necessário, para que *policymakers* tomem suas decisões com mais informações e empreguem recursos de

maneira mais responsável.

Para isso, realizou-se um experimento em escolas públicas da rede estadual de São Paulo, no qual os alunos preencheram uma série de instrumentos de autorelato de competências socioemocionais em diferentes contextos. Procurou-se investigar se a disseminação de informação acerca das competências socioemocionais e se o fato dos professores receberem as respostas individuais dos alunos podem afetar o desejo de aceitação social dos mesmos, viesando suas respostas e afetando propriedades desejáveis dos instrumentos. Ademais, como forma de controlar possíveis vieses, outros grupos de comparação foram criados para checar se uma notificação de que havia um mecanismo presente no questionário para identificar respostas inconsistentes seria suficiente para reduzir a prevalência desse comportamento.

Portanto, o trabalho irá responder quatro perguntas referentes aos fatores que influenciam o comportamento manipulador: motivação, oportunidade e capacidade. Em primeiro lugar, as abordagens investigadas tornaram os alunos mais motivados em manipular suas respostas? Em segundo lugar, os alunos são capazes de manipular os instrumentos? Se sim, qual o efeito disso nas propriedades do instrumento? Por fim, a notificação é suficiente para barrar o comportamento manipulador?

O trabalho está estruturado da seguinte forma: o capítulo 2 traz informações sobre o experimento: construção do instrumento, escolha das abordagens e estratégia de identificação. Já o capítulo 3 descreve a amostra e apresenta algumas propriedades dos instrumentos no grupo de controle. O capítulo 4 apresenta os resultados e o capítulo 5 conclui, respondendo às perguntas acima e faz algumas recomendações para a mensuração dessas competências no contexto escolar.

2 O experimento

2.1 Construção do instrumento

Existem diferentes formas de se mensurar competências socioemocionais nos indivíduos: questionários de autorelato, hétero-relato, entrevistas, dinâmicas, entre outros. Contudo, várias delas mostram-se inviáveis para mensurações de larga escala. A mais factível nesse sentido com certeza são os questionários de autorelato, onde os próprios indivíduos avaliam suas competências através de sentenças que devem ser classificadas em escalas de percepção em relação às suas personalidades.

Além disso, sendo as competências socioemocionais relacionadas à personalidade e essa por sua vez ter inúmeros traços, há uma grande variedade nas dimensões que são medidas por esses instrumentos.

Dessa forma, a dificuldade que surge é na escolha do(s) instrumento(s) a ser(em) avaliado(s) no trabalho. Mesmo considerando apenas os questionários de autorelato, são muitos os instrumentos catalogados na literatura. Para superar essa dificuldade inicial, partiu-se do trabalho de [Primi et al. \(2016a\)](#) no qual os autores realizaram rigorosos levantamento e classificação desses instrumentos com base nos critérios de poder preditivo e maleabilidade dos constructos medidos e viabilidade e robustez (boas propriedades psicométricas) dos instrumentos.

Esse levantamento dos autores resulta em oito melhores instrumentos de acordo com os critérios acima, sendo em seguida adaptados para a faixa etária desejada e traduzidos com o objetivo de se criar um instrumento único cientificamente robusto e economicamente viável para mensuração em larga escala dessas competências em escolas no Brasil. O presente trabalho utiliza essas versões adaptadas e traduzidas de sete dos oito instrumentos para criação do questionário a ser aplicado nas escolas¹.

Além dos instrumentos que medem competências socioemocionais, estão presentes

¹ Por motivos de prazo, não foi possível a utilização do instrumento SENNA desenvolvido por [Primi et al. \(2016a\)](#). Além disso, optou-se por excluir o Big Five for Children ([Barbaranelli et al. \(2003\)](#)) devido à sobreposição de constructos que haveria se este fosse utilizado juntamente com o Big Five Inventory.

nos questionários itens referentes à desejabilidade social, aspectos cognitivos e socioeconômicos, todos com justificativas próprias que serão apresentadas nas respectivas subseções. Construir um questionário único e comum a todas as crianças mostrou-se inviável pelo tempo que seria necessário para seu preenchimento. Dessa forma, os instrumentos foram divididos em dois questionários A e B, sendo que diferem apenas nos instrumentos de mensuração das competências socioemocionais como pode ser visto na tabela 1.

Além das outras medidas não relacionadas às competências socioemocionais que eram necessárias a todos os alunos e portanto estão em ambos os questionários, optou-se por também manter o instrumento Big Five Inventory em ambos. Com isso é possível analisar padrões de correlações de todas as demais medidas com as dimensões da teoria dos 5 grandes fatores nos grupos do experimento.

Tabela 1 – Estrutura dos questionários

Questionário A					Questionário B				
Instrumento	Sigla	Dimensão	Subescalas	Itens	Instrumento	Sigla	Dimensão	Subescalas	Itens
Big Five Inventory	BFI	Inventário de personalidade	Abertura a novas experiências; Conscienciosidade; Extroversão; Amabilidade; Neuroticismo.	44	Big Five Inventory	BFI	Inventário de personalidade	Abertura a novas experiências; Conscienciosidade; Extroversão; Amabilidade; Neuroticismo.	44
Self-Efficacy Questionnaire for Children	SEQ-C	Autoeficácia	Acadêmica; Social; Emocional.	21	Strengths and Difficulties Questionnaire	SDQ	Capacidades e dificuldades	Emocional; Conduta; Relacionamento; Hiperatividade; Comportamento pró-social.	25
Nowicki-Schickland Locus of Control Scale	NSLCS	Lócus de controle		21	Core Self Evaluation Scale	CSE	Autoconceito		12
Short Grit Scale	GS	Garra		8	Rosenberg Self-Esteem Scale	RSES	Autoestima		10
Desejabilidade Social	DS	Desejabilidade social		12	Desejabilidade Social	DS	Desejabilidade social		12
Cognitive Refletion Test	CRT	Reflexo cognitivo		3	Cognitive Refletion Test	CRT	Reflexo cognitivo		3
Raven's Standard Progressive Matrices	MR	Habilidade cognitiva		4	Raven's Standard Progressive Matrices	MR	Habilidade cognitiva		4
Questionário Socioeconômico				19	Questionário Socioeconômico				19
Total				132	Total				129

Dessa forma, metade dos alunos respondeu o questionário A e a outra metade o questionário B. Eles foram distribuídos de maneira equiparada em todas as salas de todos os grupos de tratamento e controle para que o número de pessoas respondendo cada instrumento não fosse muito diferente e para que houvesse o mesmo nível de heterogeneidade nas análises de todos eles.

2.1.1 Big Five Inventory

O "Big Five Inventory"(BFI) é um instrumento proposto por [John, Donahue e Kentle \(1991\)](#) para mapear os traços de personalidade de acordo com a teoria dos 5 grandes fatores. Essa teoria, apesar de não ser consensual, é a mais bem aceita atualmente. De acordo com ela, todos os traços de personalidade podem ser agrupados em cinco grandes dimensões que estão descritas na Tabela 2, juntamente com suas descrições, características e evidências científicas que as relacionam com aspectos futuros da vida dos indivíduos.

Tabela 2 – Dimensões da teoria do Big Five

Dimensão	Descrição dicionário APA ¹	Características do Indivíduo	Evidências com indicadores futuros
Abertura a novas experiências	Tendência a ser aberto a novas experiências estéticas, culturais e intelectuais.	Imaginativo, artístico, excitável, curioso, não convencional e com amplos interesses.	Indicadores escolares.
Conscienciosidade	Tendência a ser organizado, esforçado e responsável.	Eficiente, organizado, autônomo, disciplinado, não impulsivo e orientado para seus objetivos.	Desempenho escolar, escolaridade, empregabilidade, rendimento, indicadores de saúde e menor criminalidade.
Extroversão	Orientação de interesses e energia em direção ao mundo externo e pessoas e coisas.	Amigável, sociável, autoconfiante, energético, aventureiro e entusiasmado.	Desempenho no mercado de trabalho e maior facilidade de alcançar postos de liderança.
Amabilidade	Tendência a agir de modo cooperativo e não egoisticamente.	Tolerante, altruísta, modesto, simpático, não teimoso e objetivo.	Indicadores de saúde e menor criminalidade.
Neuroticismo	Previsibilidade e consistência de reações emocionais, sem mudanças bruscas de humor.	Preocupado, irritadiço, introspectivo, impulsivo, não autoconfiante, podendo manifestar depressão e desordens de ansiedade.	Menor escolaridade, menor rendimento e piores indicadores de saúde.

¹: Associação Brasileira de Psicologia.

Nota: Elaborado com base no trabalho de [Santos \(2014\)](#).

O instrumento é composto por 44 sentenças relacionadas a comportamentos diversos que devem ser avaliadas de acordo com a frequência que ocorrem na vida do indivíduo, variando de "quase nunca"(1) a "quase sempre"(5).

As dimensões "extroversão" e "neuroticismo" possuem 8 itens cada. Já "amabilidade" e "conscienciosidade" possuem 9 itens cada e por fim, "abertura a novas experiências", 10 itens. As pontuações são obtidas somando-se os itens de cada dimensão.

Em todas as escalas, quanto maior a pontuação, maior o nível dessa competência. Vale notar que alto nível de neuroticismo está relacionado a pessoas mais emocionalmente instáveis, de forma que são encontradas evidências de relações negativas com indicadores desejáveis, ao contrário das demais dimensões. Dessa forma, as dimensões estão positivamente relacionadas com desejabilidade social, com exceção da dimensão "neuroticismo".

2.1.2 Strengths and Difficulties Questionnaire

Esse instrumento elaborado por [Goodman \(1997\)](#) se propõe a mapear comportamentos, emoções e relacionamentos dos indivíduos de acordo com suas capacidades e dificuldades. É uma alternativa de mensuração via autorelato, além de ser mais recente e com aspectos mais relevantes do mundo atual em relação ao questionário proposto por [Rotter \(1966\)](#).

O instrumento possui 5 subescalas com 5 itens cada: "problemas de conduta", "sintomas emocionais", "hiperatividade", "relacionamento com colegas" e "comportamento pró-social". As sentenças devem ser respondidas de acordo com o nível de semelhança em relação ao próprio comportamento em uma escala que varia de "falso"(0) a "verdadeiro"(2).

As subescalas são pontuadas via soma dos itens, de forma que cada uma varia de 0 a 10, sendo maior quanto maior a dificuldade do indivíduo em relação àquele aspecto da vida e portanto negativamente correlacionadas com desejabilidade social, com exceção do comportamento pró-social, que é maior quanto mais preocupado o indivíduo é em relação às suas ações na sociedade. A escala total, por sua vez, é obtida através da soma das subescalas, excluindo a de comportamento pró-social, de forma que varia de 0 a 40.

2.1.3 Self-Efficacy Questionnaire for Children

De acordo com [Bandura \(1994\)](#) autoeficácia pode ser entendida como a percepção da capacidade que o próprio indivíduo tem de realizar comportamentos por ele desejados. Esse traço de personalidade é apontado pelo mesmo autor como sendo bastante relevante para explicar o desenvolvimento de distúrbios afetivos, em especial a depressão.

Como esses distúrbios tem início geralmente no início na juventude, [Muris \(2001\)](#) desenvolveu o Self-Efficacy Questionnaire for Children (SEQ-C) para medir esse traço de personalidade em crianças. O instrumento é composto por 21² sentenças relacionadas a fatos cotidianos em que as crianças devem classificá-las de acordo com sua percepção

² Originalmente o instrumento possui 24 itens, mas sugerem tirar 3 deles que não tiveram propriedades tão satisfatórias quanto os demais.

sobre sua capacidade de lidar com eles em uma escala de 5 pontos que varia de "de jeito nenhum"(1) a "muito bem"(5).

Além disso, o instrumento pode ser fragmentado em 3 subescalas compostas por 7 itens cada: "autoeficácia emocional", entendida como a percepção que as crianças tem em relação à capacidade de lidar com emoções negativas; "autoeficácia acadêmica", que é a percepção da capacidade em administrar seu próprio comportamento de aluno, em realizar obrigações e atingir expectativas acadêmicas e "autoeficácia social", percepção da capacidade em situações que envolvem colegas e em agir com positividade. Todas elas positivamente correlacionadas com desejabilidade social.

A pontuação dessas subescalas é obtida através da soma dos itens, de forma que cada uma varia de 7 a 40 e a escala de autoeficácia total pela soma das 3 subescalas, variando de 21 a 120. Quanto maior a autoeficácia, maior a pontuação das escalas.

2.1.4 Nowicki-Schickland Locus of Control Scale

Lócus de controle refere-se à atribuição que o indivíduo dá em relação às suas conquistas e fracassos. Se os indivíduos atribuem suas conquistas e fracassos às suas ações, então possuem controle interno. Por outro lado, se atribuem à sorte, destino, acaso, entre outros, diz-se que esses indivíduos possuem controle externo.

De acordo com [Rotter \(1966\)](#) o esforço só é determinante do comportamento do indivíduo se ele enxerga uma relação causal entre esse comportamento e às conquistas e fracassos de sua vida.

Com o objetivo de criar um instrumento que medisse lócus de controle em crianças, [Nowicki e Strickland \(1973\)](#) criaram o Nowicki-Schickland Locus of Control Scale (NSLCS) que é composto por 40 itens baseados na definição de [Rotter \(1966\)](#) sobre controle interno e externo que devem ser respondidos com "sim"(1) ou "não"(0). O trabalho utiliza a versão reduzida de 21 itens sugerida pelo autor para a faixa etária mais condizente com a do experimento.

A pontuação da escala é obtida pela soma dos itens, variando de 0 a 21. Quanto mais

alta, mais externo o controle e portanto negativamente correlacionado com desejabilidade social.

2.1.5 Rosenberg Self-Esteem Scale

De acordo com [Branden \(1969\)](#), autoestima é a avaliação emocional que o indivíduo tem sobre si mesmo. É o julgamento de valor mais importante que o indivíduo pode ter, pois afeta todas as suas reações emocionais.

O instrumento Rosenberg Self-Esteem Scale (RSES) foi elaborado por [Rosenberg \(1986\)](#) e possui 10 itens que devem ser respondidos de acordo com o nível de concordância em uma escala que varia de "discordo fortemente"(0) a "concordo fortemente"(3).

A pontuação da escala é a soma dos itens, variando de 0 a 30. Ela é maior quanto maior a autoestima do indivíduo e portanto positivamente correlacionado com desejabilidade social.

2.1.6 Core Self Evaluation Scale

O termo "Core Self Evaluation" refere-se a uma dimensão de personalidade bastante abrangente que engloba traços de autoestima, autoeficácia, neuroticismo e locus de controle. Entendendo que esses 4 traços são bastante relacionados entre si e apoiados nas evidências de que eles carregam um fator comum único ([Judge e Bono \(2001\)](#)), justificou-se a busca por um conceito latente de maior ordem. Dessa forma, essa dimensão de maior ordem representaria uma autoavaliação do próprio valor, eficiência e capacidade como indivíduo e estaria relacionada à performance profissional e satisfação profissional e pessoal ([Judge et al. \(2003\)](#)).

Apesar desse fator único comum, os autores justificam que os 4 traços não são redundantes, mas sim diferentes formas de como esse conceito de ordem maior é manifestado.

Para medir esse novo conceito, [Judge et al. \(2003\)](#) criaram a "Core Self Evaluation Scale"(CSE) composta por 12 itens que devem ser respondidos em uma escala de 5 pontos

que varia de "discordo fortemente"(1) a "concordo fortemente"(5).

A escala é obtida pela média dos itens, de forma que varia de 1 a 5. Quanto mais próximo de 5, maior o nível do indivíduo em relação a esse conceito, correlacionando-se de forma positiva com desejabilidade social.

2.1.7 Short Grit Scale

Esse instrumento é uma versão reduzida com 4 itens a menos e com melhores propriedades psicométricas do questionário do instrumento Grit Scale ([Duckworth et al. \(2007\)](#)). Esse traço chamado garra está relacionado à capacidade do indivíduo em manter esforço e interesse em objetivos de longo prazo, o que representaria um traço distinto da dimensão conscienciosidade da teoria dos 5 fatores. Assim como a dimensão conscienciosidade, se relaciona positivamente com desejabilidade social.

O instrumento é composto por apenas 8 sentenças avaliadas de acordo com a identificação que o indivíduo tem com elas em uma escala de 5 pontos. Essa medida é relacionada a conquistas educacionais, desempenho acadêmico e menor mudanças de carreira ([Duckworth e Quinn \(2009\)](#)).

A pontuação da escala é a média dos itens, logo varia de 1 a 5. Quanto maior a garra do indivíduo, maior a pontuação da escala.

2.1.8 Escala de Desejabilidade Social

Como o comportamento manipulador geralmente é viesado para o sentido da desejabilidade social, também foi incluída uma escala que se propõe a medir essa propensão a responder de acordo com o desejado socialmente e a negar associação pessoal com opiniões ou comportamentos considerados socialmente negativos. A escala escolhida foi a de Marlowe-Crowne ([Crowne e Marlowe \(1960\)](#)).

Optou-se por usar a versão reduzida com apenas 13 dos 33 itens da escala original proposta por [Ballard \(1992\)](#) e traduzida por [Scagliusi et al. \(2004\)](#).

Essa escala será utilizada para checar a existência de um componente comum relacionado ao comportamento manipulador nas medidas do grupos de tratamento. Além disso, pode servir como mecanismo de ajuste nos instrumentos que medem competências socioemocionais para eliminar o viés causado pelo comportamento manipulador.

2.1.9 Cognitive Reflection Test

Esse teste foi desenvolvido por [Frederick \(2005\)](#), é composto por 3 problemas e tem por objetivo medir a reflexão cognitiva dos indivíduos. Esse traço metacognitivo captura a tendência em refletir sobre respostas de problemas para chegar nas soluções corretas ao invés de escolher as respostas mais intuitivas porém erradas ([Welsh, Burns e Delfabbro \(2013\)](#)).

Dessa forma, o instrumento está relacionado ao processo de tomada de decisão e pode ser relevante para explicar o comportamento manipulador dos indivíduos. Por essa razão, essa pequena escala foi incluída no instrumento para controlar esse traço entre os alunos.

Além da tradução para o português, para a escala melhor se adaptar às condições do experimento, foi feita uma alteração na forma de resposta dos problemas: ao invés de respostas livres, criou-se 5 alternativas de respostas incluindo a correta e a mais intuitiva mas errada para cada um dos 3 problemas.

2.1.10 Raven's Standard Progressive Matrices

As matrizes progressivas de Raven ([Raven, Raven e Court \(1998\)](#)) estão entre os testes psicométricos mais utilizados para medir o quociente de inteligência (QI) dos indivíduos. Cada matriz contém algumas figuras, um espaço em branco e alternativas de figuras em que uma delas completa a sequência lógica daquela matriz. Além disso, o nível de dificuldade aumenta conforme se avança no teste.

No instrumento, as matrizes foram introduzidas com o objetivo de medir habilidade cognitiva dos alunos uma vez que essa habilidade pode influenciar o comportamento

manipulador ([Ellingson \(2011\)](#)).

Novamente, para se adequar às condições do experimento, utilizou-se uma versão reduzida do teste composta por 4 matrizes de diferentes níveis de dificuldade.

2.1.11 Questionário socioeconômico

Por fim, também foram incluídas 19 perguntas referentes ao aluno e ambiente familiar para melhor compreensão da amostra, assim como análise de balanceamento entre grupos uma vez que há evidências que o comportamento manipulador é mais prevalente em alguns tipos de indivíduos. Há evidências por exemplo que homens e indivíduos mais novos são mais propícios a reportarem respostas extremas no sentido da desejabilidade social nessas medidas ([Ziegler \(2007\)](#)).

Assim, além dessas questões relacionadas à idade e gênero, incluiu-se outras que poderiam afetar a predisposição ao comportamento manipulador, como o fato de possuir irmãos, morar com o pai ou com a mãe, *proxies* de renda, entre outros.

2.2 Escolha das abordagens

O ambiente escolar é complexo por inúmeras razões. Os alunos estão em constante e forte processo de desenvolvimento pessoal e intelectual e se relacionam com inúmeros agentes: colegas, professores, diretores, família, entre outros. Essa complexidade, que pode ser traduzida em uma rede de incentivos bastante diversa, pode trazer dificuldades na mensuração dessas competências nesse ambiente uma vez que parte desses incentivos pode afetar a forma com que os alunos respondem a esses questionários de autorelato.

Dessa forma, muitas abordagens explorando diversos incentivos poderiam ser escolhidas para analisar a possível manipulação nesses questionários por parte dos alunos. No presente trabalho foram escolhidos apenas duas delas explicadas a seguir.

2.2.1 Disseminação da informação

São muitas as evidências da importância das competências socioemocionais para o sucesso e bem-estar dos indivíduos. Consequentemente fala-se mais sobre elas no meio acadêmico, nas esferas governamentais e também na mídia geral, principalmente quando são descobertas novas evidências acerca dessas competências ou quando novas iniciativas focadas em promovê-las são anunciadas. Nesse último caso especificamente, as competências socioemocionais tornam-se pauta inclusive dentro da sala de aula.

Não é de se espantar que hoje em dia esse assunto seja muito mais popular do que anteriormente³ e é esperado que se torne ainda mais ao longo do tempo acompanhando a tendência de agentes diversos em promover essas competências e de trabalhos na área sendo publicados. Ademais, a informação geralmente é focada nos benefícios de se desenvolver essas competências.

Com isso em vista, e sendo o controle da informação algo inviável e até indesejável, é necessário analisar se os alunos sofrem algum tipo de impacto que reflita na forma com que respondem aos questionários. É possível que ao receberem essas informações e principalmente pelo teor positivo com que são abordadas essas competências, ocorra algum tipo de pressão interna para que os alunos desejem ser mais socialmente aceitos, o que faria com que respondessem às questões de maneira viesada no sentido da desejabilidade social, de forma deliberada ou não.

Vale esclarecer em quais condições esse impacto seria ou não problemático. Por um lado, se a informação for disseminada de forma homogênea no território, todos indivíduos seriam afetados igualmente e portanto não haveria grandes problemas em comparar as medidas entre alunos ou escolas naquele momento do tempo específico. Por outro lado, se determinada população receber maior nível de informação do que outra, a comparação já torna-se problemática uma vez que o componente viesado estaria presente de forma

³ Utilizando a plataforma Google Trends, é possível checar a popularidade de termos buscados no Google em relação a todas as pesquisas no site ao longo do tempo. O termo "competência socioemocionais" quase não era buscado relativamente às demais pesquisas até meados de 2014, apresentando uma elevação em nível desde então. Já o termo "non-cognitive skills" apresentou trajetória de queda até meados de 2012 e ascendente desde então.

diferente nas medidas dessas competências dos indivíduos. Além disso, se o objetivo da mensuração for de acompanhar a trajetória dessas competências ao longo do tempo, novamente a comparação torna-se problemática, já que é esperado que a informação seja mais disseminada conforme novas iniciativas ou descobertas sobre essas competências forem acontecendo.

Para capturar esse efeito, os alunos do grupo de tratamento receberam uma palestra de cerca de 5 minutos sobre o que são e como são relacionadas as competências socioemocionais com aspectos futuros da vida do indivíduo, tendo como plano de fundo a teoria dos 5 grandes fatores. Esse tratamento seria, portanto, um choque no conjunto de informações que os alunos possuem sobre esse tema, simulando o que é esperado que aconteça ao longo do tempo.

2.2.2 Professores receberem as respostas individuais

A relação entre alunos e professores pode estar baseada em vários fatores que vai desde a admiração, até o conflito, e para alguns, completa indiferença. A depender de como essa relação é construída, as motivações e os incentivos dos alunos em preencher esses questionários variam muito. Tal fato, obviamente, não afeta somente as avaliações de competências socioemocionais, mas qualquer outra que se faça na sala de aula, mas nessas, a liberdade que o aluno tem em dar respostas diferentes das reais é muito maior.

Dada a relação dos alunos com seus professores, é provável imaginar que os incentivos se alteram quando são informados que serão avaliados pelos seus professores através das suas respostas individuais aos questionários em comparação com o contexto em que devem preencher os questionários apenas para um diagnóstico geral da escola ou da rede de ensino sem que suas respostas estejam disponíveis para seus professores.

Para investigar esse efeito, o grupo de tratamento foi informado, durante as instruções de preenchimento do questionário, que seus professores receberiam suas respostas individuais e só então os gabaritos seriam levados para a equipe que estaria fazendo o

mapeamento dessas competências na rede estadual de escolas⁴. A investigação desse efeito é justificável, pois como as ações relacionadas ao mapeamento dessas competências ainda são muito recentes, é importante ter indícios se determinados usos dessas informações tornam o contexto mais apropriado para o comportamento manipulador, criando vieses adicionais nas medidas.

2.2.3 Notificação

Além de investigar os dois canais adicionais de ativação do comportamento manipulador expostos acima, procurou-se com o trabalho encontrar mecanismos factíveis para desestimular esse comportamento nas medições de larga escala.

Inspirado no modelo de [McFarland e Ryan \(2000\)](#) e nas evidências de [Doll \(1971\)](#) e [Schrader e Osburn \(1977\)](#), parte dos alunos recebeu questionários cujos cabeçalhos continham uma notificação de que havia um mecanismo capaz de identificar se as respostas eram inconsistentes ou mentirosas e que caso essas respostas fossem identificadas, o questionário todo seria invalidado⁵.

Na possibilidade de o comportamento manipulador ter sido ativado pelas abordagens acima, mas esse artifício mostrar-se razoável e suficiente para desencorajar esse comportamento mitigando o viés da manipulação, então ter-se-ia uma solução bastante simples para superar esse problema.

Assim, metade dos alunos de cada grupo de tratamento recebeu o questionário com essa notificação, de forma que ao todo existem 5 grupos de comparação: controle, tratamento da palestra, tratamento da palestra com notificação, tratamento do professor receber os resultados individuais e tratamento do professor receber os resultados individuais com notificação.

⁴ Os outros grupos eram informados que estavam participando de uma pesquisa envolvendo a rede de escolas estaduais de São Paulo e a Universidade de São Paulo que tinha como objetivo mapear essas competências sem mencionar nenhuma avaliação por parte dos professores ou da escola muito menos a disponibilidade dos gabaritos para os mesmos.

⁵ O apêndice F contém os questionários utilizados no experimento.

2.3 Estratégia de identificação

Como a maior dificuldade para identificar o comportamento manipulador e estimar seu viés causado vem da impossibilidade de se observar o mesmo indivíduo em um ambiente neutro onde esse comportamento não é estimulado e um ambiente motivado, *high* ou *low stake*, que o faça externalizar esse comportamento, optou-se pelo desenho de um experimento aleatório (CRT), criando o contrafactual através da aleatorização dos alunos entre os grupos de controle e tratamento e eliminando o viés de seleção dos grupos.

Todos os alunos, independente dos grupos, receberam orientações gerais de preenchimento do questionário e foram informados que aquela atividade tinha como objetivo a medição dessas características nos alunos da rede pública estadual de São Paulo, em parceria com a Universidade de São Paulo.

Já os tratamentos foram descritos acima e variavam de acordo com o conteúdo que era apresentado para os alunos no momento das instruções pelo aplicador e com a presença ou não da notificação.

Seja R_{ij}^1 o resultado potencial do indivíduo i em uma medida da competência j na situação simulada pelo tratamento e R_{ij}^0 o resultado potencial do indivíduo i em uma medida da competência j no contexto escolar atual, tem-se:

$$R_{ij}^1 = r_{ij} + m_{ij}^e + m_{ij} + \varepsilon_{ij} \quad (2.1)$$

$$R_{ij}^0 = r_{ij} + m_{ij}^e + \varepsilon_{ij} \quad (2.2)$$

onde r_{ij} é o real nível da competência j no indivíduo i , m_{ij}^e é o viés de manipulação que o contexto escolar causa por si só de maneira generalizada, m_{ij} o viés de manipulação causada pela abordagem adicional que configura o grupo de tratamento e ε_{ij} um erro de medida tal que $E(\varepsilon_{ij}) = 0$, esses três últimos componentes também relacionados à competência j no indivíduo i .

Em primeiro lugar, nota-se que não se exclui a possibilidade de o ambiente escolar possuir uma rede de incentivos por si só que seja suficiente para estimular o comportamento manipulador (m_{ij}^e), mas o que se investiga no trabalho é o viés adicional que determinadas abordagens podem causar nessas medidas, m_{ij} . Espera-se que $m_{ij}^e > 0$ e $m_{ij} > 0$ para aquelas competências que estão positivamente correlacionadas com desejabilidade social e $m_{ij}^e < 0$ e $m_{ij} < 0$ para as negativamente correlacionadas (Ziegler, MacCann e Roberts (2011)).

Como só é possível observar R_{ij} de cada indivíduo i , isso é, o resultado observado e não potencial, e sendo T_i uma dummy para definir grupo de tratamento (1) ou controle (0), é possível estimar o efeito médio do tratamento $\widehat{m}_j$ em cada competência j através da comparação de médias entre grupos, ou seja:

$$E(R_{ij}|T_i = 1) - E(R_{ij}|T_i = 0) = \widehat{m}_j \quad (2.3)$$

uma vez que a aleatorização dos grupos é garantida.

2.3.1 Análise dos itens

Uma das formas de análise do impacto das abordagens da informação e do professor é em relação à quantidade de itens de um questionário que o tratamento causou diferenças significativas entre os grupos. Com isso, tem-se evidência de quão vulnerável tal instrumento é em relação àquelas situações.

Introduz-se com isso, os subscritos k e l nas definições acima. Ou seja, observa-se R_{ijkl} , o resultado do indivíduo i na competência j medido pelo item k do instrumento l . Considere também θ_+ o conjunto dos instrumentos apresentados na seção anterior que medem dimensões positivamente correlacionadas com desejabilidade social e θ_- o conjunto daqueles que medem dimensões negativamente correlacionadas, ou seja, $\theta_+ = [\text{BFI-O}, \text{BFI-C}, \text{BFI-E}^6, \text{BFI-A}, \text{SEQC-Ac}, \text{SEQC-Sc}, \text{SEQC-Em}, \text{GS}, \text{RSES}, \text{CSE}, \text{SDQ-PS}]$ e $\theta_- = [\text{BFI-N},$

⁶ Apesar de o sentido da correlação dessa dimensão com desejabilidade social não ser consensual na literatura e nem nos dados do trabalho como ficará claro na descrição da amostra, optou-se por incluí-lo

SDQ-Em, SDQ-Col, SDQ-Cond, SDQ-Hip, NSLCS]

Para cada item k dos instrumentos em θ_+ ou θ_- , comparou-se a média entre os grupos de controle e cada tratamento. Para $l \in \theta_+$, foi considerado efeito estatisticamente significativo no sentido esperado se $\widehat{m}_{jkl} > 0$ e no sentido contrário se $\widehat{m}_{jkl} < 0$ a um nível de 10% de significância. Já para $l \in \theta_-$, os sinais da desigualdade são trocados.

Em seguida, criou-se a estatística X_l . Se cada instrumento l possui K_l itens em cada dimensão, então

$$X_l = \sum_{k=1}^{K_l} x_{jkl}, \text{ onde } \begin{cases} x_{jkl} = 1, \text{ se } \widehat{m}_{jkl} > 0^* \text{ e } l \in \theta_+ \text{ ou } \widehat{m}_{jkl} < 0^* \text{ e } l \in \theta_- \\ x_{jkl} = -1, \text{ se } \widehat{m}_{jkl} < 0^* \text{ e } l \in \theta_+ \text{ ou } \widehat{m}_{jkl} > 0^* \text{ e } l \in \theta_- \\ x_{jkl} = 0 \text{ caso contrário} \end{cases} \quad (2.4)$$

O asterisco serve para lembrar que deve valer a um nível de 10% de significância. Dessa forma, X_l pode ser interpretado como o efeito líquido da intervenção nos itens do instrumento l . Se $X_l > 0$, então houve mais indício de manipulação no sentido esperado e se $X_l < 0$ há mais evidência de manipulação nos itens no sentido contrário do esperado.

Para determinar se esse efeito é suficiente e estatisticamente grande para afirmar que de fato o instrumento não se mostrou robusto àquela intervenção, criou-se uma distribuição da estatística, X_l^b , através de repetições com reposição com as respostas dos alunos da amostra de controle⁷ para cada instrumento l e todas as amostras.

Portanto, o instrumento l será considerado não robusto à determinado intervenção caso X_l esteja na região crítica da distribuição de X_l^b em qualquer uma das caudas delimitadas por um nível de significância de 10%.

Já para análise do impacto da notificação, foi criada outra estatística. Se a notificação de fato inibe o comportamento manipulador adicional causado por uma das duas abordagens propostas, então deve-se observar $\widehat{m}_{jkl} = 0$. Contudo, esse não é o único efeito. É possível que a notificação iniba o comportamento manipulador como um todo, de tal

nesse conjunto

⁷ Bootstrap com reposição com N=1000

forma que $m_{jkl} < 0$ para compensar um possível $m_{jkl}^e > 0$, ou seja, que o viés causado pela intervenção composta (com notificação) seja negativo para compensar o viés causado no contexto escolar, positivo, que não é possível isolar com o desenho do experimento (isso para aquelas dimensões correlacionadas positivamente com desejabilidade social. Para as correlacionadas negativamente, os sinais se invertem).

A estatística Z_l procura medir esse efeito composto da notificação. Seja x_{jkl} definido de acordo com 2.4 e o sobrescrito n um indicador que aquele componente foi extraído da comparação do grupo que além de uma das duas abordagens também recebeu a notificação com o grupo de controle, então:

$$Z_l = \sum_{k=1}^{K_l} (x_{jkl} - x_{jkl}^n), \text{ onde } \begin{cases} x_{jkl}^n = 1, & \text{se } \widehat{m}_{jkl}^n > 0^* \text{ e } l \in \theta_+ \text{ ou } \widehat{m}_{jkl}^n < 0^* \text{ e } l \in \theta_- \\ x_{jkl}^n = -1, & \text{se } \widehat{m}_{jkl}^n < 0^* \text{ e } l \in \theta_+ \text{ ou } \widehat{m}_{jkl}^n > 0^* \text{ e } l \in \theta_- \\ x_{jkl}^n = 0 & \text{caso contrário} \end{cases} \quad (2.5)$$

Dessa forma, se $Z_l = 0$, então a notificação não teve nenhum impacto: a mesma quantidade de itens com evidências de manipulação foi encontrada no grupo que recebeu uma das duas abordagens e no grupo que recebeu uma das abordagens e a notificação. Se, contudo, $Z_l > 0$, então a notificação reduziu o número de itens manipulados no sentido esperado pelas abordagens e/ou apresentou diferenças significantes entre os grupos no sentido inverso da desejabilidade social, o que seria evidência de que inibiu também o viés causado pelo contexto escolar. Por outro lado, se $Z_l < 0$ então a notificação teve efeito contrário do esperado, ou seja, aumentou a manipulação no sentido da desejabilidade social.

Da mesma forma, para inferir se a notificação foi estatisticamente eficiente para inibir o comportamento manipulador, a estatística Z_l foi comparada com a distribuição de Z_l^b , construída através de repetições com reposição com a amostra do grupo de controle.

2.3.2 Análise dos constructos

Apesar da análise individual dos itens dar um indício sobre a intensificação do comportamento manipulador com as abordagens nos instrumentos, ainda é necessário checar se na agregação dos itens para construção dos constructos daquele instrumento esse viés permanece ou é de certa forma recompensado ou minimizado pelos demais. Para isso, comparou-se os constructos agregados das formas como seus respectivos autores instruem entre os grupos dos tratamentos de controle.

Foi considerado indício de intensificação do comportamento manipulador causado pelas abordagens quando a diferença entre as médias dos constructos ($\widehat{m}_{jl}$) mostrou-se positiva e estatisticamente significativa a um nível de 10% para $l \in \theta_+$ e negativa nas mesmas condições para $l \in \theta_-$.

Além disso, calculou-se o tamanho desse viés quantitativamente em termos de desvios-padrão⁸ e qualitativamente em número e porcentagem de alunos do grupo de controle que precisariam alterar sua resposta no sentido da desejabilidade social na escala daquele constructo para atingir a média do mesmo constructo no grupo de tratamento.

2.3.3 Análise da consistência interna

Outra forma de se avaliar o impacto das intervenções é através da consistência interna dos instrumentos, que é uma das formas de se checar a confiabilidade do mesmo. De acordo com Hayes (1998), a confiabilidade é o grau com que o instrumento reflete o resultado verdadeiro. A estatística utilizada é o coeficiente α de Cronbach (Cronbach (1951)), que de acordo com Guttman (1953) mede o limite inferior da consistência interna dos instrumentos:

$$\alpha_l = \left[\frac{K_l}{K_l - 1} \right] * \left[1 - \frac{\sum_{k_l=1}^{K_l} S_k^2}{S_{tl}^2} \right] \quad (2.6)$$

onde K_l é o número de itens do instrumento l , S_k^2 a variância de cada item e S_{tl}^2 a variância

⁸ Delta Glass em módulo

total do instrumento l .

De acordo com Cortina (1993), o coeficiente deve ser interpretado com cautela pois é função, entre outras coisas, do número de itens do questionário. Contudo, isso não é um problema pois a comparação que será feita é entre amostras diferentes para o mesmo questionário.

Esse indicador é negativamente correlacionado com a prevalência de comportamento manipulador na amostra (Pastore e Lombardi (2014)). Assim, espera-se que o coeficiente alfa seja menor nos grupos em que os alunos receberam a palestra ou foram informados que os professores receberiam suas respostas individuais. Para checar se isso de fato ocorre, estimou-se também os limites do intervalo de confiança ao nível de 90%. Com isso é possível saber, com 90% de confiança se os coeficientes α_l nos grupos de tratamento são menores do que os do grupo de controle.

2.3.4 Análise das correlações

Outra propriedade que deve ser afetada com o comportamento manipulador é a correlação dos itens ou constructos que medem diferentes competências. Isso porque em um contexto onde o comportamento manipulador está presente, a medida observada R_{ij} contém não somente o componente relacionado ao que a medida se propõe a medir, r_{ij} , mas também os vieses causados pela manipulação, m_{ij} e m_{ij}^e .

Como nesse contexto onde há manipulação R_{ij} também captura o comportamento manipulador e ele está presente em outras medidas, é natural que as correlações entre as medidas observadas nessa situação sejam maiores do que no contexto livre de manipulação (Ziegler, MacCann e Roberts (2011)). Para checar esse efeito, *heat maps* foram construídos com as correlações dos constructos⁹ em cada grupo. Se esse efeito de fato existe e o comportamento manipulador foi intensificado nos grupos que receberam as abordagens do grupo de tratamento, então espera-se que as cores dos mapas sejam mais intensas.

Além disso, são feitas análises de correlações entre os constructos relacionados às

⁹ Também foram feitos *heat maps* com os itens dos questionários, apresentados no apêndice E.

competências socioemocionais e a escala de desejabilidade social. Se essa escala possui alto nível de confiabilidade, então as correlações entre ela e as medidas das competências devem ser maiores quando o comportamento manipulador estiver presente.

Por fim, espera-se que em geral, os instrumentos possuam melhores propriedades no grupo onde não há manipulação. Tendo isso em vista, em todas as comparações entre grupos é feito um rápido levantamento da quantidade de correlações entre instrumentos que possuem sinais contrários ao esperado pelas evidências na literatura. Quanto maior esse número maior o indício de que aquela abordagem piorou essa propriedade dos instrumentos de maneira geral.

3 Amostra

O experimento ocorreu em 10 escolas públicas da rede estadual de São Paulo que fazem parte de um conjunto de escolas parceiras da ONG Parceiros da Educação, de forma que não são representativas da rede estadual de São Paulo como um todo. Como pode ser visto na Figura 1, o desempenho dos alunos dessas escolas em provas padronizadas é maior do que na rede estadual e no Brasil nas duas disciplinas medidas pelo SARESP (Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo) e pela ANEB/Prova Brasil.

Figura 1 – Comparativo de desempenho em provas padronizadas

Nota: SARESP edição 2015 e ANEB/Prova Brasil edição 2013

Por outro lado, é possível notar na tabela 3 que apesar de em média os alunos apresentarem desempenho superior que a média da rede e do país, há muita heterogeneidade entre as escolas participantes, havendo, inclusive escolas com desempenho abaixo da rede estadual e da média nacional.

Fizeram parte do experimento 2929 alunos do sexto ao nono ano dessas 10 escolas. Foram retirados da amostra 76 casos com problemas de gabarito¹, de forma que a amostra é composta por 2853 alunos. Além disso, para fins de teste de robustez, uma amostra mais restrita foi criada somente com os gabaritos que foram inteiramente respondidos, totalizando apenas 781 alunos, equivalente a 27,4% da amostra completa. Esse número

¹ Especificamente, casos em que o aplicador entregou ao mesmo aluno folha de gabarito e questionário não condizentes entre si.

Tabela 3 – Desempenho em provas padronizadas - Escolas da amostra

Escola	SARESP 2015		ANEB/Prova Brasil 2013	
	Língua Portuguesa	Matemática	Língua Portuguesa	Matemática
Escola 1	231,6	254,6	245,5	243,2
Escola 2	262,3	283,0	254,1	252,5
Escola 3	253,7	264,6	262,5	271,1
Escola 4	242,9	263,0	248,1	251,3
Escola 5	242,9	267,1	240,2	251,7
Escola 6	254,0	273,4	250,4	267,8
Escola 7	262,0	276,6	223,2	232,4
Escola 8	232,8	249,8	236,5	243,7
Escola 9	230,6	249,6	238,4	246,4
Escola 10	255,6	273,2	243,4	246,1
Média	246,8	265,5	244,2	250,6
Variância	136,8	117,9	103,9	119,2

bastante reduzido é reflexo por um lado da dificuldade dos alunos em compreender alguns termos do questionário, que apesar de adaptados para a faixa etária ainda cria alguns obstáculos ao pleno entendimento² e por outro, pela dificuldade enfrentada pela equipe de aplicadores nas escolas em despertar atenção e interesse dos alunos em preencher os questionários de forma rigorosa. Pontua-se que essa amostra pode ser considerada mais limpa de ruídos, uma vez que faz sentido acreditar que nessa amostra estão selecionados os alunos menos displicentes. Não obstante, é bastante pequena, principalmente considerando as fragmentações feitas para capturar os diferentes efeitos pretendidos pelo trabalho, o que afeta consideravelmente o efeito mínimo detectável³.

A tabela 4 apresenta a distribuição de alunos por grupo e questionário nas diferentes amostras. Como pode ser visto, houve uma distribuição bastante equilibrada entre grupos, resultado do processo de aleatorização do experimento. O total de alunos varia de 18,4% (tratamento professor) a 22,6% (controle) entre os grupos. Além disso, mesmo na amostra restrita há uma distribuição bastante equitativa entre os grupos, variando de 18,3% (tratamento professor) a 22,0% (tratamento palestra + notificação).

² Em todas as salas os aplicadores colocavam uma espécie de glossário na lousa para tentar superar esse problema.

³ O maior MDE (teste monocaudal com poder igual a 90% e significância 10%) entre os instrumentos e tipo de tratamento na amostra principal é de 0,25 dp enquanto na amostra reduzida é de 0,50 dp.

Tabela 4 – Distribuição de alunos por grupo, amostra e questionário

Grupo	Amostra completa			Amostra restrita		
	A	B	Total	A	B	Total
Controle	313	332	645	67	92	159
Palestra	265	285	550	71	74	145
Palestra + Notificação	282	284	566	82	90	172
Professor	255	271	526	76	67	143
Professor + Notificação	284	282	566	81	81	162
Total	1399	1454	2853	377	404	781

Ademais, é possível notar uma distribuição bastante equilibrada de alunos entre os diferentes tipos de questionários: 49,0% preencheram o questionário A e 51,0% o questionário B, resultado do processo alternado de entrega dos questionários nas salas de aula pelos aplicadores. Na amostra restrita houve uma maior diferença, mas ainda assim bastante pequena, de 48,3% e 51,7% de questionários inteiramente preenchidos nos questionários A e B respectivamente⁴.

Para realizar as análises discutidas na seção anterior e aproveitando ao máximo as respostas dos alunos, chamou-se de gabarito AB as respostas aos itens do BFI e aquelas não referentes às competências socioemocionais que eram comuns aos questionários A e B (checar tabela 1). Já o que está sendo chamado de gabarito A' e B' são as respostas aos questionários A e B excluindo aquelas do BFI.

Ao utilizar a amostra completa para fazer as análises, é importante ter como garantia que a ausência de respostas é aleatória entre os grupos, uma vez que esse padrão de resposta pode estar correlacionado com características não observadas que poderiam afetar a predisposição ao comportamento manipulador. A tabela 5 apresenta os resultados dessa análise dos missings entre os grupos e gabaritos.

Tabela 5 – Diferença de média da porcentagem de missing entre grupos e gabaritos

% Resposta	Média Controle	Palestra		Palestra + Notificação		Professor		Professor + Notificação	
		Média	P-Valor	Média	P-Valor	Média	P-Valor	Média	P-Valor
A'	0,93	0,94	0,36	0,95	0,17	0,96	0,02	0,92	0,56
B'	0,95	0,95	0,40	0,95	0,53	0,94	0,41	0,93	0,25
AB	0,95	0,95	0,13	0,95	0,18	0,95	0,41	0,94	0,36

⁴ As distribuições dos alunos entre grupos e tipos de questionário por escola estão no apêndice B.

Como pode ser visto, a porcentagem de resposta entre os gabaritos no grupo de controle varia de 93% a 95%. Dados os totais de questões em cada um deles ($A'=89$, $B'=86$ e $AB=83$), em média os alunos de cada tipo de gabarito deixaram de responder entre 4,4 e 6,1 questões. Além disso, das 12 comparações possíveis (4 tratamentos e 3 gabaritos), somente em uma delas a diferença de média entre as porcentagens de respostas foi estatisticamente significativa (tratamento professor no gabarito A'), equivalente a 3 p.p. ou apenas 2,4 questões. Com isso, optou-se por considerar a amostra completa para a análise principal e a amostra restrita de questionários inteiramente respondidos para testar a robustez como supracitado. Entende-se que a análise de robustez é suficiente para superar o problema pontual dessa única comparação onde houve diferença estatisticamente significativa.

A tabela 6 apresenta as estatísticas descritivas para as duas amostras. A amostra completa é composta por uma maioria de meninas (53%), de raça não branca ou amarela (58%) e com idade média de 12,3 anos. Além disso, os alunos frequentam em média o primeiro semestre do sétimo ano (média 7,4). Em relação ao ambiente familiar, mais de 90% deles moram com as mães e nesses casos, apenas 2% das mães não sabem ler ou escrever. Por outro lado, apenas 71% das crianças moram com os pais e em 5% desses casos o pai não sabe ler ou escrever. Além disso, praticamente metade da amostra possui irmãos mais novos e irmãos mais velhos. O número de moradores médio nas casas desses alunos é de aproximadamente 4 pessoas e essas casas possuem pouco mais de 2 quartos, menos de 2 banheiros em média e só possuem empregada em 12% dos casos. Ademais, em média, esses lares possuem menos de 1 carro e mais de 1 computador por família. Em 81% dos casos suas casas estão em ruas asfaltadas e em 4% não há energia elétrica.

Por fim, em relação às medidas de traços cognitivos, pode-se notar que os alunos tiveram muitas dificuldades, principalmente em relação ao CRT, com apenas 7% de respostas corretas em média. Em relação às matrizes de Raven, o percentual de acerto foi de 42%.

Em relação à amostra reduzida, há muita semelhança. As únicas diferenças rele-

vantes são que os alunos em sua totalidade moram com as mães e os pais e as variáveis que podem estar mais correlacionadas com condição econômica da família são levemente mais favoráveis. Já em relação às medidas de cognição, houve menor percentual de acerto no CRT e maior no MR.

Tabela 6 – Estatísticas descritivas das amostras

Variável	Amostra Completa					Amostra Restrita				
	Média	Variância	Min	Max	N	Média	Variância	Min	Max	N
Gênero (homem=1)	0,47	0,25	0	1	2657	0,43	0,25	0	1	781
Raça (branco e amarelo =1)	0,42	0,24	0	1	2641	0,46	0,25	0	1	781
Idade	12,30	1,13	11	15	2702	12,39	1,01	11	15	781
Ano escolar	7,40	1,33	6	9	2561	7,59	1,29	6	9	781
CRT (%)	0,07	0,02	0	1	2633	0,05	0,02	0	0,67	781
MR (%)	0,42	0,07	0	1	2622	0,47	0,07	0	1	781
Mora mãe	0,91	0,09	0	1	2603	1,00	0,00	1	1	781
Mora pai	0,71	0,21	0	1	2557	1,00	0,00	1	1	781
Mãe ler e escrever	0,98	0,02	0	1	2270	0,99	0,01	0	1	781
Pai ler e escrever	0,95	0,05	0	1	1734	0,98	0,02	0	1	781
Irmãos	0,82	0,15	0	1	2604	0,87	0,12	0	1	781
Irmãos novos	0,49	0,25	0	1	2608	0,51	0,25	0	1	781
Irmãos velhos	0,52	0,25	0	1	2624	0,55	0,25	0	1	781
Pessoas na casa	3,99	1,34	2	6	2611	4,14	1,03	2	6	781
Empregada	0,12	0,11	0	1	2584	0,11	0,10	0	1	781
Carros	0,99	0,76	0	3	2608	1,05	0,62	0	3	781
Banheiros	1,43	0,60	0	3	2607	1,49	0,57	0	3	781
Quartos	2,04	0,61	0	3	2605	2,16	0,48	0	3	781
Computador	1,38	0,85	0	3	2604	1,45	0,80	0	3	781
Asfalto	0,81	0,15	0	1	2331	0,88	0,11	0	1	781
Energia	0,96	0,04	0	1	2527	0,98	0,02	0	1	781

Apesar de o processo de aleatorização ter sido garantido como explicado anteriormente, as tabelas de 21 a 28, dispostas no apêndice C, apresentam os testes de balanceamento para todas as comparações possíveis. Em geral, não parece haver problema de desbalanceamento entre os grupos⁵, principalmente em relação às variáveis mais relevantes para o comportamento manipulador de acordo com as evidências apontadas na introdução. De toda forma, como mais um teste de robustez dos resultados, foi investigado o impacto dos tratamentos no contexto de regressão considerando todas as variáveis não balanceadas em seus respectivos grupos como covariadas. Esses resultados estão no apêndice D.

Já em relação às competências socioemocionais, a tabela 7 apresenta as médias e desvios-padrão dos constructos por instrumento nas duas amostras do grupo de controle

⁵ Cada amostra possui 12 grupos, de forma que ao todo são 24 grupos

(amostra 1 refere-se à completa e amostra 2 à restrita), comparando-as com as mesmas estatísticas encontradas na literatura para os mesmos instrumentos. Sempre que possível, o trabalho escolhido como referência foi aquele que originou o instrumento. Quando essas estatísticas não estavam reportadas nesses trabalhos, referenciou-se outro cuja faixa etária da amostra mais se aproximava da amostra desse experimento.

Em geral, não há muita diferença entre as competências observadas no grupo de controle com aquelas reportadas nos trabalhos de referência. As diferenças mais significativas são em relação aos instrumentos SDQ e NSLCS. No primeiro, os alunos reportaram níveis de dificuldades muito maiores, com diferenças superiores a 1 d.p. Em relação ao locus de controle, os alunos da amostra reportaram controle muito mais interno.

Outra forma de avaliar a amostra é em relação às propriedades dos instrumentos nos grupos de controle. A tabela 8 apresenta os coeficientes α de Cronbach calculados nas amostras completa e restrita. Pode-se notar que em termos de consistência interna praticamente em todos os instrumentos essa propriedade é pior em comparação com os respectivos trabalhos de referência. Ademais, utilizando a recomendação de α no mínimo igual a 0,7 para ser considerado razoável (Nunnally, Bernstein e Berge (1967)), apenas entre 6 e 9 escalas passariam no teste.

Esses valores baixos de α podem em parte ser considerado um problema da amostra relacionado à displicência com que os questionários foram preenchidos, uma vez que os coeficientes em geral são superiores na amostra restrita, onde há uma seleção nesse sentido. Por outro lado, de acordo com McFarland e Sparks (1985), a consistência interna é positivamente correlacionada com idade e escolaridade, o que explicaria parte do problema uma vez que os alunos da amostra em geral são mais novos e menos escolarizados que aqueles nos trabalhos de referência.

Por fim, as figuras 2 e 3 apresentam as matrizes de correlação dos constructos nas amostras completa e restrita respectivamente para o grupo de controle⁶. É possível notar

⁶ As diagonais principais das matrizes não são inteiramente preenchidas pois os instrumentos foram divididos em dois questionários de forma que não é possível cruzar a informação do mesmo aluno em todos os instrumentos.

Tabela 7 – Médias e desvios-padrão dos constructos na amostras de controle: comparação com trabalhos de referência

Instrumento	Referência	N	Faixa etária	Local	Média (DP)	Grupo Controle			
						Amostra 1		Amostra 2	
						Média (DP)	N	Média (DP)	N
BFI - O ¹					3,7 (0,6)	3,3 (0,6)	645	3,4 (0,8)	159
BFI - C	Benet-Martinez e John (1998)	711	Alunos de graduação	EUA	3,6 (0,7)	3,4 (0,7)	645	3,6 (0,7)	159
BFI - E					3,2 (0,8)	3,3 (0,7)	645	3,3 (0,7)	159
BFI - A					3,8 (0,6)	3,7 (0,6)	645	3,9 (0,6)	159
BFI - N					3,0 (0,8)	2,7 (0,6)	645	2,7 (0,8)	159
GS					3,4 (0,7)	3,4 (0,6)	301	3,5 (0,5)	67
SDQ - Em					2,0 (2,1)	3,8 (2,2)	329	3,7 (2,2)	92
SDQ - Cond	Muris, Meesters e Berg (2003)	562	9 a 15 anos	Holanda	1,3 (1,3)	3,2 (2,1)	329	2,8 (1,9)	92
SDQ - Hip					2,9 (2,2)	4,7 (2,3)	329	4,4 (2,3)	92
SDQ - Col					1,3 (1,6)	2,9 (2,1)	329	2,3 (1,8)	92
SDQ - PS					8,2 (1,6)	7,3 (2,1)	329	7,7 (1,9)	92
NSLCS					11,4 - 18,8 (3,1 - 5,3)	9,9 (3,2)	308	8,8 (3,1)	67
SEQC - Ac	Muris (2001)	330	14 a 17 anos	EUA	23,6 (5,8)	25,0 (4,9)	309	26,1 (3,9)	67
SEQC - Em					25,0 (5,0)	22,6 (5,2)	308	22,9 (4,7)	67
SEQC - Sc					28,2 (4,3)	24,3 (4,7)	307	24,7 (4,3)	67
CSE	Judge et al. (2003)	126 a 205	Alunos de graduação	EUA	3,8 - 4,0 (0,5 - 0,6)	3,4 (0,5)	328	3,5 (0,5)	92
RSES	Sinclair et al. (2010)	503	Variado	EUA	22,6 (5,8)	18,7 (4,6)	323	19,7 (4,3)	92

¹: os valores dos constructos do BFI estão reportados como médias dos itens para serem compatíveis com o do trabalho de referência. Contudo, ao longo do trabalho esses constructos foram calculados através da soma dos itens conforme orientação no trabalho de John, Donahue e Kentle (1991).

que em geral os instrumentos se comportam de acordo com o esperado em relação ao sinal das correlações. Na amostra completa, a única correlação com sinal contrário ao esperado foi entre as dimensões SDQ-Em e SDQ-PS. Por outro lado, contraintuitivamente parece haver maior perturbação nas correlações na amostra restrita: são 5 correlações com sinal contrário do esperado, em geral nos instrumentos SDQ e NSLCS, exatamente aqueles que apresentaram menor consistência interna na amostra.

Tabela 8 – Consistência interna dos instrumentos nas amostras de controle: comparação com trabalhos de referência

Instrumento	Referência	N	Faixa etária	Local	α	Grupo Controle			
						Amostra 1		Amostra 2	
						α	N	α	N
BFI - O	John, Donahue e Kentle (1991)	462	Alunos de graduação	EUA	0,81	0,66	645	0,75	159
BFI - C					0,82	0,73	645	0,77	159
BFI - E					0,88	0,59	645	0,61	159
BFI - A					0,79	0,63	645	0,63	159
BFI - N					0,84	0,60	645	0,71	159
GS	Duckworth e Quinn (2009)	179 e 1308	Variado	EUA	0,73 - 0,83	0,57	301	0,65	67
SDQ - Em	Goodman (1997)	3983	5 a 15 anos	Inglaterra	0,66	0,53	329	0,48	92
SDQ - Cond					0,6	0,51	329	0,50	92
SDQ - Hip					0,67	0,56	329	0,59	92
SDQ - Col					0,41	0,45	329	0,36	92
SDQ - PS					0,66	0,63	329	0,65	92
NSLCS ¹	Nowicki e Strickland (1973)	1017	Alunos de "elementary" e "high school"	EUA	0,63 - 0,81	0,56	308	0,58	67
SEQC - Ac	Muris (2001)	330	14 a 17 anos	EUA	0,88	0,73	309	0,73	67
SEQC - Em					0,86	0,71	308	0,73	67
SEQC - Sc					0,85	0,65	307	0,67	67
CSE	Judge et al. (2003)	126 a 205	Alunos de graduação	EUA	0,83 - 0,87	0,68	328	0,70	92
RSES	Sinclair et al. (2010)	503	Variado	EUA	0,91	0,72	323	0,72	92

¹: Coeficiente de Spearman Brown para cálculo da consistência interna uma vez que é uma escala bivariada.

Figura 2 – Matriz de correlação dos constructos - Grupo de controle - Amostra completa

		BFI					DS	SEQC			NSLCS	GS	SDQ					CSE	RSES
		E	A	C	N	O		Ac	Sc	Em			Em	Cond	Hip	Col	PS		
BFI	E	0.15																	
	A		0.43																
	C	0.00		0.43															
	N	0.01	-0.36	-0.22															
	O	0.20	0.28	0.33	-0.10														
DS		-0.03	0.34	0.46	-0.32	0.21													
SEQC	Ac	0.02	0.38	0.59	-0.18	0.27	0.30												
	Sc	0.37	0.35	0.25	-0.17	0.33	0.15	0.44											
	Em	0.20	0.39	0.31	-0.43	0.30	0.34	0.44	0.48										
NSLCS		-0.09	-0.24	-0.23	0.26	-0.13	-0.30	-0.24	-0.23	-0.15									
GS		0.01	0.39	0.57	-0.29	0.22	0.48	0.50	0.25	0.34	-0.27								
SDQ	Em	-0.10	-0.11	-0.11	0.37	0.00	-0.20												
	Cond	0.17	-0.34	-0.27	0.34	-0.12	-0.38						0.29						
	Hip	0.27	-0.26	-0.52	0.35	-0.13	-0.35						0.22	0.34					
	Col	-0.19	-0.17	-0.11	0.11	-0.05	-0.13						0.35	0.28	0.06				
	PS	0.09	0.48	0.30	-0.11	0.26	0.23						0.14	-0.27	-0.14	-0.16			
CSE		0.20	0.28	0.39	-0.25	0.21	0.35						0.41	-0.27	-0.31	-0.39	0.20		
RSES		0.13	0.25	0.35	-0.24	0.14	0.28						0.34	-0.28	-0.20	-0.35	0.14	0.68	

Nota: Em destaque correlações com sinais contrários ao esperado pela literatura. BFI-E está marcado pois não há muito consenso sobre o sinal de suas correlações com os demais constructos.

Figura 3 – Matriz de correlação dos constructos - Grupo de controle - Amostra reduzida

		E	BFI					DS	SEQC			NSLCS	GS	SDQ					CSE	RSES
			A	C	N	O			Ac	Sc	Em			Em	Cond	Hip	Col	PS		
BFI	E																			
	A	0.14																		
	C	-0.02	0.41																	
	N	0.03	-0.34	-0.11																
	O	0.18	0.29	0.33	-0.10															
DS		-0.01	0.33	0.34	-0.25	0.18														
SEQC	Ac	0.07	0.48	0.62	-0.03	0.26	0.14													
	Sc	0.34	0.45	0.36	-0.12	0.33	0.09	0.48												
	Em	0.22	0.58	0.33	-0.49	0.28	0.27	0.38	0.41											
NSLCS		-0.09	0.08	-0.02	0.11	-0.08	-0.20	-0.06	-0.17	0.11										
GS		0.16	0.47	0.55	-0.21	0.34	0.33	0.54	0.41	0.45	-0.12									
SDQ	Em	-0.21	-0.08	-0.02	0.27	0.03	-0.19													
	Cond	0.15	-0.43	-0.19	0.37	-0.20	-0.26							0.25						
	Hip	0.37	-0.32	-0.59	0.33	-0.23	-0.31							0.17	0.37					
	Col	-0.23	-0.06	-0.02	0.02	-0.17	-0.19							0.27	0.12	-0.01				
	PS	0.01	0.56	0.31	-0.24	0.33	0.29							0.24	-0.23	-0.14	-0.04			
CSE		0.17	0.21	0.41	-0.16	0.30	0.44							-0.41	-0.19	-0.32	-0.29	0.15		
RSES		0.19	0.22	0.34	-0.19	0.26	0.41							-0.41	-0.17	-0.26	-0.27	0.16	0.76	

Nota: Em destaque correlações com sinais contrários ao esperado pela literatura. BFI-E está marcado pois não há muito consenso sobre o sinal de suas correlações com os demais constructos.

4 Resultados

4.1 Efeito da informação

A tabela 9 apresenta o efeito da palestra nas respostas dos alunos através da análise dos itens dos instrumentos. A segunda coluna mostra a quantidade de itens cujas médias entre os grupos de controle e tratamento mostraram-se estatisticamente significantes no sentido da desejabilidade social, ou seja, melhores competências socioemocionais no grupo de alunos que assistiu a palestra, enquanto a quarta se refere aos itens com diferenças estatisticamente significativas no sentido contrário da desejabilidade social, o que seria contraintuitivo. O efeito líquido da palestra por instrumento é, como explicado anteriormente, calculado pela estatística X_l e apresentada na terceira coluna que está sendo comparada com a região crítica construída conforme explicado anteriormente. Estatísticas X_l positivas e dentro da região crítica indicam que há indícios de que o comportamento manipulador foi bem sucedido naqueles itens, ou seja, alunos reportaram sistematicamente melhores competências socioemocionais após assistirem a palestra.

Conforme pode ser visto, a estatística X_l é positiva em 12 dos 17 constructos conforme o esperado. Contudo, é estatisticamente significativa em apenas 5 deles: BFI-C, NSLCS, SDQ-Cond, SEQC-Ac e SEQC-Em, todos identificados em negrito na planilha. Não há nenhum caso de estatística X_l negativa e significativa. Nesses casos onde houve indício de manipulação, a porcentagem de itens manipulados varia de 22% no BFI-C a 43% nas dimensões do SEQC. Quando utiliza-se a amostra restrita, há menos evidência de manipulação nos itens, aparecendo somente na dimensão “abertura a novas experiências” do BFI, com 5 itens (50%) com diferenças significativas no sentido da desejabilidade social.

Em relação à análise dos constructos depois da agregação dos itens, os resultados se encontram na tabela 10. O viés causado pela manipulação permanece significativo em apenas 2 dos 5 casos mencionados anteriormente: NSLCS e SEQC-Em, cujos efeitos em termos de desvio padrão são de apenas 0,20 e 0,13 d.p. respectivamente. Já qualitativamente, em ambos os casos, representa em torno de 65% dos alunos do grupo de controle alterarem suas

Tabela 9 – Efeito da palestra sobre os itens dos constructos

	Amostra completa				Amostra restrita			
	Sentido DS	Sentido contrário	Efeito descontado		Sentido DS	Sentido contrário	Efeito descontado	
	N Itens	N Itens	X_i	RC	N Itens	N Itens	X_i	RC
BFI - E (8 itens)	0	1	-1	$(-\infty, -2] \cup [3, +\infty)$	0	0	0	$(-\infty, -2] \cup [3, +\infty)$
BFI - A (9 itens)	2	1	1	$(-\infty, -2] \cup [2, +\infty)$	1	1	0	$(-\infty, -2] \cup [3, +\infty)$
BFI - C (9 itens)	2	0	2	$(-\infty, -2] \cup [2, +\infty)$	1	0	1	$(-\infty, -2] \cup [2, +\infty)$
BFI - N (8 itens)	1	2	-1	$(-\infty, -2] \cup [3, +\infty)$	2	2	0	$(-\infty, -2] \cup [2, +\infty)$
BFI - O (10 itens)	2	0	2	$(-\infty, -3] \cup [3, +\infty)$	5	0	5	$(-\infty, -3] \cup [3, +\infty)$
NSLCS (21 itens)	6	0	6	$(-\infty, -4] \cup [3, +\infty)$	0	1	-1	$(-\infty, -3] \cup [3, +\infty)$
GS (8 itens)	2	0	2	$(-\infty, -3] \cup [3, +\infty)$	1	0	1	$(-\infty, -2] \cup [2, +\infty)$
SDQ - Em (5 itens)	0	1	-1	$(-\infty, -2] \cup [2, +\infty)$	0	1	-1	$(-\infty, -2] \cup [2, +\infty)$
SDQ - Cond (5 itens)	2	0	2	$(-\infty, -2] \cup [2, +\infty)$	0	0	0	$(-\infty, -2] \cup [2, +\infty)$
SDQ - Col (5 itens)	0	1	-1	$(-\infty, -2] \cup [2, +\infty)$	0	0	0	$(-\infty, -2] \cup [2, +\infty)$
SDQ - Hip (5 itens)	0	0	0	$(-\infty, -2] \cup [2, +\infty)$	0	1	-1	$(-\infty, -2] \cup [2, +\infty)$
SDQ - PS (5 itens)	1	0	1	$(-\infty, -2] \cup [2, +\infty)$	1	1	0	$(-\infty, -2] \cup [2, +\infty)$
RSES (10 itens)	2	1	1	$(-\infty, -3] \cup [2, +\infty)$	0	1	-1	$(-\infty, -3] \cup [3, +\infty)$
CSE (12 itens)	2	0	2	$(-\infty, -3] \cup [3, +\infty)$	1	1	0	$(-\infty, -2] \cup [3, +\infty)$
SEQC - Ac (7 itens)	3	0	3	$(-\infty, -2] \cup [2, +\infty)$	1	0	1	$(-\infty, -3] \cup [2, +\infty)$
SEQC - Sc (7 itens)	2	1	1	$(-\infty, -2] \cup [2, +\infty)$	2	3	-1	$(-\infty, -3] \cup [2, +\infty)$
SEQC - Em (7 itens)	3	0	3	$(-\infty, -3] \cup [2, +\infty)$	1	0	1	$(-\infty, -2] \cup [2, +\infty)$

respostas em uma pontuação acima da escala em um dos itens que compoem a dimensão. Nos outros 3 constructos que apresentaram indícios de manipulação dos itens, o viés foi de certa forma recompensado pelos demais itens que compoem a dimensão. Por outro lado, a dimensão BFI-O que anteriormente não aparecia como manipulada na análise dos itens na amostra completa apresentou viés significante na análise dos constructos nas duas amostras. Na amostra completa, o viés equivale a apenas 0,07 d.p. ou 47,8% dos alunos do grupo de controle alterando suas respostas em uma pontuação acima de um dos itens da escala. Na amostra restrita esses efeitos são maiores, respectivamente a 0,21 d.p. e quase 146% dos alunos. Essa porcentagem maior que 100% pode ser melhor entendida como metade dela, no caso 73% dos alunos alterando as respostas de dois itens da escala em uma pontuação acima no sentido da desejabilidade social.

A tabela 11 apresenta a consistência interna dos instrumentos nos dois grupos de tratamento referentes à palestra nas duas amostras. Na segunda coluna estão os coeficientes α do grupo de controle, seguidos pelos limites inferiores dos respectivos intervalos de confiança no nível de 90%. Comparando a terceira e a quarta colunas podemos inferir se a consistência interna dos instrumentos no grupo que recebeu a palestra é menor que a do grupo de controle com 90% de confiança. Das 17 escalas que medem competências socioemocionais, em apenas 3 delas pode-se dizer que houve uma redução da consistência

Tabela 10 – Efeito da palestra nos constructos

	Amostra completa						Amostra restrita							
	Média controle	Média tratado	P-Valor ¹	Tamanho do efeito		Média controle	Média tratado	P-Valor ¹	Tamanho do efeito					
				Quant (dp)	Qualitativo				Quant (dp)	Qualitativo				
BFI - E ²	26,6	26,4	0,72	0,03	122	18,9%	26,6	26,9	0,32	0,06	53	33,5%		
BFI - A	33,5	33,6	0,42	0,01	43	6,7%	34,9	34,9	0,51	0,00	3	1,9%		
BFI - C	30,5	30,9	0,15	0,06	268	41,6%	32,5	33,1	0,20	0,10	110	69,0%		
BFI - N ³	21,9	21,9	0,48	0,00	10	1,6%	21,4	21,5	0,52	0,01	7	4,1%		
BFI - O	33,5	34,0	0,10	*	0,07	309	47,8%	33,9	35,3	0,04	**	0,21	232	145,9%
NSLCS ³	9,9	9,3	0,01	***	0,20	198	64,2%	8,8	8,6	0,32		0,09	19	27,7%
GS	3,4	3,4	0,29		0,05	8	2,7%	3,5	3,6	0,17		0,18	6	8,6%
SDQ - Em ³	3,8	3,9	0,72		0,05	36	11,0%	3,7	4,0	0,81		0,16	31	33,9%
SDQ - Cond ³	3,2	3,1	0,22		0,06	42	12,7%	2,8	2,8	0,45		0,02	4	4,3%
SDQ - Col ³	2,9	2,9	0,49		0,00	1	0,3%	2,3	2,7	0,89		0,21	35	37,7%
SDQ - Hip ³	4,7	4,7	0,64		0,03	21	6,4%	4,4	4,7	0,80		0,13	28	30,4%
SDQ - PS	7,3	7,4	0,30		0,04	31	9,4%	7,7	7,8	0,47		0,02	3	2,8%
RSES	18,7	18,8	0,35		0,03	47	14,6%	19,6	19,1	0,77		0,12	48	51,9%
CSE	3,4	3,5	0,19		0,07	13	4,0%	3,5	3,5	0,62		0,05	3	2,8%
SEQC - Ac	25,0	25,3	0,17		0,08	117	37,9%	26,1	26,5	0,29		0,11	28	41,4%
SEQC - Sc	24,3	24,7	0,19		0,07	109	35,4%	24,7	24,6	0,53		0,01	4	5,4%
SEQC - Em	22,6	23,3	0,06	*	0,13	203	65,9%	22,9	23,8	0,16		0,19	61	90,3%
DS	7,6	7,7	0,19		0,05	81	13,3%	7,9	8,1	0,32		0,06	24	15,3%

¹: P-valor de teste unicaudal no sentido da desejabilidade social.

²: A dimensão extroversão é a menos consensual em relação ao sentido da desejabilidade social, principalmente em relação à monotonicidade do efeito. Considerou-se para a análise, maior extroversão relacionada a maior desejabilidade social.

³: Constructo em que há correlação negativa com desejabilidade social.

*, **, ***: estatisticamente significativa a 10%, 5% e 1% respectivamente.

★, ★★, ★★★: estatisticamente significativa a 10%, 5% e 1% respectivamente, mas no sentido oposto da desejabilidade social.

interna (SDQ-Col, RSES e SEQC-Em), sendo que a única que também apresentou evidência de manipulação foi a dimensão “autoeficácia emocional” do SEQC. Além dessas, vale notar que a escala de desejabilidade social também apresenta menor consistência interna no grupo de tratamento. Já em relação à amostra restrita (coluna 8), há evidências de menor consistência interna somente na dimensão “hiperatividade” do SDQ.

Por fim, na análise das correlações (figura 4) é possível notar em primeiro lugar que houve maior perturbação nos sinais das correlações em comparação com o grupo de controle (figura 2). Esperar-se-ia que as cores na figura fossem mais intensas, indicando maior correlação em módulo devido ao fator de manipulação, mas visualmente não é possível estabelecer um padrão de imediato. Ainda assim, é possível checar a magnitude das correlações com a escala de desejabilidade social: em 10 das 17 escalas houve um aumento dessas correlações em módulo no grupo de tratamento.

A figura 5 apresenta a matriz de correlação para a amostra restrita. O número de correlações com sinais contrários ao esperado pela literatura na amostra de tratamento é igual ao observado na amostra de controle (figura 3). Já em relação às correlações com

Tabela 11 – Consistência interna dos instrumentos nos diferentes grupos e amostras

	Amostra completa				Amostra restrita			
	Controle		Palestra	Palestra + Notificação	Controle		Palestra	Palestra + Notificação
	α	IC (90%)	α	α	α	IC (90%)	α	α
BFI - E	0,59	0,55	0,61	0,61	0,61	0,54	0,68	0,70
BFI - A	0,63	0,60	0,62	0,57	0,63	0,58	0,66	0,57
BFI - C	0,73	0,71	0,77	0,76	0,77	0,74	0,75	0,82
BFI - N	0,60	0,57	0,60	0,62	0,71	0,66	0,70	0,69
BFI - O	0,66	0,63	0,69	0,68	0,75	0,71	0,73	0,71
NSLCS	0,56	0,52	0,53	0,54	0,58	0,48	0,62	0,62
GS	0,57	0,52	0,55	0,57	0,65	0,57	0,58	0,64
SDQ - Em	0,53	0,48	0,59	0,58	0,48	0,37	0,66	0,57
SDQ - Cond	0,51	0,46	0,50	0,38	0,50	0,39	0,55	0,32
SDQ - Col	0,45	0,39	0,33	0,40	0,36	0,21	0,36	0,47
SDQ - Hip	0,56	0,51	0,52	0,54	0,59	0,50	0,47	0,69
SDQ - PS	0,63	0,59	0,70	0,58	0,65	0,57	0,71	0,55
RSES	0,72	0,69	0,65	0,72	0,72	0,66	0,68	0,78
CSE	0,68	0,65	0,71	0,65	0,70	0,64	0,76	0,68
SEQC - Ac	0,73	0,70	0,71	0,76	0,73	0,66	0,68	0,79
SEQC - Sc	0,65	0,61	0,66	0,58	0,67	0,58	0,68	0,62
SEQC - Em	0,71	0,67	0,66	0,66	0,73	0,66	0,73	0,69
DS	0,60	0,57	0,55	0,51	0,63	0,57	0,62	0,58

desejabilidade social, em 12 dos 17 constructos houve aumento absoluto dessa medida no grupo de tratamento.

Com isso, podemos inferir que o comportamento manipulador foi intensificado pela palestra. Esse comportamento é percebido de maneira mais acentuada na análise dos itens, o que de certa forma é menos preocupante pois parte do viés é diluído na agregação dos constructos. Contudo, mesmo após a agregação dos itens, houve viés significativo em alguns instrumentos. A manipulação causada pela palestra parece ter afetado negativamente ao menos 4 instrumentos em relação à consistência interna na amostra completa e apenas 1 na amostra restrita. Já em relação às correlações, parece ter tido um efeito em tornar as respostas dos alunos desse grupo mais correlacionadas com desejabilidade social.

4.2 Efeito da notificação sobre a informação

Quando são comparadas as respostas aos itens dos grupo de controle e aquele que recebeu a palestra e teve a notificação de que respostas inconsistentes anulariam o questionário (tabela 12), tem-se evidências de que a notificação de fato reduziu a

Figura 4 – Matriz de correlação dos constructos - Grupo de tratamento: palestra - Amostra completa

	E	BFI					DS	SEQC			NSLCS	GS	SDQ					CSE	RSES
		A	C	N	O			Ac	Sc	Em			Em	Cond	Hip	Col	PS		
E																			
A	0,08																		
BFI																			
C	0,02	0,35																	
N	0,06	-0,35	-0,14																
O	0,23	0,20	0,34	0,08															
DS	-0,09	0,39	0,36	-0,27	0,03														
Ac	0,06	0,37	0,60	0,02	0,33	0,31													
SEQC																			
Sc	0,46	0,22	0,15	0,09	0,44	-0,02	0,36												
Em	0,13	0,27	0,17	-0,39	0,18	0,15	0,22	0,42											
NSLCS	-0,05	-0,18	-0,30	0,22	-0,04	-0,30	-0,18	-0,11	-0,19										
GS	-0,01	0,27	0,58	-0,10	0,27	0,45	0,46	0,19	0,23	-0,27									
Em	-0,09	-0,09	-0,00	0,41	0,15	-0,30													
Cond	0,01	-0,35	-0,22	0,35	-0,04	-0,39							0,38						
SDQ																			
Hip	0,25	-0,14	-0,55	0,21	-0,07	-0,37							0,19	0,32					
Col	-0,11	-0,21	-0,06	0,14	-0,04	-0,27							0,28	0,32	0,15				
PS	0,22	0,50	0,33	-0,06	0,30	0,27							0,08	-0,18	-0,09	-0,16			
CSE	0,10	0,37	0,40	-0,33	0,06	0,47							-0,40	-0,43	-0,33	-0,36	0,30		
RSES	0,15	0,26	0,27	-0,34	0,15	0,32							-0,39	-0,34	-0,26	-0,34	0,10	0,64	

Nota: Em destaque correlações com sinais contrários ao esperado pela literatura. BFI-E está marcado pois não há muito consenso sobre o sinal de suas correlações com os demais constructos.

Figura 5 – Matriz de correlação dos constructos - Grupo de tratamento: palestra - Amostra restrita

	E	BFI					DS	SEQC			NSLCS	GS	SDQ					CSE	RSES
		A	C	N	O			Ac	Sc	Em			Em	Cond	Hip	Col	PS		
E																			
A	0,07																		
BFI																			
C	-0,08	0,27																	
N	0,06	-0,36	-0,10																
O	0,23	0,08	0,36	0,10															
DS	-0,13	0,51	0,38	-0,34	0,03														
Ac	0,05	0,32	0,61	0,03	0,28	0,25													
SEQC																			
Sc	0,52	0,14	-0,09	0,11	0,38	0,03	0,18												
Em	0,22	0,22	0,12	-0,39	0,23	0,16	0,07	0,43											
NSLCS	-0,21	-0,22	-0,39	0,28	-0,10	-0,38	-0,21	-0,27	-0,26										
GS	0,01	0,28	0,61	0,01	0,29	0,47	0,48	0,16	0,24	-0,33									
Em	0,07	-0,20	-0,04	0,46	0,12	-0,41													
Cond	0,07	-0,38	-0,28	0,41	0,02	-0,47							0,58						
SDQ																			
Hip	0,32	-0,06	-0,59	0,26	0,07	-0,39							0,22	0,42					
Col	-0,13	-0,26	-0,08	0,20	0,02	-0,23							0,47	0,36	0,16				
PS	0,33	0,49	0,16	-0,11	0,22	0,21							0,07	-0,27	-0,04	-0,16			
CSE	-0,06	0,32	0,36	-0,38	0,03	0,51							-0,57	-0,56	-0,36	-0,39	0,26		
RSES	-0,03	0,15	0,32	-0,34	0,07	0,44							-0,48	-0,42	-0,39	-0,31	0,06	0,82	

Nota: Em destaque correlações com sinais contrários ao esperado pela literatura. BFI-E está marcado pois não há muito consenso sobre o sinal de suas correlações com os demais constructos.

manipulação na dimensão “conscienciosidade” do BFI, ou seja, seus 2 itens manipulados no grupo sem notificação deixaram de ser e houve uma redução dos níveis reportados de

conscienciosidade no grupo com notificação em comparação com o grupo de controle em 4 itens, de forma que a estatística Z_l é igual a 6, dentro da região crítica. Pode-se afirmar também que a notificação foi estatisticamente eficaz para reduzir a manipulação nos itens da dimensão “autoeficácia acadêmica” do SEQC ($Z_l = 3$) e no instrumento NSLCS, com 5 de seus 6 itens manipulados no grupo sem notificação deixando de ser e outros 5 itens aparecendo com diferenças significantes no sentido inverso da desejabilidade social ($Z_l = 10$). Além disso, a notificação parece ter tido efeito estatisticamente significativo também no instrumento GS, retirando a diferença que existia no sentido da manipulação em 2 itens e tornando 1 item com diferença no sentido contrário da desejabilidade ($Z_l = 3$). Na amostra restrita houve impacto da notificação nas medidas BFI-C e BFI-O.

Tabela 12 – Efeito da palestra e notificação sobre os itens dos constructos

	Amostra completa				Amostra restrita			
	Sentido DS	Sentido contrário	Efeito descontado		Sentido DS	Sentido contrário	Efeito descontado	
	N Itens	N Itens	Z_l	RC	N Itens	N Itens	Z_l	RC
BFI - E (8 itens)	1	1	-1	$(-\infty, -3] \cup [3, +\infty)$	0	0	0	$(-\infty, -3] \cup [3, +\infty)$
BFI - A (9 itens)	1	0	0	$(-\infty, -3] \cup [3, +\infty)$	0	2	2	$(-\infty, -3] \cup [3, +\infty)$
BFI - C (9 itens)	0	4	6	$(-\infty, -3] \cup [3, +\infty)$	0	6	7	$(-\infty, -3] \cup [3, +\infty)$
BFI - N (8 itens)	1	1	-1	$(-\infty, -3] \cup [3, +\infty)$	1	1	0	$(-\infty, -3] \cup [3, +\infty)$
BFI - O (10 itens)	1	0	1	$(-\infty, -4] \cup [3, +\infty)$	0	0	5	$(-\infty, -3] \cup [3, +\infty)$
NSLCS (21 itens)	1	5	10	$(-\infty, -3] \cup [3, +\infty)$	2	3	0	$(-\infty, -4] \cup [4, +\infty)$
GS (8 itens)	0	1	3	$(-\infty, -2] \cup [2, +\infty)$	0	0	1	$(-\infty, -3] \cup [3, +\infty)$
SDQ - Em (5 itens)	1	0	-2	$(-\infty, -3] \cup [2, +\infty)$	1	0	-2	$(-\infty, -3] \cup [2, +\infty)$
SDQ - Cond (5 itens)	1	0	1	$(-\infty, -3] \cup [3, +\infty)$	0	0	0	$(-\infty, -3] \cup [3, +\infty)$
SDQ - Col (5 itens)	1	1	-1	$(-\infty, -2] \cup [2, +\infty)$	0	1	1	$(-\infty, -2] \cup [3, +\infty)$
SDQ - Hip (5 itens)	0	0	0	$(-\infty, -3] \cup [3, +\infty)$	0	2	1	$(-\infty, -2] \cup [3, +\infty)$
SDQ - PS (5 itens)	1	0	0	$(-\infty, -3] \cup [3, +\infty)$	0	1	1	$(-\infty, -3] \cup [3, +\infty)$
RSES (10 itens)	2	0	-1	$(-\infty, -3] \cup [3, +\infty)$	1	1	-1	$(-\infty, -3] \cup [3, +\infty)$
CSE (12 itens)	1	0	1	$(-\infty, -3] \cup [3, +\infty)$	0	0	0	$(-\infty, -3] \cup [3, +\infty)$
SEQC - Ac (7 itens)	0	0	3	$(-\infty, -2] \cup [2, +\infty)$	0	1	2	$(-\infty, -3] \cup [3, +\infty)$
SEQC - Sc (7 itens)	2	0	-1	$(-\infty, -2] \cup [2, +\infty)$	2	1	-2	$(-\infty, -3] \cup [3, +\infty)$
SEQC - Em (7 itens)	2	0	1	$(-\infty, -2] \cup [2, +\infty)$	0	0	1	$(-\infty, -3] \cup [3, +\infty)$

Já a tabela 13 apresenta os resultados do efeito da notificação sobre as médias dos constructos. A medida de "conscienciosidade" no grupo de controle foi de 30,5 e no grupo de tratamento 30,0, diferença significativa no sentido contrário da desejabilidade social, efeito equivalente a 0,08 d.p. ou 51,6% dos alunos do grupo de controle alterando sua resposta para uma pontuação a menos no sentido da desejabilidade social de um item. A evidência de que a média nessa dimensão é inferior no grupo de tratamento é corroborada e amplificada na amostra restrita. Além dessa, a dimensão "amabilidade" do mesmo instrumento também apresenta o mesmo tipo de evidência na amostra restrita. Esse tipo de comportamento

entre grupos será discutido ao fim da subseção. No sentido oposto, houve diferença estatisticamente significativa na dimensão “autoeficácia social” do instrumento SEQC, com o grupo de tratamento apresentando médias superiores ao grupo de controle, 24,9 ante 24,3, efeito de 0,12 d.p. ou 57,2% dos alunos do grupo de controle alterando em um ponto para mais um dos itens no sentido da desejabilidade social. Esse último resultado é bastante contra intuitivo uma vez que aparentemente o comportamento manipulador dos alunos não ocorreu nessa dimensão quando ouviram a palestra, mas ocorreu quando ouviram a palestra e foram notificados de que o comportamento manipulador identificado anulava o questionário.

Tabela 13 – Efeito da palestra e notificação nos constructos

	Amostra completa						Amostra restrita							
	Média controle	Média tratado	P-Valor ¹	Tamanho do efeito		Média controle	Média tratado	P-Valor ¹	Tamanho do efeito					
				Quant (dp)	Qualitativo				Quant (dp)	Qualitativo				
BFI - E ²	26,6	26,7	0,33	0,03	94	14,6%	26,6	26,7	0,45	0,02	15	9,4%		
BFI - A	33,5	33,5	0,48	0,00	11	1,7%	34,9	34,1	0,90	★	0,15	131	82,4%	
BFI - C	30,5	30,0	0,91	★	0,08	333	51,6%	32,5	30,4	0,99	★ ★ ★	0,30	329	206,9%
BFI - N ³	21,9	22,1	0,75		0,04	145	22,4%	21,4	22,1	0,80		0,10	102	64,0%
BFI - O	33,5	33,5	0,49		0,00	6	1,0%	33,9	34,2	0,33		0,05	52	32,6%
NSLCS ³	9,9	10,1	0,73		0,05	47	15,3%	8,8	9,3	0,83		0,16	33	48,9%
GS	3,4	3,3	0,87		0,09	17	5,5%	3,5	3,5	0,52		0,01	0	0,4%
SDQ - Em ³	3,8	3,7	0,29		0,05	34	10,3%	3,7	3,3	0,13		0,18	37	39,8%
SDQ - Cond ³	3,2	3,2	0,50		0,00	0	0,1%	2,8	2,9	0,59		0,03	6	6,6%
SDQ - Col ³	2,9	2,8	0,25		0,05	36	10,8%	2,3	2,7	0,87		0,17	30	32,1%
SDQ - Hip ³	4,7	4,8	0,79		0,06	49	14,8%	4,4	4,6	0,69		0,09	19	20,2%
SDQ - PS	7,3	7,5	0,12		0,09	65	19,7%	7,7	7,5	0,75		0,11	19	20,7%
RSES	18,7	19,0	0,20		0,07	103	31,8%	19,6	19,7	0,43		0,03	12	12,6%
CSE	3,4	3,5	0,29		0,04	8	2,4%	3,5	3,5	0,48		0,01	0	0,4%
SEQC - Ac	25,0	25,1	0,35		0,03	46	14,8%	26,1	25,4	0,82		0,18	47	70,7%
SEQC - Sc	24,3	24,9	0,06	*	0,12	176	57,2%	24,7	25,1	0,27		0,11	31	46,4%
SEQC - Em	22,6	23,1	0,15		0,08	134	43,6%	22,9	23,0	0,42		0,03	11	16,3%
DS	7,6	7,5	0,70		0,03	48	7,8%	8,0	8,1	0,39		0,05	18	11,4%

¹: P-valor de teste unicaudal no sentido da desejabilidade social.

²: A dimensão extroversão é a menos consensual em relação ao sentido da desejabilidade social, principalmente em relação à monotonicidade do efeito. Considerou-se para a análise, maior extroversão relacionada a maior desejabilidade social.

³: Constructo em que há correlação negativa com desejabilidade social.

*, **, ***: estatisticamente significativa a 10%, 5% e 1% respectivamente.

★, ★★, ★★*: estatisticamente significativa a 10%, 5% e 1% respectivamente, mas no sentido oposto da desejabilidade social.

Voltando à tabela 11 é possível comparar a consistência interna dos instrumentos nesse grupo (coluna 5) em relação ao grupo de controle. No caso anterior, de comparação entre grupo que recebeu palestra sem notificação e grupo de controle, era esperado que o comportamento manipulador fosse estimulado pela palestra através do mecanismo da desejabilidade social e, portanto, a consistência interna dos instrumentos fosse reduzida no grupo de tratamento. Já nesse caso, o tratamento é composto pela palestra, que agiria

no sentido da desejabilidade social intensificando o comportamento manipulador, e da notificação, que agiria como barreira para esse comportamento, de forma que não se sabe qual o efeito mais preponderante. Caso a consistência interna dos instrumentos não seja menor no grupo de tratamento, é provável que o efeito da notificação tenha sido eficaz para reduzir o comportamento manipulador. Comparando esses coeficientes entre grupos é possível notar que houve redução da consistência interna em 5 escalas que medem competências socioemocionais, além da escala de desejabilidade social no grupo de tratamento, número maior, inclusive, do que no grupo em que os alunos receberam a palestra, mas não a notificação. Contudo, nas 4 escalas em que a estatística Z_I mostrou-se significativa não houve redução na consistência interna, o que indicaria efeito preponderante da notificação em detrimento do efeito da palestra nesses constructos. Já na amostra reduzida são 4 escalas que apresentam menos consistência interna nesse grupo em comparação com o grupo de controle com 90% de confiabilidade, e nesse caso inclui-se uma das dimensões cuja estatística Z_I mostrou-se significativa, não deixando claro qual o efeito mais preponderante nesse tratamento composto pela palestra e notificação.

A figura 6 apresenta a matriz de correlação para o grupo de tratamento que recebeu a palestra e a notificação. Houve alteração de sinal da correlação em relação ao esperado em 4 casos identificados na figura. Comparando-a com a figura 2, tem-se que há um aumento absoluto das correlações com a escala de desejabilidade social em 9 constructos, número menor do que aquele observado na amostra que recebeu a palestra, mas não a notificação (12).

Já em relação à amostra restrita (figura 7), pode-se notar que houve uma perturbação causada pelo tratamento, alterando o sinal esperado de 5 correlações. Além disso, 10 das 18 medidas aumentaram a correlação em módulo com desejabilidade social em comparação com o grupo de controle, contra 12 do grupo de tratamento que não recebeu a notificação.

O fato de haver maior correlação dos itens com desejabilidade social nesse tratamento em relação ao grupo de controle, mas em intensidade menor do que no grupo

Figura 6 – Matriz de correlação dos constructos - Grupo de tratamento: palestra e notificação - Amostra completa

	E	A	C	N	O	DS	Ac	Sc	Em	NSLCS	GS	Em	Cond	Hip	Col	PS	CSE	RSES
E																		
A	-0,04																	
BFI C	-0,06	0,33																
N	0,09	-0,40	-0,24															
O	0,19	0,21	0,21	-0,03														
DS	-0,07	0,40	0,40	-0,29	0,11													
SEQC Ac	0,06	0,24	0,63	-0,02	0,25	0,24												
Sc	0,40	0,07	0,15	0,09	0,42	-0,02	0,32											
Em	0,02	0,24	0,21	-0,28	0,19	0,22	0,26	0,28										
NSLCS	0,00	-0,13	-0,21	0,20	-0,11	-0,32	-0,09	0,02	-0,05									
GS	0,01	0,28	0,57	-0,17	0,27	0,39	0,54	0,11	0,21	-0,36								
SDQ Em	-0,09	-0,14	-0,21	0,42	-0,06	-0,23												
Cond	0,09	-0,45	-0,32	0,42	-0,10	-0,44						0,32						
Hip	0,16	-0,36	-0,50	0,38	-0,06	-0,39						0,21	0,34					
Col	-0,19	-0,26	-0,18	0,18	-0,03	-0,22						0,28	0,22	0,14				
PS	0,06	0,45	0,38	-0,13	0,18	0,26						0,10	-0,18	-0,19	-0,10			
CSE	0,14	0,33	0,40	-0,38	0,18	0,39						-0,44	-0,36	-0,31	-0,35	0,20		
RSES	0,22	0,20	0,22	-0,27	0,16	0,28						-0,40	-0,22	-0,15	-0,30	0,11	0,61	

Nota: Em destaque correlações com sinais contrários ao esperado pela literatura. BFI-E está marcado pois não há muito consenso sobre o sinal de suas correlações com os demais constructos.

Figura 7 – Matriz de correlação dos constructos - Grupo de tratamento: palestra e notificação - Amostra restrita

	E	A	C	N	O	DS	Ac	Sc	Em	NSLCS	GS	Em	Cond	Hip	Col	PS	CSE	RSES
E																		
A	-0,04																	
BFI C	-0,07	0,32																
N	0,11	-0,39	-0,18															
O	0,21	0,14	0,18	0,01														
DS	-0,13	0,48	0,52	-0,29	0,15													
SEQC Ac	0,20	0,25	0,68	-0,01	0,37	0,33												
Sc	0,53	0,04	0,20	0,15	0,45	0,06	0,30											
Em	0,08	0,27	0,29	-0,29	0,21	0,33	0,32	0,25										
NSLCS	0,03	-0,13	-0,41	0,09	-0,30	-0,51	-0,20	-0,12	-0,01									
GS	0,09	0,29	0,70	-0,05	0,38	0,51	0,61	0,18	0,35	-0,31								
SDQ Em	-0,03	-0,12	-0,33	0,38	-0,03	-0,16												
Cond	0,15	-0,49	-0,38	0,41	0,01	-0,52						0,29						
Hip	0,32	-0,33	-0,67	0,40	0,14	-0,46						0,20	0,41					
Col	-0,16	-0,25	-0,23	0,23	0,16	-0,29						0,28	0,23	0,21				
PS	-0,00	0,46	0,35	-0,25	0,08	0,28						-0,11	-0,32	-0,17	0,02			
CSE	0,10	0,33	0,52	-0,34	0,06	0,42						-0,48	-0,38	-0,35	-0,46	0,24		
RSES	0,15	0,23	0,37	-0,32	0,18	0,25						-0,48	-0,29	-0,26	-0,43	0,11	0,63	

Nota: Em destaque correlações com sinais contrários ao esperado pela literatura. BFI-E está marcado pois não há muito consenso sobre o sinal de suas correlações com os demais constructos.

que assistiu à palestra mas não recebeu a notificação na amostra completa, sugere que a notificação não foi suficiente para barrar por completo o comportamento manipulador, mas

serviu para reduzi-lo como um todo. Essa evidência é suportada também pela análise dos itens e médias dos constructos. Dimensões e instrumentos que se mostraram impactados de forma significativa no grupo que apenas assistiu à palestra, deixaram de apresentar manipulação no grupo com a notificação. Mais do que isso, em alguns instrumentos o efeito da notificação foi ainda mais intenso, criando um padrão de resposta em que os alunos reportam sistematicamente menores níveis de competências do que no grupo de controle. Esse resultado parece indicar que a notificação barra não somente o comportamento manipulador intensificado pela palestra (m_{ij}), mas também aquele relacionado ao ambiente escolar como um todo (m_{ij}^e).

4.3 Efeito dos professores receberem os resultados

O fato de informar os alunos que seus professores receberiam suas respostas dos questionários impactou significativamente os itens dos instrumentos SDQ-PS e RSES apenas, como pode ser visto na tabela 14. Já na amostra reduzida essa evidência é encontrada em 3 constructos: RSES, SEQC-Sc e SEQC-Em. Além disso, vale notar os valores de X_l mesmo quando não mostram-se significantes: há muitos casos em que $X_l = 0$ (10 de 17 na amostra completa e 7 de 17 na amostra restrita) indicando que o esse tratamento parece ter impacto menor nos itens dos instrumentos vis a vis aquele da palestra.

A tabela 15 apresenta a análise dos constructos. Nota-se viés estatisticamente significativo em 2 escalas em cada amostra. Na amostra completa há indícios de viés na dimensão "pró-social" do SDQ da magnitude de 0,12 d.p. ou pouco mais de 25% dos alunos do grupo de controle alterarem suas respostas em uma pontuação no sentido da desejabilidade social em um dos itens da escala e na dimensão "emocional" do SEQC, efeito equivalente a 0,12 d.p. ou 62,6% dos alunos do grupo de controle alterando suas respostas. Já em relação à amostra restrita, há indícios de viés novamente no constructo SEQC-Em, em magnitudes maiores que aquelas da amostra completa e também na dimensão "social" do mesmo instrumento. Impacto esse bastante alto, de 0,5 d.p. equivalente a todos os alunos

Tabela 14 – Efeito dos professores receberem respostas individuais sobre os itens dos constructos

	Amostra completa					Amostra restrita				
	Sentido DS	Sentido contrário	Efeito descontado			Sentido DS	Sentido contrário	Efeito descontado		
	N Itens	N Itens	X_i	RC		N Itens	N Itens	X_i	RC	
BFI - E (8 itens)	1	2	-1	$(-\infty, -2]$	$U [3, +\infty)$	0	0	0	$(-\infty, -2]$	$U [3, +\infty)$
BFI - A (9 itens)	1	1	0	$(-\infty, -2]$	$U [2, +\infty)$	2	1	1	$(-\infty, -2]$	$U [3, +\infty)$
BFI - C (9 itens)	1	1	0	$(-\infty, -2]$	$U [2, +\infty)$	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$
BFI - N (8 itens)	2	0	2	$(-\infty, -2]$	$U [3, +\infty)$	2	1	1	$(-\infty, -2]$	$U [2, +\infty)$
BFI - O (10 itens)	2	2	0	$(-\infty, -3]$	$U [3, +\infty)$	1	2	-1	$(-\infty, -3]$	$U [3, +\infty)$
NSLCS (21 itens)	1	2	-1	$(-\infty, -4]$	$U [3, +\infty)$	2	3	-1	$(-\infty, -3]$	$U [3, +\infty)$
GS (8 itens)	0	1	-1	$(-\infty, -3]$	$U [3, +\infty)$	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$
SDQ - Em (5 itens)	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$	0	1	-1	$(-\infty, -2]$	$U [2, +\infty)$
SDQ - Cond (5 itens)	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$
SDQ - Col (5 itens)	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$
SDQ - Hip (5 itens)	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$
SDQ - PS (5 itens)	2	0	2	$(-\infty, -2]$	$U [2, +\infty)$	2	1	1	$(-\infty, -2]$	$U [2, +\infty)$
RSES (10 itens)	2	0	2	$(-\infty, -3]$	$U [2, +\infty)$	4	0	4	$(-\infty, -3]$	$U [3, +\infty)$
CSE (12 itens)	0	1	-1	$(-\infty, -3]$	$U [3, +\infty)$	0	0	0	$(-\infty, -2]$	$U [3, +\infty)$
SEQC - Ac (7 itens)	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$	1	1	0	$(-\infty, -3]$	$U [2, +\infty)$
SEQC - Sc (7 itens)	0	0	0	$(-\infty, -2]$	$U [2, +\infty)$	3	0	3	$(-\infty, -3]$	$U [2, +\infty)$
SEQC - Em (7 itens)	0	0	0	$(-\infty, -3]$	$U [2, +\infty)$	3	0	3	$(-\infty, -2]$	$U [2, +\infty)$

do grupo de controle mudarem suas respostas em uma pontuação acima no sentido da desejabilidade social em mais de 2 itens do constructo.

Tabela 15 – Efeito dos professores receberem respostas individuais nos constructos

	Amostra completa					Amostra restrita				
	Média controle	Média tratado	P-Valor ¹	Tamanho do efeito		Média controle	Média tratado	P-Valor ¹	Tamanho do efeito	
				Quant (dp)	Qualitativo				Quant (dp)	Qualitativo
BFI - E ²	26,6	26,2	0,85	0,06	218 33,8%	26,6	26,2	0,70	0,06	57 36,1%
BFI - A	33,5	33,3	0,69	0,03	110 17,0%	34,9	34,9	0,47	0,01	9 5,6%
BFI - C	30,5	30,4	0,57	0,01	47 7,3%	32,5	32,0	0,68	0,06	64 40,1%
BFI - N ³	21,9	21,9	0,47	0,00	18 2,8%	21,4	21,0	0,31	0,06	62 38,9%
BFI - O	33,5	33,5	0,44	0,01	38 5,8%	33,9	33,9	0,51	0,00	3 1,9%
NSLCS ³	9,9	10,0	0,52	0,00	4 1,3%	8,8	9,0	0,59	0,05	9 13,9%
GS	3,4	3,4	0,55	0,01	2 0,7%	3,5	3,5	0,25	0,15	5 7,3%
SDQ - Em ³	3,8	4,0	0,83	0,08	58 17,5%	3,7	3,8	0,59	0,04	7 7,6%
SDQ - Cond ³	3,2	3,3	0,70	0,04	30 9,2%	2,8	2,8	0,47	0,01	2 2,4%
SDQ - Col ³	2,9	2,9	0,38	0,02	16 4,9%	2,3	2,6	0,75	0,11	19 20,8%
SDQ - Hip ³	4,7	4,9	0,84	0,08	61 18,5%	4,4	4,7	0,79	0,14	29 32,1%
SDQ - PS	7,3	7,5	0,07 *	0,12	83 25,2%	7,7	7,9	0,33	0,08	14 15,2%
RSES	18,7	19,0	0,19	0,07	105 32,6%	19,6	20,3	0,16	0,17	65 70,5%
CSE	3,4	3,4	0,68	0,04	7 2,0%	3,5	3,6	0,32	0,07	4 3,8%
SEQC - Ac	25,0	25,0	0,47	0,01	11 3,5%	26,1	26,7	0,22	0,15	39 58,4%
SEQC - Sc	24,3	24,6	0,23	0,06	91 29,6%	24,7	26,8	0,00 ***	0,49	141 211,2%
SEQC - Em	22,6	23,2	0,07 *	0,12	193 62,6%	22,9	24,3	0,05 **	0,30	95 142,3%
DS	7,6	7,8	0,23	0,06	48 15,4%	7,9	8,3	0,21	0,14	62 39,1%

¹: P-valor de teste unicaudal no sentido da desejabilidade social.

²: A dimensão extroversão é a menos consensual em relação ao sentido da desejabilidade social, principalmente em relação à monotonicidade do efeito. Considerou-se para a análise, maior extroversão relacionada a maior desejabilidade social.

³: Constructo em que há correlação negativa com desejabilidade social.

*, **, ***: estatisticamente significativa a 10%, 5% e 1% respectivamente.

★, ★★, ★★★: estatisticamente significativa a 10%, 5% e 1% respectivamente, mas no sentido oposto da desejabilidade social.

Analisando os indicadores de consistência interna dos instrumentos apresentados na tabela 16, pode-se notar que houve redução da consistência interna com 90% de confiança

em 3 constructos, sendo que em somente um deles (SEQC-Em) houve indicativo de manipulação nas análises via itens ou médias apresentadas anteriormente. Já em relação à amostra de questionários inteiramente respondidos, houve redução em 5 escalas, sendo que em 2 delas houve indício de manipulação sistemática nas análises anteriores (SEQC-Em e RSES).

Tabela 16 – Consistência interna dos instrumentos nos diferentes grupos e amostras

	Amostra completa				Amostra restrita			
	Controle		Professor	Professor + Notificação	Controle		Professor	Professor + Notificação
	α	IC (90%)	α	α	α	IC (90%)	α	α
BFI - E	0,59	0,55	0,59	0,60	0,61	0,54	0,61	0,68
BFI - A	0,63	0,60	0,59	0,63	0,63	0,58	0,61	0,63
BFI - C	0,73	0,71	0,73	0,76	0,77	0,74	0,79	0,74
BFI - N	0,60	0,57	0,62	0,60	0,71	0,66	0,69	0,67
BFI - O	0,66	0,63	0,66	0,69	0,75	0,71	0,59	0,73
NSLCS	0,56	0,52	0,62	0,55	0,58	0,48	0,72	0,64
GS	0,57	0,52	0,58	0,55	0,65	0,57	0,73	0,69
SDQ - Em	0,53	0,48	0,49	0,58	0,48	0,37	0,33	0,58
SDQ - Cond	0,51	0,46	0,43	0,43	0,50	0,39	0,60	0,55
SDQ - Col	0,45	0,39	0,41	0,36	0,36	0,21	0,37	0,34
SDQ - Hip	0,56	0,51	0,52	0,55	0,59	0,50	0,57	0,59
SDQ - PS	0,63	0,59	0,63	0,57	0,65	0,57	0,74	0,44
RSES	0,72	0,69	0,71	0,70	0,72	0,66	0,64	0,74
CSE	0,68	0,65	0,66	0,74	0,70	0,64	0,61	0,77
SEQC - Ac	0,73	0,70	0,77	0,76	0,73	0,66	0,79	0,78
SEQC - Sc	0,65	0,61	0,68	0,66	0,67	0,58	0,65	0,65
SEQC - Em	0,71	0,67	0,64	0,61	0,73	0,66	0,65	0,65
DS	0,60	0,57	0,57	0,60	0,63	0,57	0,67	0,67

Em relação às correlações dos instrumentos, há alteração de sinal em relação ao teoricamente esperado em dois casos apenas apontados na figura 8, indicando que esse tratamento parece não ter afetado de forma tão significativa essa propriedade dos instrumentos. Além disso, houve maior número de constructos cujas correlações com desejabilidade social caíram em módulo em comparação com o grupo de controle (figura 2), resultado contraintuitivo. Já utilizando a amostra de questionários inteiramente respondidos (figura 9), é possível notar uma perturbação bastante grande nos sinais nas correlações, 9 casos. Fora isso, há um aumento da correlação absoluta de 9 escalas com desejabilidade social em comparação com o grupo de controle (figura 3).

De acordo com esses resultados, pode-se afirmar que o efeito de os professores receberem as respostas individuais dos alunos é menor do que o observado através do

Figura 8 – Matriz de correlação dos constructos - Grupo de tratamento: professor receber respostas individuais - Amostra completa

	E	A	C	N	O	DS	Ac	Sc	Em	NSLCS	GS	Em	Cond	Hip	Col	PS	CSE	RSES
E																		
A	0,09																	
BFI C	-0,06	0,39																
N	0,04	-0,34	-0,24															
O	0,27	0,26	0,25	-0,04														
DS	-0,06	0,40	0,38	-0,35	0,10													
SEQC Ac	-0,09	0,39	0,63	-0,16	0,26	0,40												
Sc	0,41	0,26	0,10	0,06	0,41	0,08	0,16											
Em	0,16	0,35	0,26	-0,34	0,29	0,30	0,32	0,37										
NSLCS	0,01	-0,32	-0,40	0,21	-0,18	-0,35	-0,35	-0,16	-0,20									
GS	-0,04	0,33	0,56	-0,18	0,20	0,41	0,55	0,13	0,29	-0,38								
Em	-0,08	-0,22	-0,04	0,43	0,02	-0,20												
Cond	0,01	-0,43	-0,33	0,35	-0,07	-0,38						0,34						
SDQ Hip	0,10	-0,28	-0,57	0,21	-0,14	-0,32						0,26	0,32					
Col	-0,18	-0,26	-0,14	0,20	-0,07	-0,27						0,27	0,23	0,13				
PS	0,08	0,43	0,33	-0,05	0,31	0,23						-0,00	-0,29	-0,25	-0,23			
CSE	0,16	0,30	0,37	-0,32	0,30	0,30						-0,38	-0,33	-0,36	-0,33	0,26		
RSES	0,15	0,26	0,28	-0,24	0,28	0,17						-0,26	-0,22	-0,27	-0,29	0,21	0,59	

Nota: Em destaque correlações com sinais contrários ao esperado pela literatura. BFI-E está marcado pois não há muito consenso sobre o sinal de suas correlações com os demais constructos.

Figura 9 – Matriz de correlação dos constructos - Grupo de tratamento: professor receber respostas individuais - Amostra restrita

	E	A	C	N	O	DS	Ac	Sc	Em	NSLCS	GS	Em	Cond	Hip	Col	PS	CSE	RSES
E																		
A	0,11																	
BFI C	-0,03	0,49																
N	0,11	-0,33	-0,23															
O	0,26	0,10	0,15	0,09														
DS	-0,04	0,40	0,47	-0,37	0,00													
SEQC Ac	-0,08	0,44	0,79	-0,13	0,18	0,46												
Sc	0,51	0,14	-0,01	-0,00	0,11	0,09	0,01											
Em	0,12	0,40	0,37	-0,35	0,34	0,40	0,42	0,17										
NSLCS	0,20	-0,39	-0,50	0,37	-0,15	-0,49	-0,37	-0,05	-0,23									
GS	-0,04	0,37	0,73	-0,14	0,30	0,49	0,76	0,13	0,37	-0,56								
Em	-0,11	-0,11	-0,06	0,34	0,09	-0,17												
Cond	-0,08	-0,55	-0,17	0,28	0,03	-0,42						0,25						
SDQ Hip	0,10	-0,33	-0,62	0,08	-0,13	-0,39						0,24	0,31					
Col	-0,31	-0,26	-0,02	0,19	0,05	-0,06						0,19	0,27	0,02				
PS	0,22	0,42	0,33	0,06	0,42	0,07						0,10	-0,26	-0,28	-0,22			
CSE	0,14	0,44	0,38	-0,19	0,30	0,18						-0,44	-0,28	-0,30	-0,26	0,25		
RSES	0,12	0,36	0,05	0,01	0,24	-0,15						-0,03	-0,08	-0,11	-0,29	0,20	0,56	

Nota: Em destaque correlações com sinais contrários ao esperado pela literatura. BFI-E está marcado pois não há muito consenso sobre o sinal de suas correlações com os demais constructos.

tratamento da palestra na amostra completa. Além de o número de instrumentos cujos itens ou médias que se mostraram manipulados ser menor do que aqueles observados no

caso da palestra, a correlação com desejabilidade social mais caiu do que aumentou nos instrumentos. Esse efeito menor no comportamento manipulador pode ser resultado de um tratamento mais fraco, causando *partial compliance*, ou seja, alunos considerados tratados podem na verdade ser do grupo de controle uma vez que o mecanismo que configura o tratamento é uma simples informação aos alunos antes do preenchimento do questionário de duração de no máximo 15 segundos. Dessa forma, é possível que alguns alunos da sala de aula não tenham ouvido essa informação, ou seja, tenham preenchido os questionários sem terem sido expostos de forma efetiva a nenhum tratamento. Vale notar, contudo, que mesmo em menor efeito, houve viés significativo em ao menos 2 constructos mesmo após a agregação dos itens.

A conclusão retirada através da amostra restrita é diferente: nessa amostra, o tratamento de avisar os alunos que os professores receberiam os resultados individuais parece ter um maior efeito do que a palestra no comportamento manipulador nas análises de itens e médias, além de ter afetado de forma mais significativa a consistência interna das escalas, assim como a correlação esperada entre elas. Uma possível explicação para esse resultado é que ao considerar apenas os gabaritos inteiramente respondidos, pode-se ter resolvido em grande parte o problema de *partial compliance* uma vez que deve haver correlação entre responder o questionário completamente e ter prestado atenção nas instruções do aplicador, ambos os fatores relacionados ao comprometimento dos alunos.

4.4 Efeito da Notificação sobre os professores receberem os resultados

Analisando o efeito da notificação nas respostas dos alunos que foram informados que os professores receberiam suas respostas individuais (tabela 17), é possível notar impacto significativo no sentido esperado, isso é, redução do comportamento manipulador, em 3 constructos: BFI-C, SDQ-PS e RSES. Por outro lado, o instrumento NSLCS apresentou estatística Z_l negativa e estatisticamente significante, indicando que a notificação estimulou as respostas no sentido da desejabilidade social, o que é bastante contraintuitivo. Já na

amostra restrita a notificação afetou número maior de instrumentos no sentido esperado: BFI-C, RSES, SEQC-Ac e SEQC-Em.

Tabela 17 – Efeito da notificação e dos professores receberem respostas individuais sobre os itens dos constructos

	Amostra completa				Amostra restrita			
	Sentido	Sentido	Efeito		Sentido	Sentido	Efeito	
	DS	contrário	descontado		DS	contrário	descontado	
	N Itens	N Itens	Z_l	RC	N Itens	N Itens	Z_l	RC
BFI - E (8 itens)	1	3	1	$(-\infty, -3] \cup [3, +\infty)$	0	1	1	$(-\infty, -3] \cup [3, +\infty)$
BFI - A (9 itens)	0	1	1	$(-\infty, -3] \cup [3, +\infty)$	0	1	2	$(-\infty, -3] \cup [3, +\infty)$
BFI - C (9 itens)	0	3	3	$(-\infty, -3] \cup [3, +\infty)$	0	6	6	$(-\infty, -3] \cup [3, +\infty)$
BFI - N (8 itens)	1	0	1	$(-\infty, -3] \cup [3, +\infty)$	0	1	2	$(-\infty, -3] \cup [3, +\infty)$
BFI - O (10 itens)	1	2	1	$(-\infty, -4] \cup [3, +\infty)$	0	1	0	$(-\infty, -3] \cup [3, +\infty)$
NSLCS (21 itens)	4	2	-3	$(-\infty, -3] \cup [3, +\infty)$	2	3	0	$(-\infty, -4] \cup [4, +\infty)$
GS (8 itens)	1	1	-1	$(-\infty, -2] \cup [2, +\infty)$	0	1	1	$(-\infty, -3] \cup [3, +\infty)$
SDQ - Em (5 itens)	1	1	0	$(-\infty, -3] \cup [2, +\infty)$	0	1	0	$(-\infty, -3] \cup [2, +\infty)$
SDQ - Cond (5 itens)	1	0	-1	$(-\infty, -3] \cup [3, +\infty)$	0	0	0	$(-\infty, -3] \cup [3, +\infty)$
SDQ - Col (5 itens)	0	0	0	$(-\infty, -2] \cup [2, +\infty)$	0	1	1	$(-\infty, -2] \cup [3, +\infty)$
SDQ - Hip (5 itens)	0	0	0	$(-\infty, -3] \cup [3, +\infty)$	0	0	0	$(-\infty, -2] \cup [3, +\infty)$
SDQ - PS (5 itens)	0	1	3	$(-\infty, -3] \cup [3, +\infty)$	0	1	2	$(-\infty, -3] \cup [3, +\infty)$
RSES (10 itens)	0	3	5	$(-\infty, -3] \cup [3, +\infty)$	0	3	7	$(-\infty, -3] \cup [3, +\infty)$
CSE (12 itens)	0	1	0	$(-\infty, -3] \cup [3, +\infty)$	0	1	1	$(-\infty, -3] \cup [3, +\infty)$
SEQC - Ac (7 itens)	0	1	1	$(-\infty, -2] \cup [2, +\infty)$	0	1	1	$(-\infty, -3] \cup [3, +\infty)$
SEQC - Sc (7 itens)	1	1	0	$(-\infty, -2] \cup [2, +\infty)$	1	1	3	$(-\infty, -3] \cup [3, +\infty)$
SEQC - Em (7 itens)	0	0	0	$(-\infty, -2] \cup [2, +\infty)$	0	2	5	$(-\infty, -3] \cup [3, +\infty)$

A tabela 18 apresenta os resultados dos constructos após a agregação dos itens. Pode-se notar que há evidências de que a notificação agiu no sentido de reportarem menores níveis das competências “conscienciosidade” do BFI e autoestima (RSES). No primeiro caso, o efeito é equivalente a 0,09 d.p. ou a 61,0% dos alunos do grupo de controle alterarem em uma pontuação a menos no sentido da desejabilidade social um item da escala. Já no instrumento RSES, o viés é da magnitude de 0,12 d.p. ou qualitativamente à mudança nas respostas de um item de 54,0% dos alunos do grupo de controle em uma pontuação a menos no sentido da desejabilidade social. Na amostra restrita o efeito é ainda maior tanto em número de constructos quanto em magnitude do viés: são 4 constructos (BFI-C, RSES, SEQC-Ac e SEQC-Em) sendo que na dimensão "conscienciosidade" do BFI o viés representa uma alteração nas respostas do total de alunos do grupo de controle em mais de 2 itens da escala em uma pontuação abaixo no sentido da desejabilidade social.

Em relação à consistência interna dos instrumentos, tem-se que esse tratamento composto reduziu com 90% de confiança a consistência interna de 4 escalas (tabela 16), mesmo número observado no tratamento em que os alunos foram informados que os

Tabela 18 – Efeito dos professores receberem respostas individuais e notificação nos constructos

	Amostra completa						Amostra restrita					
	Média controle	Média tratado	P-Valor ¹	Tamanho do efeito		Média controle	Média tratado	P-Valor ¹	Tamanho do efeito			
				Quant (dp)	Qualitativo				Quant (dp)	Qualitativo		
BFI - E ²	26,6	26,3	0,76	0,04	147 22,7%	26,6	25,8	0,86	0,14	126 79,1%		
BFI - A	33,5	33,3	0,73	0,03	130 20,1%	34,9	34,7	0,62	0,04	33 20,9%		
BFI - C	30,5	29,9	0,94 *	0,09	394 61,0%	32,5	30,4	0,99 ***	0,30	322 202,4%		
BFI - N ³	21,9	21,9	0,52	0,00	13 2,0%	21,4	21,8	0,68	0,06	56 35,1%		
BFI - O	33,5	33,1	0,87	0,07	271 42,0%	33,9	33,8	0,55	0,02	17 10,7%		
NSLCS ³	9,9	9,9	0,43	0,01	15 4,8%	8,8	9,3	0,79	0,15	31 46,2%		
GS	3,4	3,3	0,83	0,08	14 4,8%	3,5	3,4	0,72	0,12	4 5,9%		
SDQ - Em ³	3,8	3,9	0,64	0,03	22 6,8%	3,7	3,8	0,64	0,06	12 12,6%		
SDQ - Cond ³	3,2	3,0	0,13	0,09	62 18,9%	2,8	2,6	0,24	0,12	20 22,2%		
SDQ - Col ³	2,9	2,7	0,10	0,10	67 20,3%	2,3	2,5	0,76	0,11	19 20,4%		
SDQ - Hip ³	4,7	4,7	0,58	0,02	12 3,7%	4,4	4,3	0,48	0,01	2 2,0%		
SDQ - PS	7,3	7,2	0,64	0,03	20 6,1%	7,7	7,5	0,79	0,13	22 23,4%		
RSES	18,7	18,1	0,93 *	0,12	174 54,0%	19,6	18,5	0,94 *	0,26	101 110,1%		
CSE	3,4	3,4	0,80	0,07	12 3,8%	3,5	3,4	0,89	0,22	10 11,2%		
SEQC - Ac	25,0	24,7	0,71	0,04	67 21,6%	26,1	24,9	0,93 *	0,29	77 114,5%		
SEQC - Sc	24,3	24,2	0,58	0,02	24 8,0%	24,7	24,6	0,55	0,03	7 10,7%		
SEQC - Em	22,6	22,7	0,45	0,01	17 5,4%	22,9	21,8	0,91 *	0,23	72 108,0%		
DS	7,5	7,6	0,46	0,01	7 2,4%	8,0	7,6	0,76	0,14	56 35,2%		

¹: P-valor de teste unicaudal no sentido da desejabilidade social.

²: A dimensão extroversão é a menos consensual em relação ao sentido da desejabilidade social, principalmente em relação à monotonicidade do efeito. Considerou-se para a análise, maior extroversão relacionada a maior desejabilidade social.

³: Constructo em que há correlação negativa com desejabilidade social.

*, **, ***: estatisticamente significativa a 10%, 5% e 1% respectivamente.

★, ★★, ★★★: estatisticamente significativa a 10%, 5% e 1% respectivamente, mas no sentido oposto da desejabilidade social.

professores receberiam os resultados individuais, mas não notificados da anulação devido à manipulação. Já em relação à amostra restrita, houve redução da consistência interna do instrumento em relação ao grupo de controle com 90% de confiança em apenas 2 escalas.

Por fim, as figuras 10 e 11 apresentam as matrizes de correlação dos instrumentos para as duas amostras, completa e restrita, respectivamente. Em relação à mudança de sinal das correlações em relação ao esperado, houve 3 casos na amostra completa e 2 na amostra reduzida. Já em relação à desejabilidade social, são 10 casos onde houve redução absoluta da correlação dos itens com desejabilidade social. Por outro lado, na amostra de questionários completamente respondidos, o resultado foi contrário: 10 escalas apresentaram aumento da correlação com desejabilidade social em módulo em relação ao grupo de controle.

Esses resultados parecem sustentar a ideia de que a notificação também exerce algum efeito sobre o comportamento manipulador do contexto escolar de maneira generalizada e não somente o causado pelas abordagens do experimento. Outro fator que pode explicar

Figura 10 – Matriz de correlação dos constructos - Grupo de tratamento: professor receber respostas individuais e notificação - Amostra completa

	E	BFI					DS	SEQC			NSLCS	GS	SDQ					CSE	RSES
		A	C	N	O			Ac	Sc	Em			Em	Cond	Hip	Col	PS		
E																			
A	-0,12																		
BFI																			
C	-0,04	0,40																	
N	0,12	-0,46	-0,24																
O	0,20	0,25	0,28	-0,07															
DS	-0,17	0,43	0,42	-0,33	0,11														
Ac	0,01	0,44	0,67	-0,18	0,28	0,42													
SEQC																			
Sc	0,41	0,16	0,15	0,01	0,37	0,11	0,34												
Em	0,03	0,34	0,37	-0,36	0,28	0,35	0,38	0,30											
NSLCS	-0,06	-0,18	-0,28	0,14	-0,05	-0,33	-0,19	-0,09	-0,08										
GS	0,01	0,39	0,62	-0,27	0,21	0,47	0,56	0,19	0,32	-0,29									
Em	-0,13	-0,14	-0,04	0,35	0,03	-0,13													
Cond	0,20	-0,48	-0,35	0,46	-0,12	-0,41							0,27						
SDQ																			
Hip	0,35	-0,31	-0,57	0,33	-0,01	-0,38							0,17	0,43					
Col	-0,21	-0,18	-0,14	0,04	-0,01	-0,10							0,32	0,23	0,11				
PS	0,07	0,39	0,30	-0,03	0,24	0,31							0,11	-0,24	-0,19	-0,15			
CSE	0,13	0,31	0,40	-0,25	0,16	0,32							-0,42	-0,33	-0,37	-0,27	0,27		
RSES	0,09	0,16	0,23	-0,19	0,12	0,16							-0,42	-0,18	-0,27	-0,24	0,18	0,65	

Nota: Em destaque correlações com sinais contrários ao esperado pela literatura. BFI-E está marcado pois não há muito consenso sobre o sinal de suas correlações com os demais constructos.

Figura 11 – Matriz de correlação dos constructos - Grupo de tratamento: professor receber respostas individuais e notificação - Amostra restrita

	E	BFI					DS	SEQC			NSLCS	GS	SDQ					CSE	RSES
		A	C	N	O			Ac	Sc	Em			Em	Cond	Hip	Col	PS		
E																			
A	-0,19																		
BFI																			
C	-0,04	0,44																	
N	0,20	-0,53	-0,18																
O	0,23	0,24	0,36	-0,09															
DS	-0,24	0,49	0,44	-0,30	0,14														
Ac	-0,05	0,33	0,66	-0,15	0,25	0,48													
SEQC																			
Sc	0,44	0,10	0,19	0,08	0,37	0,16	0,33												
Em	-0,03	0,46	0,36	-0,52	0,26	0,46	0,31	0,28											
NSLCS	-0,05	-0,23	-0,39	0,04	-0,24	-0,38	-0,31	-0,17	-0,13										
GS	-0,10	0,50	0,81	-0,29	0,29	0,59	0,66	0,14	0,32	-0,40									
Em	-0,11	-0,13	-0,04	0,34	0,08	-0,07													
Cond	0,33	-0,43	-0,30	0,52	-0,12	-0,38							0,25						
SDQ																			
Hip	0,29	-0,40	-0,60	0,31	-0,26	-0,41							0,24	0,50					
Col	-0,24	-0,16	-0,05	0,13	0,00	-0,01							0,31	0,19	0,17				
PS	0,08	0,30	0,27	-0,08	0,40	0,27							0,14	-0,15	-0,29	-0,05			
CSE	0,10	0,36	0,43	-0,23	0,10	0,37							-0,47	-0,23	-0,45	-0,35	0,15		
RSES	0,05	0,18	0,27	-0,08	0,07	0,17							-0,41	-0,07	-0,35	-0,24	0,14	0,74	

Nota: Em destaque correlações com sinais contrários ao esperado pela literatura. BFI-E está marcado pois não há muito consenso sobre o sinal de suas correlações com os demais constructos.

esse viés negativo causado pela notificação associada ao fato de os professores receberem as respostas individuais é a possibilidade de os alunos avaliarem que seus professores

subestimam suas competências socioemocionais. Caso essa hipótese fosse válida, faria sentido os alunos reportarem menores níveis do que de fato possuem para que os professores não invalidassem seus questionários, já que estes receberiam e poderiam avaliar de acordo com suas percepções as respostas individuais dos alunos.

5 Conclusão

Esse trabalho procurou contribuir para a literatura de comportamento manipulador em mensurações de competências socioemocionais especificamente no ambiente escolar e em situações *low-stake* causadas por abordagens que simulam eventos prováveis nesse contexto.

Os resultados mostram que ambas as abordagens investigadas intensificaram o comportamento manipulador nos alunos, gerando viés adicional nessas medidas. Além disso, esse viés só é parcialmente diluído na agregação dos itens em seus respectivos constructos. A maior parte dos instrumentos que apresentaram evidências de viés nos itens também apresentou diferenças significativas nos constructos agregados.

Dessa forma, as respostas para as duas primeiras perguntas que nortearam o trabalho são sim, tanto a disseminação de informação quanto o fato de os professores receberem as respostas individuais tornam os alunos mais motivados a inflarem suas respostas no sentido da desejabilidade social e fatores como baixa idade e escolaridade dos alunos não representam barreiras para identificar e reportar as respostas mais desejáveis socialmente daqueles instrumentos, de forma que são capazes de manipular as respostas quando desejam.

Já sobre o efeito do aumento do comportamento manipulador nas propriedades dos instrumentos, não se tem nenhuma conclusão muito forte sobre as duas propriedades investigadas: consistência interna e correlação entre constructos. Em relação à primeira, não é possível concluir que sob as tratamentos os instrumentos como um todo apresentaram menor consistência interna. Já em relação à correlação dos constructos, o máximo que se pode concluir é que nos tratamentos relacionados à disseminação da informação, houve um aumento da correlação da maioria dos constructos com a escala de desejabilidade social nos grupos sem e com notificação, mas esse número de constructos foi menor no último. Essa evidência sugere que de fato o comportamento manipulador foi intensificado com a palestra, mas reduzido com a inclusão da notificação.

Outro ponto importante é que aspectos relacionados à estrutura de cada instrumento

podem ter algum efeito para facilitar ou dificultar esse comportamento, isso porque esse comportamento não se mostrou presente de maneira igual em todos eles. Os instrumentos GS e CSE, por exemplo, não apresentaram viés em nenhuma das análises e amostras. Por outro lado, os instrumentos menos robustos a essas intervenções foram o SEQC, SDQ e BFI. Em relação a esse último, vale notar que suas dimensões apresentaram viés de manipulação somente no tratamento relacionado à palestra, provavelmente devido ao fato de que a palestra aumentou o conhecimento do que estava sendo medido uma vez que a palestra foi estruturada com base na teoria dos 5 fatores e isso potencializou a capacidade dos indivíduos em manipularem esse instrumento, em linha com o trabalho de [Cunningham, Wong e Barbee \(1994\)](#).

Seguindo essa comparação entre escalas, vale discutir um ponto adicional: escalas geralmente apontadas como muito correlacionadas na literatura apresentaram, nesse trabalho, níveis bastante diferentes de suscetibilidade à manipulação. É o caso por exemplo da escala de garra (GS) e a dimensão conscienciosidade (BFI-C)¹. Enquanto a primeira não apresentou nenhum indício de viés de manipulação nas intervenções, a última apresentou viés com a palestra. Essa diferença pode ser causada por estruturas de cada um desses instrumentos, mas também à forma com que o questionário aplicado no experimento foi construído. Isso porque o BFI fazia parte do primeiro bloco de questões, enquanto a GS estava mais para o fim do mesmo. Esse formato pode ter gerado padrões de respostas diferentes em cada um desses instrumentos. Isolar esses efeitos seria um bom exercício para futuros trabalhos.

A evidência de viés adicional causado pela disseminação da informação é em certo sentido mais preocupante do que aquele causado pelo recebimento das respostas pelos professores. Isso porque no último caso trata-se de uma situação controlada, ou seja, pode-se limitar o uso das informações para análise individual dos alunos pelos professores. Já em relação ao primeiro efeito, é esperado e provável que se dê cada vez mais importância a essas competências, gerando e disseminando informações sobre elas. Assim, não podendo

¹ No grupo de controle essas escalas correlacionam-se em torno de 56%, mas há estudos que mostram correlação ainda mais alta.

controlar a motivação espontânea que será criada, deve-se procurar formas de desestimular o comportamento manipulador.

Nesse sentido, cabe responder à terceira pergunta sobre a eficácia da notificação. E a resposta para isso com base nos resultados é também que sim, a notificação de fato reduziu o comportamento manipulador dos alunos causado pelas abordagens investigadas e também o possível viés causado pelo ambiente escolar, de forma que parece ser uma alternativa bastante factível para se obter medidas mais fieis das reais competências dos indivíduos, mostrando adequação ao modelo de [McFarland e Ryan \(2000\)](#) e aos trabalhos de [Doll \(1971\)](#) e [Schrader e Osburn \(1977\)](#).

Sobre o efeito da notificação e consequente sugestão da existência de viés de manipulação no contexto escolar sem nenhum choque adicional, vale notar que uma dimensão mostrou-se consistentemente viesada no sentido contrário ao da desejabilidade social na presença da notificação em ambas as abordagens e amostras: a dimensão "conscienciosidade" do BFI. Esse fato é interessante já que essa dimensão mede traços como organização, disciplina e outros traços mais relacionados com a educação formal que se desenvolve de maneira geral nas escolas. Assim, não seria surpreendente o fato dos alunos se sentirem motivados a inflar suas respostas quando são questionados sobre elas no ambiente escolar.

Outra sugestão geralmente apontada para detectar e controlar o viés de manipulação é o uso de escalas de desejabilidade social. Entretanto, essa não parece ser uma boa alternativa uma vez que os resultados mostraram viés de manipulação nos itens e constructos dos instrumentos mesmo quando a escala de desejabilidade social não apresentou diferença significativa entre os grupos, o que de certa forma mostra a limitação de seu uso para inferir sobre comportamento manipulador. Resultado esse alinhado ao de [Griffith e Peterson \(2008\)](#).

Dessa forma, as recomendações de política em relação à mensuração dessa matriz de competências no ambiente escolar são as seguintes: em primeiro lugar, que se dê atenção ao fato de que nesse contexto pode existir uma rede de incentivo própria que age de forma a tornar o ambiente em algum grau já motivado para o comportamento manipulador, não

precisando de nenhum choque adicional para tornar as respostas viesadas. Essa hipótese poderia ter sido testada no trabalho caso o desenho do experimento incluísse um grupo de tratamento no qual as notificações estivessem presentes nos questionários sem nenhuma abordagem adicional. Entretanto, por não ter sido considerada essa hipótese ex-ante, isso não foi feito. Por outro lado, o fato de as notificações causarem um padrão de resposta sistematicamente viesado no sentido contrário da desejabilidade social em relação ao grupo de controle sugere que essa hipótese seja válida.

Como consequência da evidência apontada acima, sugere-se a inclusão de mecanismos para barrar o comportamento manipulador nos questionários em qualquer cenário. Isso porque deve desestimular esse comportamento nas condições atuais de conhecimento sobre essas competências e também quando o conjunto de informação dos indivíduos sobre esse tema for expandido.

Outra recomendação é que se evite a utilização dessas medidas para avaliação individual dos alunos principalmente se não for incluído nenhum mecanismo para desestimular o comportamento manipulador. Os resultados do trabalho mostram algo bem intuitivo: que tal uso gera motivações fortes suficiente para que os alunos respondam os itens inflando suas respostas de forma sistemática, o que tornaria as medidas fracas instrumentos de medição das reais competências. Nesse sentido, um exercício adicional que pode ser feito futuramente é investigar o efeito dessa abordagem no ordenamento dos alunos em relação às competências de interesse.

Como limitação do trabalho destaca-se o fato da ausência de vinhetas âncoras (King et al. (2004)) nos questionários. Tal ferramenta poderia fornecer respostas mais fiéis dos alunos ao eliminar estilos de respostas individuais que viesam essas medidas e inclusive são mais presentes na faixa etária aqui estudada (Soto et al. (2008); Primi et al. (2016b)). Recomenda-se, portanto, que futuros trabalhos utilizem essa ferramenta.

Referências

- ALMLUND, M. et al. *Personality psychology and economics*. [S.l.], 2011. 11
- BALLARD, R. Short forms of the marlowe-crowne social desirability scale. *Psychological reports*, Ammons Scientific, v. 71, n. 3f, p. 1155–1160, 1992. 23
- BANDURA, A. *Self-efficacy*. [S.l.]: Wiley Online Library, 1994. 20
- BARBARANELLI, C. et al. A questionnaire for measuring the big five in late childhood. *Personality and Individual Differences*, Elsevier, v. 34, n. 4, p. 645–664, 2003. 17
- BENET-MARTINEZ, V.; JOHN, O. P. Los cinco grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the big five in spanish and english. *Journal of personality and social psychology*, American Psychological Association, v. 75, n. 3, p. 729, 1998. 43
- BRANDEN, N. The psychology of self-esteem: a new concept of man's psychology. Bantam Books, 1969. 22
- CORTINA, J. M. What is coefficient alpha? an examination of theory and applications. *Journal of applied psychology*, American Psychological Association, v. 78, n. 1, p. 98, 1993. 34
- CRONBACH, L. J. Coefficient alpha and the internal structure of tests. *psychometrika*, Springer, v. 16, n. 3, p. 297–334, 1951. 33
- CROWNE, D. P.; MARLOWE, D. A new scale of social desirability independent of psychopathology. *Journal of consulting psychology*, American Psychological Association, v. 24, n. 4, p. 349, 1960. 23
- CUNNINGHAM, M. R.; WONG, D. T.; BARBEE, A. P. Self-presentation dynamics on overt integrity tests: Experimental studies of the reid report. *Journal of Applied Psychology*, American Psychological Association, v. 79, n. 5, p. 643, 1994. 13, 66
- DOLL, R. E. Item susceptibility to attempted faking as related to item characteristic and adopted fake set. *The Journal of Psychology*, Taylor & Francis, v. 77, n. 1, p. 9–16, 1971. 13, 28, 67
- DUCKWORTH, A. L. et al. Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology*, American Psychological Association, v. 92, n. 6, p. 1087, 2007. 23
- DUCKWORTH, A. L.; QUINN, P. D. Development and validation of the short grit scale (grit-s). *Journal of personality assessment*, Taylor & Francis, v. 91, n. 2, p. 166–174, 2009. 23, 43, 44
- DUCKWORTH, A. L.; YEAGER, D. S. Measurement matters assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, SAGE Publications, v. 44, n. 4, p. 237–251, 2015. 11
- DUNCAN, G. J.; MAGNUSON, K. The nature and impact of early achievement skills, attention skills, and behavior problems. *Whither opportunity*, p. 47–70, 2011. 11

- ELLINGSON, J. E. People fake only when they need to fake. In: ZIEGLER, M.; MACCANN, C.; ROBERTS, R. (Ed.). *New perspectives on faking in personality assessment*. [S.l.]: Oxford University Press, 2011. cap. 2, p. 19–33. [25](#)
- FREDERICK, S. Cognitive reflection and decision making. *The Journal of Economic Perspectives*, JSTOR, v. 19, n. 4, p. 25–42, 2005. [24](#)
- GOODMAN, R. The strengths and difficulties questionnaire: a research note. *Journal of child psychology and psychiatry*, Wiley Online Library, v. 38, n. 5, p. 581–586, 1997. [20](#), [44](#)
- GRIFFITH, R. L.; PETERSON, M. H. The failure of social desirability measures to capture applicant faking behavior. *Industrial and Organizational Psychology*, Cambridge Univ Press, v. 1, n. 03, p. 308–311, 2008. [67](#)
- GUTTMAN, L. Image theory for the structure of quantitative variates. *Psychometrika*, Springer, v. 18, n. 4, p. 277–296, 1953. [33](#)
- HAMPSON, S. E. et al. Mechanisms by which childhood personality traits influence adult health status: educational attainment and healthy behaviors. *Health psychology*, American Psychological Association, v. 26, n. 1, p. 121, 2007. [11](#)
- HAYES, B. E. *Measuring customer satisfaction: Survey design, use, and statistical analysis methods*. [S.l.]: ASQ Quality Press, 1998. [33](#)
- HEGGESTAD, E. D. A conceptual representation of faking. In: ZIEGLER, M.; MACCANN, C.; ROBERTS, R. (Ed.). *New perspectives on faking in personality assessment*. [S.l.]: Oxford University Press, 2011. cap. 6, p. 87–101. [12](#), [13](#)
- HEGGESTAD, E. D.; GEORGE, E.; REEVE, C. L. Transient error in personality scores: Considering honest and faked responses. *Personality and individual differences*, Elsevier, v. 40, n. 6, p. 1201–1211, 2006. [12](#)
- JOHN, O. P. et al. The "little five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child development*, JSTOR, p. 160–178, 1994. [11](#)
- JOHN, O. P.; DONAHUE, E. M.; KENTLE, R. L. *The big five inventory—versions 4a and 54*. [S.l.]: Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research, 1991. [19](#), [43](#), [44](#)
- JUDGE, T. A.; BONO, J. E. Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of applied Psychology*, American Psychological Association, v. 86, n. 1, p. 80, 2001. [22](#)
- JUDGE, T. A. et al. The core self-evaluations scale: Development of a measure. *Personnel psychology*, Wiley Online Library, v. 56, n. 2, p. 303–331, 2003. [22](#), [43](#), [44](#)
- KERN, M. L.; FRIEDMAN, H. S. Do conscientious individuals live longer? a quantitative review. *Health Psychology*, American Psychological Association, v. 27, n. 5, p. 505, 2008. [11](#)
- KING, G. et al. Enhancing the validity and cross-cultural comparability of measurement in survey research. *American political science review*, Cambridge Univ Press, v. 98, n. 01, p. 191–207, 2004. [68](#)

LAUTENSCHLAGER, G. J. Accuracy and faking of background data. CPP Books, 1994. [13](#)

LOUNSBURY, J. W. et al. An investigation of personality traits in relation to adolescent school absenteeism. *Journal of Youth and Adolescence*, Springer, v. 33, n. 5, p. 457–466, 2004. [11](#)

MARTIN, B. A.; BOWEN, C.-C.; HUNT, S. T. How effective are people at faking on personality questionnaires? *Personality and Individual Differences*, Elsevier, v. 32, n. 2, p. 247–256, 2002. [14](#)

MCFARLAND, L. A.; RYAN, A. M. Variance in faking across noncognitive measures. *Journal of Applied Psychology*, American Psychological Association, v. 85, n. 5, p. 812, 2000. [12](#), [28](#), [67](#)

MCFARLAND, S. G.; SPARKS, C. M. Age, education, and the internal consistency of personality scales. *Journal of Personality and Social Psychology*, American Psychological Association, v. 49, n. 6, p. 1692, 1985. [42](#)

MCKINLEY, J. C.; HATHAWAY, S. R. The identification and measurement of the psychoneuroses in medical practice: the minnesota multiphasic personality inventory. *Journal of the American Medical Association*, American Medical Association, v. 122, n. 3, p. 161–167, 1943. [12](#)

MURIS, P. A brief questionnaire for measuring self-efficacy in youths. *Journal of Psychopathology and Behavioral Assessment*, Springer, v. 23, n. 3, p. 145–149, 2001. [20](#), [43](#), [44](#)

MURIS, P.; MEESTERS, C.; BERG, F. van den. The strengths and difficulties questionnaire (sdq). *European child & adolescent psychiatry*, Springer, v. 12, n. 1, p. 1–8, 2003. [43](#)

NOWICKI, S.; STRICKLAND, B. R. A locus of control scale for children. *Journal of consulting and clinical psychology*, American Psychological Association, v. 40, n. 1, p. 148, 1973. [21](#), [43](#), [44](#)

NUNNALLY, J. C.; BERNSTEIN, I. H.; BERGE, J. M. t. *Psychometric theory*. [S.l.]: JSTOR, 1967. v. 226. [42](#)

PASTORE, M.; LOMBARDI, L. The impact of faking on cronbach's alpha for dichotomous and ordered rating scores. *Quality & Quantity*, Springer, v. 48, n. 3, p. 1191–1211, 2014. [34](#)

PRIMI, R. et al. Development of an inventory assessing social and emotional skills in brazilian youth. *European Journal of Psychological Assessment*, Hogrefe Publishing, 2016. [17](#)

PRIMI, R. et al. Anchoring vignettes. *European Journal of Psychological Assessment*, Hogrefe Publishing, 2016. [68](#)

RAVEN, J.; RAVEN, J. C.; COURT, J. H. *Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 4 Advanced Progressive Matrices sets I and II*. 1998. ed. [S.l.]: Oxford Psychologists, Oxford, 1998. [24](#)

ROBERTS, B. W. et al. The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, SAGE Publications, v. 2, n. 4, p. 313–345, 2007. 11

ROBERTS, B. W.; WALTON, K. E.; VIECHTBAUER, W. Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological bulletin*, American Psychological Association, v. 132, n. 1, p. 1, 2006. 11

ROSENBERG, M. *Conceiving the self*. [S.l.]: RE Krieger, 1986. 22

ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied*, American Psychological Association, v. 80, n. 1, p. 1, 1966. 20, 21

SANTOS, D. *A importância socioeconômica das características de Personalidade*. 2014. Disponível em: <<http://educacaosec21.org.br/wp-content/uploads/2013/08/A-import%C3%A2ncia-socioecon%C3%B4mica-das-caracter%C3%ADsticas-de-Personalidade.pdf>>. 11, 19

SCAGLIUSI, F. B. et al. Tradução da escala de desejo de aceitação social de marlowe & crowne para a língua portuguesa. *Revista de Psiquiatria Clínica*, v. 31, n. 6, p. 272–278, 2004. 23

SCHRADER, A. D.; OSBURN, H. Biodata faking: Effects of induced subtlety and position specificity. *Personnel Psychology*, Wiley Online Library, v. 30, n. 3, p. 395–404, 1977. 13, 28, 67

SINCLAIR, S. J. et al. Psychometric properties of the rosenberg self-esteem scale: Overall and across demographic groups living within the united states. *Evaluation & the Health Professions*, SAGE Publications, v. 33, n. 1, p. 56–80, 2010. 43, 44

SMITH, D. B.; ELLINGSON, J. E. Substance versus style: a new look at social desirability in motivating contexts. *Journal of Applied Psychology*, American Psychological Association, v. 87, n. 2, p. 211, 2002. 14

SOTO, C. J. et al. The developmental psychometrics of big five self-reports: acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of personality and social psychology*, American Psychological Association, v. 94, n. 4, p. 718, 2008. 68

STÖRMER, S.; FAHR, R. Individual determinants of work attendance: Evidence on the role of personality. *Applied Economics*, Taylor & Francis, v. 45, n. 19, p. 2863–2875, 2013. 11

VAZSONYI, A. T. et al. An empirical test of a general theory of crime: A four-nation comparative study of self-control and the prediction of deviance. *Journal of Research in Crime and Delinquency*, Sage Publications, v. 38, n. 2, p. 91–131, 2001. 11

VISWESVARAN, C.; ONES, D. S. Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and psychological measurement*, Sage Publications, v. 59, n. 2, p. 197–210, 1999. 14

WELSH, M.; BURNS, N.; DELFABBRO, P. The cognitive reflection test: how much more than numerical ability. In: COGNITIVE SCIENCE SOCIETY AUSTIN, TX. *Proceedings of the 35th Annual Conference of the Cognitive Science Society*. [S.l.], 2013. p. 1587–1592. [24](#)

ZIEGLER, M. *Situational demand and its impact on construct and criterion validity of a personality questionnaire: State and trait, a couple you just can't study separately!* Tese (Doutorado) — lmu, 2007. [13](#), [14](#), [25](#)

ZIEGLER, M.; MACCANN, C.; ROBERTS, R. *New perspectives on faking in personality assessment*. [S.l.]: Oxford University Press, 2011. [11](#), [30](#), [34](#)

Apêndice A - Detalhes do experimento

Uma vez estabelecida a participação das 10 escolas no experimento, procurou-se agendar o experimento em todas as escolas no menor intervalo de tempo entre elas para evitar problemas de transbordamento. Isso é, evitar a possibilidade que os agentes das escolas, principalmente os alunos que já tivessem participado do experimento falassem sobre ele, especialmente sobre as abordagens dos grupos de tratamento para alunos de outras escolas que ainda iriam participar do experimento em outra data. Se essas informações fossem passadas para um aluno aleatoriamente alocado no grupo de controle, então criaria-se problema de *partial compliance* uma vez que na verdade esse aluno recebeu algum dos tratamentos por uma canal diferente, contaminando o grupo de controle. Tendo isso em vista, o experimento ocorreu nas 10 escolas no intervalo de uma semana, sendo que escolas mais próximas foram agendadas para o mesmo dia para facilitar a logística e reduzir o risco citado anteriormente.

O objetivo do experimento foi informado para todos os diretores das escolas participantes, porém todos foram orientados a não passar essa informação para nenhum outro funcionário da escola e alunos, novamente para não criar ruídos na mensuração. O que podiam fazer, e em muitos casos fizeram, foi informar os professores que em determinado horário e data uma equipe passaria nas salas de aula para medir o nível de competências socioemocionais dos alunos.

O experimento ocorreu na semana do dia 14 de Março de 2016, uma semana depois da reunião com os diretores. Nesse intervalo, todos os diretores enviaram as listas com nomes de todos os alunos por turma. Com isso, realizou-se a aleatorização dos alunos entre grupos. De maneira geral, os alunos de cada turma eram divididos de forma aleatória em 3 grupos, controle, tratamento palestra e tratamento professor, sendo os dois últimos com o dobro de alunos do primeiro para que, depois da fragmentação em relação à notificação, todos os grupos tivessem o mesmo número de alunos².

² Em uma escola não foi permitido realizar o deslocamento entre alunos, apesar da concordância por parte da diretora na reunião em que foi combinado os termos do experimento.

A equipe era formada por 19 aplicadores mais 2 supervisores, todos remunerados e estudantes de graduação ou pós graduação, em grande parte do curso de economia da USP. A equipe recebeu treinamento prévio em algumas reuniões com o autor da pesquisa, em que foi apresentada ao conteúdo da mesma, ou seja, o que são as competências socioemocionais, evidências científicas das suas relações com aspectos futuros da vida dos indivíduos, salientando a importância de desenvolvê-las no ambiente escolar, exemplos de ações nesse sentido, esforços realizados para se mapear essa matriz de competências e a importância de se fazer isso de forma segura e correta, mas principalmente o objetivo da pesquisa, de forma que ficasse claro que a abordagem que eles teriam na hora de dar as instruções aos alunos era de profunda importância uma vez que o efeito investigado dependeria disso. A estratégia foi dar o máximo de informação à equipe para que soubessem o que era permitido falar e o que era proibido para não gerar nenhum problema adicional de incompatibilidade entre os grupos de controle e tratamento a não ser pelas informações adicionais que identificam os grupos de tratamento.

Ademais, foram orientados sobre aspectos gerais de aplicação dos questionários, muitos deles colhidos durante a etapa piloto da pesquisa. Além de informações sobre forma de preenchimento dos questionários, tempo para tal, erros comuns e postura desejada frente a situações problemáticas, também foram orientados a fazerem uma espécie de glossário no quadro negro com termos identificados como de difícil compreensão por parte dos alunos para que agilizasse o processo de resolução de dúvidas dos alunos e não atrapalhasse no controle da sala de aula. Os aplicadores alocados no grupo de tratamento da palestra receberam adicionalmente um quadro resumo contendo as evidências científicas sobre dimensões do Big Five e aspectos da vida futura estimulados por elas para ajudar na criação da palestra. Foram orientados a fazerem essa palestra para os alunos após as instruções gerais de preenchimento do questionário e antes da entrega deles aos alunos em duração de 5 minutos em média. Nota-se que o objetivo não foi de padronizar a palestra através de um modelo feito por uma única pessoa, mas preferiu-se certa heterogeneidade em como a palestra podia ser transmitida, tendo como base obviamente um núcleo de informações comum composto por evidências cientificamente corroboradas, para que cada

aplicador conseguisse passar essas informações de forma mais natural e menos engessada para atrair a atenção da classe. Além disso, sendo a palestra uma simulação da disseminação da informação por diversos agentes, não faz sentido acreditar que essa informação é/seria transmitida de forma padronizada. Já os aplicadores alocados no grupo de tratamento do professor foram orientados a informarem, depois das instruções gerais de preenchimento do questionário, que os professores receberiam aqueles questionários individuais para análise interna na escola e só então seriam levados para os pesquisadores.

Os aplicadores eram, portanto, responsáveis pela aplicação do questionário, enquanto os 2 supervisores eram responsáveis por circular entre as salas para garantir que as informações passadas estavam dentro do programado, além de ajudar em situações problemáticas relacionadas ao comportamento dos alunos, quando uma pessoa apenas não era suficiente para controlar.

Uma vez na escola, todas as turmas preenchiam o questionário simultaneamente, novamente para barrar o possível transbordamento de informações entre alunos. Para tanto, os aplicadores eram levados para suas turmas e com os alunos em classe informavam que seria feita uma pesquisa que levaria em torno de 1h e que fazia parte de um trabalho realizado pela universidade de São Paulo nas escolas da rede para checar qual o nível das competências socioemocionais (que seriam explicadas posteriormente a eles) nessas escolas. Em seguida, os aplicadores chamavam os alunos que eram designados para outros grupos para trocarem de sala com base na aleatorização feita previamente. Esse deslocamento de alunos entre salas era ajudado pelos supervisores, assim como funcionários da escola e durou em torno de 5 minutos em média.

Quando os aplicadores recebiam os alunos das outras turmas que pertenciam àquele grupo a ele designado, começavam as instruções gerais e as instruções específicas daquele grupo. Vale notar que em termos de tipos de aplicadores e tipos de salas, havia apenas 3 grupos, controle, tratamento palestra e tratamento professor, enquanto na pesquisa são 5 grupos diferentes. Isso porque nos grupos de tratamento houve uma fragmentação em relação à presença da notificação ou não. Isso se dava apenas na distribuição dos

questionários, com metade dos alunos dos grupos de tratamento recebendo questionários com a notificação impressa e metade sem ela. Esse desenho do experimento garante que o efeito da notificação seja identificado de maneira limpa do tratamento da palestra ou do professor. Caso esses alunos que receberam a notificação estivessem em salas separadas daqueles que também receberam um dos tratamentos, mas sem a notificação, não seria possível saber se o efeito observado tem a ver com a notificação em si ou é um resultado da forma ou intensidade que a palestra ou a informação sobre o professor foi passada por diferentes aplicadores. A contrapartida é que como dois tratamentos estão sendo investigados na mesma turma, poderia ser o caso de transbordamento entre tratamentos se algum aluno que recebeu a notificação falasse em voz alta sobre ela, afetando alunos que estão alocados no tratamento sem notificação. Contudo, essa ameaça é facilmente identificável pelo aplicador e não foi notificado nenhum caso desse.

Apêndice B - Distribuição dos alunos entre grupos e questionários por escola

Tabela 19 – Distribuição de alunos por escola, grupo e amostra

Escola	Controle	Palestra	Palestra + Notificação	Professor	Professor + Notificação	Total
AM	103 [39]	87 [25]	85 [38]	97 [37]	94 [31]	466 [170]
CM	93 [12]	58 [16]	56 [17]	71 [14]	69 [17]	347 [76]
LG	65 [13]	75 [19]	79 [24]	45 [15]	90 [28]	354 [99]
MR	68 [14]	34 [3]	32 [9]	50 [13]	53 [13]	237 [52]
OB	43 [13]	46 [16]	45 [9]	28 [8]	22 [6]	184 [52]
OC	52 [21]	47 [12]	39 [5]	32 [6]	22 [3]	192 [47]
OA	64 [7]	39 [3]	57 [14]	49 [5]	67 [15]	276 [44]
PI	30 [2]	61 [19]	66 [20]	41 [10]	30 [7]	228 [58]
RR	61 [18]	56 [23]	56 [25]	55 [21]	55 [20]	283 [107]
TA	66 [20]	47 [9]	51 [11]	58 [14]	64 [22]	286 [76]
Total	645 [159]	550 [145]	566 [172]	526 [143]	566 [162]	2853 [781]

[.]: amostra restrita de questionários inteiramente respondidos.

Tabela 20 – Distribuição de alunos por escola, questionário e amostra

Escola	A	B	Total
AM	228 [81]	238 [89]	466 [170]
CM	174 [47]	173 [29]	347 [76]
LG	140 [35]	214 [64]	354 [99]
MR	114 [29]	123 [23]	237 [52]
OB	88 [28]	96 [24]	184 [52]
OC	102 [24]	90 [23]	192 [47]
OA	148 [21]	128 [23]	276 [44]
PI	124 [28]	104 [30]	228 [58]
RR	139 [50]	144 [57]	283 [107]
TA	142 [34]	144 [42]	286 [76]
Total	1399 [377]	1454 [404]	2853 [781]

[.]: amostra restrita de questionários inteiramente respondidos.

Apêndice C - Testes de balanceamento

Tabela 21 – Teste de balanceamento - Tratamento Palestra

Variável	Gabarito A'					Gabarito B'					Gabarito AB				
	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença
Gênero	0,43	287	0,47	246	-0,04	0,43	317	0,53	268	-0,10 **	0,43	604	0,50	514	-0,07 **
Raça	0,41	290	0,46	248	-0,05	0,40	311	0,41	264	-0,01	0,40	601	0,43	512	-0,03
Idade	12,31	293	12,26	251	0,04	12,24	320	12,17	273	0,07	12,27	613	12,22	524	0,06
Ano escolar	7,39	274	7,29	233	0,10	7,35	304	7,25	259	0,09	7,37	578	7,27	492	0,10
CRT (%)	0,08	283	0,08	251	0,00	0,08	311	0,06	271	0,02 *	0,08	594	0,07	522	0,01
MR (%)	0,43	278	0,43	245	-0,01	0,44	307	0,42	269	0,02	0,43	585	0,43	514	0,00
Mora mãe	0,90	279	0,87	246	0,03	0,90	305	0,91	259	-0,01	0,90	584	0,89	505	0,01
Mora pai	0,71	276	0,70	240	0,02	0,75	300	0,74	253	0,01	0,73	576	0,72	493	0,01
Mãe ler e escrever	0,99	240	0,99	207	0,00	0,98	265	0,96	230	0,02	0,98	505	0,97	437	0,01
Pai ler e escrever	0,97	183	0,92	158	0,05 **	0,94	216	0,94	176	0,00	0,95	399	0,93	334	0,02
Irmãos novos	0,53	284	0,46	245	0,07	0,48	311	0,48	261	0,00	0,50	595	0,47	506	0,04
Irmãos velhos	0,56	282	0,47	247	0,09 **	0,53	310	0,50	262	0,03	0,54	592	0,49	509	0,06 *
Pessoas na casa	4,04	281	3,84	244	0,20 *	4,13	312	3,87	258	0,26 **	4,09	593	3,86	502	0,23 ***
Empregada	0,11	276	0,13	238	-0,02	0,15	308	0,15	259	0,00	0,13	584	0,14	497	-0,01
Carros	0,96	287	0,97	244	-0,01	1,06	308	1,05	263	0,01	1,01	595	1,01	507	0,00
Banheiros	1,45	285	1,46	245	-0,01	1,50	310	1,42	263	0,08	1,48	595	1,44	508	0,04
Quartos	2,09	281	2,05	244	0,04	2,07	309	2,11	262	-0,04	2,08	590	2,08	506	0,00
Computador	1,37	282	1,27	244	0,10	1,44	309	1,44	263	0,00	1,41	591	1,36	507	0,05
Asfalto	0,83	241	0,80	219	0,03	0,79	274	0,84	239	-0,05	0,81	515	0,82	458	-0,01
Energia	0,96	276	0,97	239	-0,01	0,96	294	0,96	258	0,00	0,96	570	0,97	497	-0,01

Tabela 22 – Teste de balanceamento - Tratamento Palestra + Notificação

Variável	Gabarito A'					Gabarito B'					Gabarito AB				
	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença
Gênero	0,43	287	0,49	259	-0,06	0,43	317	0,42	267	0,01	0,43	604	0,46	526	-0,02
Raça	0,41	290	0,41	258	0,00	0,40	311	0,43	269	-0,03	0,40	601	0,42	527	-0,02
Idade	12,31	293	12,29	261	0,02	12,24	320	12,24	272	0,00	12,27	613	12,26	533	0,01
Ano escolar	7,39	274	7,34	247	0,05	7,35	304	7,41	258	-0,07	7,37	578	7,38	505	-0,01
CRT (%)	0,08	283	0,06	261	0,02	0,08	311	0,06	266	0,02	0,08	594	0,06	527	0,02
MR (%)	0,43	278	0,40	265	0,03	0,44	307	0,44	266	-0,01	0,43	585	0,42	531	0,01
Mora mãe	0,90	279	0,92	257	-0,02	0,90	305	0,93	263	-0,03	0,90	584	0,93	520	-0,03
Mora pai	0,71	276	0,67	257	0,05	0,75	300	0,76	259	-0,02	0,73	576	0,72	516	0,02
Mãe ler e escrever	0,99	240	0,97	231	0,02	0,98	265	0,98	234	0,00	0,98	505	0,97	465	0,01
Pai ler e escrever	0,97	183	0,98	166	0,00	0,94	216	0,94	187	0,00	0,95	399	0,96	353	0,00
Irmãos novos	0,53	284	0,52	259	0,01	0,48	311	0,47	262	0,01	0,50	595	0,50	521	0,01
Irmãos velhos	0,56	282	0,58	260	-0,02	0,53	310	0,52	266	0,01	0,54	592	0,55	526	-0,01
Pessoas na casa	4,04	281	3,97	258	0,07	4,13	312	3,99	264	0,14	4,09	593	3,98	522	0,11
Empregada	0,11	276	0,12	255	-0,01	0,15	308	0,15	265	-0,01	0,13	584	0,14	520	-0,01
Carros	0,96	287	0,93	258	0,03	1,06	308	0,95	261	0,11	1,01	595	0,94	519	0,07
Banheiros	1,45	285	1,40	252	0,04	1,50	310	1,46	261	0,05	1,48	595	1,43	513	0,05
Quartos	2,09	281	2,02	257	0,07	2,07	309	2,03	260	0,05	2,08	590	2,02	517	0,06
Computador	1,37	282	1,45	254	-0,08	1,44	309	1,38	263	0,07	1,41	591	1,41	517	-0,01
Asfalto	0,83	241	0,82	235	0,01	0,79	274	0,80	236	-0,01	0,81	515	0,81	471	0,00
Energia	0,96	276	0,97	247	-0,02	0,96	294	0,98	253	-0,02	0,96	570	0,98	500	-0,02

**

Tabela 23 – Teste de balanceamento - Tratamento Professor

Variável	Gabarito A'					Gabarito B'					Gabarito AB						
	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença		
Gênero	0,43	287	0,48	245	-0,05	0,43	317	0,46	252	-0,03		0,43	604	0,47	497	-0,04	
Raça	0,41	290	0,44	240	-0,03	0,40	311	0,39	245	0,01		0,40	601	0,41	485	-0,01	
Idade	12,31	293	12,34	247	-0,03	12,24	320	12,41	257	-0,16	*	12,27	613	12,38	504	-0,10	
Ano escolar	7,39	274	7,46	240	-0,07	7,35	304	7,47	241	-0,12		7,37	578	7,47	481	-0,10	
CRT (%)	0,08	283	0,08	237	0,00	0,08	311	0,07	246	0,01		0,08	594	0,07	483	0,01	
MR (%)	0,43	278	0,41	236	0,02	0,44	307	0,40	251	0,04		0,43	585	0,40	487	0,03	
Mora mãe	0,90	279	0,90	240	0,00	0,90	305	0,89	244	0,01		0,90	584	0,90	484	0,00	
Mora pai	0,71	276	0,73	230	-0,01	0,75	300	0,66	242	0,09	**	0,73	576	0,69	472	0,04	
Mãe ler e escrever	0,99	240	0,97	208	0,02	0,98	265	0,99	209	-0,01		0,98	505	0,98	417	0,00	
Pai ler e escrever	0,97	183	0,92	159	0,05	**	0,94	216	0,95	152	-0,01	0,95	399	0,94	311	0,02	
Irmãos novos	0,53	284	0,52	234	0,01		0,48	311	0,50	241	-0,02	0,50	595	0,51	475	-0,01	
Irmãos velhos	0,56	282	0,46	239	0,10	**	0,53	310	0,49	247	0,04	0,54	592	0,48	486	0,07	**
Pessoas na casa	4,04	281	4,00	241	0,03		4,13	312	4,00	244	0,13	4,09	593	4,00	485	0,08	
Empregada	0,11	276	0,10	239	0,01		0,15	308	0,11	237	0,04	0,13	584	0,10	476	0,03	
Carros	0,96	287	1,03	239	-0,07		1,06	308	0,95	244	0,10	1,01	595	0,99	483	0,02	
Banheiros	1,45	285	1,39	242	0,06		1,50	310	1,40	240	0,11	1,48	595	1,39	482	0,08	*
Quartos	2,09	281	2,03	239	0,05		2,07	309	1,90	244	0,18	***	2,08	1,96	483	0,11	**
Computador	1,37	282	1,44	243	-0,07		1,44	309	1,28	242	0,16	**	1,41	591	1,36	485	0,05
Asfalto	0,83	241	0,80	211	0,04		0,79	274	0,80	219	-0,01		0,81	515	0,80	430	0,01
Energia	0,96	276	0,96	231	0,00		0,96	294	0,95	237	0,01		0,96	570	0,96	468	0,00

Tabela 24 – Teste de balanceamento - Tratamento Professor + Notificação

Variável	Gabarito A'					Gabarito B'					Gabarito AB				
	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença
Gênero	0,43	287	0,48	256	-0,05	0,43	317	0,46	260	-0,03	0,43	604	0,47	516	-0,04
Raça	0,41	290	0,45	254	-0,04	0,40	311	0,40	262	0,00	0,40	601	0,42	516	-0,02
Idade	12,31	293	12,38	262	-0,07	12,24	320	12,35	266	-0,11	12,27	613	12,36	528	-0,09
Ano escolar	7,39	274	7,52	252	-0,13	7,35	304	7,49	253	-0,15	7,37	578	7,50	505	-0,14 *
CRT (%)	0,08	283	0,07	250	0,00	0,08	311	0,09	257	-0,01	0,08	594	0,08	507	0,00
MR (%)	0,43	278	0,44	248	-0,01	0,44	307	0,43	257	0,01	0,43	585	0,43	505	0,00
Mora mãe	0,90	279	0,91	254	-0,01	0,90	305	0,92	256	-0,02	0,90	584	0,92	510	-0,02
Mora pai	0,71	276	0,71	249	0,01	0,75	300	0,70	251	0,05	0,73	576	0,70	500	0,03
Mãe ler e escrever	0,99	240	0,97	219	0,02	0,98	265	0,98	227	0,00	0,98	505	0,97	446	0,01
Pai ler e escrever	0,97	183	0,96	171	0,01	0,94	216	0,98	166	-0,04 **	0,95	399	0,97	337	-0,02
Irmãos novos	0,53	284	0,49	253	0,04	0,48	311	0,43	258	0,05	0,50	595	0,46	511	0,04
Irmãos velhos	0,56	282	0,50	253	0,06	0,53	310	0,52	258	0,01	0,54	592	0,51	511	0,03
Pessoas na casa	4,04	281	4,04	252	0,00	4,13	312	3,94	257	0,19 *	4,09	593	3,99	509	0,09
Empregada	0,11	276	0,12	250	-0,01	0,15	308	0,10	257	0,05 *	0,13	584	0,11	507	0,02
Carros	0,96	287	1,03	250	-0,07	1,06	308	1,00	254	0,06	1,01	595	1,01	504	-0,01
Banheiros	1,45	285	1,50	248	-0,05	1,50	310	1,34	261	0,16 **	1,48	595	1,42	509	0,05
Quartos	2,09	281	2,10	250	-0,01	2,07	309	1,98	259	0,09	2,08	590	2,04	509	0,04
Computador	1,37	282	1,38	251	-0,02	1,44	309	1,36	253	0,08	1,41	591	1,37	504	0,04
Asfalto	0,83	241	0,84	232	0,00	0,79	274	0,82	225	-0,03	0,81	515	0,83	457	-0,02
Energia	0,96	276	0,98	245	-0,02	0,96	294	0,95	247	0,01	0,96	570	0,96	492	-0,01

Tabela 25 – Teste de balanceamento - Tratamento Palestra - Amostra restrita

Variável	Gabarito A'					Gabarito B'					Gabarito AB				
	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença
Gênero	0,37	67	0,41	71	-0,04	0,42	92	0,43	74	-0,01	0,40	159	0,42	145	-0,02
Raça	0,48	67	0,55	71	-0,07	0,40	92	0,39	74	0,01	0,43	159	0,47	145	-0,04
Idade	12,39	67	12,42	71	-0,03	12,29	92	12,22	74	0,08	12,33	159	12,32	145	0,02
Ano escolar	7,54	67	7,49	71	0,04	7,45	92	7,42	74	0,03	7,48	159	7,46	145	0,03
CRT (%)	0,05	67	0,05	71	0,00	0,05	92	0,06	74	-0,01	0,05	159	0,06	145	0,00
MR (%)	0,51	67	0,49	71	0,02	0,45	92	0,48	74	-0,02	0,48	159	0,48	145	0,00
Mora mãe	1,00	67	1,00	71	0,00	1,00	92	1,00	74	0,00	1,00	159	1,00	145	0,00
Mora pai	1,00	67	1,00	71	0,00	1,00	92	1,00	74	0,00	1,00	159	1,00	145	0,00
Mãe ler e escrever	1,00	67	0,99	71	0,01	1,00	92	1,00	74	0,00	1,00	159	0,99	145	0,01
Pai ler e escrever	0,99	67	0,97	71	0,01	1,00	92	0,97	74	0,03	0,99	159	0,97	145	0,02
Irmãos novos	0,57	67	0,45	71	0,12	0,59	92	0,47	74	0,11	0,58	159	0,46	145	0,12 **
Irmãos velhos	0,63	67	0,54	71	0,09	0,53	92	0,50	74	0,03	0,57	159	0,52	145	0,06
Pessoas na casa	4,21	67	3,94	71	0,27	4,33	92	4,12	74	0,20	4,28	159	4,03	145	0,24 **
Empregada	0,10	67	0,11	71	-0,01	0,12	92	0,11	74	0,01	0,11	159	0,11	145	0,00
Carros	0,91	67	0,97	71	-0,06	1,12	92	1,08	74	0,04	1,03	159	1,03	145	0,00
Banheiros	1,48	67	1,54	71	-0,06	1,43	92	1,41	74	0,03	1,45	159	1,47	145	-0,02
Quartos	2,30	67	2,18	71	0,12	2,23	92	2,15	74	0,08	2,26	159	2,17	145	0,09
Computador	1,36	67	1,32	71	0,03	1,48	92	1,49	74	-0,01	1,43	159	1,41	145	0,02
Asfalto	0,91	67	0,92	71	-0,01	0,82	92	0,92	74	-0,10 *	0,86	159	0,92	145	-0,06 *
Energia	0,97	67	0,96	71	0,01	1,00	92	0,99	74	0,01	0,99	159	0,97	145	0,02

Tabela 26 – Teste de balanceamento - Tratamento Palestra + Notificação - Amostra restrita

Variável	Gabarito A'					Gabarito B'					Gabarito AB				
	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença
Gênero	0,37	67	0,45	82	-0,08	0,42	92	0,40	90	0,02	0,40	159	0,42	172	-0,02
Raça	0,48	67	0,50	82	-0,02	0,40	92	0,46	90	-0,05	0,43	159	0,48	172	-0,04
Idade	12,39	67	12,32	82	0,07	12,29	92	12,47	90	-0,17	12,33	159	12,40	172	-0,06
Ano escolar	7,54	67	7,46	82	0,07	7,45	92	7,73	90	-0,29 *	7,48	159	7,60	172	-0,12
CRT (%)	0,05	67	0,05	82	0,01	0,05	92	0,04	90	0,01	0,05	159	0,05	172	0,01
MR (%)	0,51	67	0,41	82	0,11 **	0,45	92	0,45	90	0,00	0,48	159	0,43	172	0,05 *
Mora mãe	1,00	67	1,00	82	0,00	1,00	92	1,00	90	0,00	1,00	159	1,00	172	0,00
Mora pai	1,00	67	1,00	82	0,00	1,00	92	1,00	90	0,00	1,00	159	1,00	172	0,00
Mãe ler e escrever	1,00	67	0,98	82	0,02	1,00	92	1,00	90	0,00	1,00	159	0,99	172	0,01
Pai ler e escrever	0,99	67	1,00	82	-0,01	1,00	92	0,98	90	0,02	0,99	159	0,99	172	0,01
Irmãos novos	0,57	67	0,45	82	0,12	0,59	92	0,51	90	0,08	0,58	159	0,48	172	0,10 *
Irmãos velhos	0,63	67	0,63	82	-0,01	0,53	92	0,53	90	0,00	0,57	159	0,58	172	-0,01
Pessoas na casa	4,21	67	3,98	82	0,23	4,33	92	4,20	90	0,13	4,28	159	4,09	172	0,18
Empregada	0,10	67	0,12	82	-0,02	0,12	92	0,17	90	-0,05	0,11	159	0,15	172	-0,03
Carros	0,91	67	1,07	82	-0,16	1,12	92	1,06	90	0,06	1,03	159	1,06	172	-0,03
Banheiros	1,48	67	1,44	82	0,04	1,43	92	1,57	90	-0,13	1,45	159	1,51	172	-0,05
Quartos	2,30	67	2,21	82	0,09	2,23	92	2,10	90	0,13	2,26	159	2,15	172	0,11
Computador	1,36	67	1,56	82	-0,20	1,48	92	1,43	90	0,04	1,43	159	1,49	172	-0,07
Asfalto	0,91	67	0,84	82	0,07	0,82	92	0,87	90	-0,05	0,86	159	0,85	172	0,00
Energia	0,97	67	0,98	82	-0,01	1,00	92	1,00	90	0,00	0,99	159	0,99	172	0,00

Tabela 27 – Teste de balanceamento - Tratamento Professor - Amostra restrita

Variável	Gabarito A'					Gabarito B'					Gabarito AB				
	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença
Gênero	0,37	67	0,41	76	-0,03	0,42	92	0,42	67	0,01	0,40	159	0,41	143	-0,01
Raça	0,48	67	0,49	76	-0,01	0,40	92	0,51	67	-0,11	0,43	159	0,50	143	-0,06
Idade	12,39	67	12,33	76	0,06	12,29	92	12,45	67	-0,15	12,33	159	12,38	143	-0,05
Ano escolar	7,54	67	7,59	76	-0,05	7,45	92	7,57	67	-0,12	7,48	159	7,58	143	-0,10
CRT (%)	0,05	67	0,07	76	-0,01	0,05	92	0,03	67	0,02	0,05	159	0,05	143	0,00
MR (%)	0,51	67	0,46	76	0,05	0,45	92	0,44	67	0,01	0,48	159	0,45	143	0,03
Mora mãe	1,00	67	1,00	76	0,00	1,00	92	1,00	67	0,00	1,00	159	1,00	143	0,00
Mora pai	1,00	67	1,00	76	0,00	1,00	92	1,00	67	0,00	1,00	159	1,00	143	0,00
Mãe ler e escrever	1,00	67	0,96	76	0,04	1,00	92	1,00	67	0,00	1,00	159	0,98	143	0,02
Pai ler e escrever	0,99	67	0,95	76	0,04	1,00	92	1,00	67	0,00	0,99	159	0,97	143	0,02
Irmãos novos	0,57	67	0,45	76	0,12	0,59	92	0,54	67	0,05	0,58	159	0,49	143	0,09
Irmãos velhos	0,63	67	0,53	76	0,10	0,53	92	0,51	67	0,03	0,57	159	0,52	143	0,05
Pessoas na casa	4,21	67	4,04	76	0,17	4,33	92	4,48	67	-0,15	4,28	159	4,24	143	0,03
Empregada	0,10	67	0,05	76	0,05	0,12	92	0,10	67	0,02	0,11	159	0,08	143	0,04
Carros	0,91	67	1,01	76	-0,10	1,12	92	1,01	67	0,10	1,03	159	1,01	143	0,02
Banheiros	1,48	67	1,45	76	0,03	1,43	92	1,67	67	-0,24	1,45	159	1,55	143	-0,10
Quartos	2,30	67	2,09	76	0,21	2,23	92	2,22	67	0,00	2,26	159	2,15	143	0,10
Computador	1,36	67	1,46	76	-0,10	1,48	92	1,36	67	0,12	1,43	159	1,41	143	0,02
Asfalto	0,91	67	0,91	76	0,00	0,82	92	0,88	67	-0,07	0,86	159	0,90	143	-0,04
Energia	0,97	67	0,99	76	-0,02	1,00	92	0,99	67	0,01	0,99	159	0,99	143	0,00

Tabela 28 – Teste de balanceamento - Tratamento Professor + Notificação - Amostra restrita

Variável	Gabarito A'					Gabarito B'					Gabarito AB				
	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença	Média Controle	Amostra Controle	Média Tratamento	Amostra Tratamento	Diferença
Gênero	0,37	67	0,47	81	-0,10	0,42	92	0,49	81	-0,07	0,40	159	0,48	162	-0,08
Raça	0,48	67	0,43	81	0,05	0,40	92	0,46	81	-0,05	0,43	159	0,44	162	-0,01
Idade	12,39	67	12,56	81	-0,17	12,29	92	12,51	81	-0,21	12,33	159	12,53	162	-0,20 *
Ano escolar	7,54	67	7,77	81	-0,23	7,45	92	7,81	81	-0,37 **	7,48	159	7,79	162	-0,31 **
CRT (%)	0,05	67	0,06	81	-0,01	0,05	92	0,05	81	0,01	0,05	159	0,05	162	0,00
MR (%)	0,51	67	0,50	81	0,01	0,45	92	0,50	81	-0,04	0,48	159	0,50	162	-0,02
Mora mãe	1,00	67	1,00	81	0,00	1,00	92	1,00	81	0,00	1,00	159	1,00	162	0,00
Mora pai	1,00	67	1,00	81	0,00	1,00	92	1,00	81	0,00	1,00	159	1,00	162	0,00
Mãe ler e escrever	1,00	67	0,98	81	0,02	1,00	92	1,00	81	0,00	1,00	159	0,99	162	0,01
Pai ler e escrever	0,99	67	0,99	81	0,00	1,00	92	0,99	81	0,01	0,99	159	0,99	162	0,01
Irmãos novos	0,57	67	0,49	81	0,07	0,59	92	0,57	81	0,02	0,58	159	0,53	162	0,05
Irmãos velhos	0,63	67	0,58	81	0,05	0,53	92	0,53	81	0,00	0,57	159	0,56	162	0,02
Pessoas na casa	4,21	67	4,01	81	0,20	4,33	92	4,12	81	0,20	4,28	159	4,07	162	0,21 *
Empregada	0,10	67	0,09	81	0,02	0,12	92	0,11	81	0,01	0,11	159	0,10	162	0,01
Carros	0,91	67	1,15	81	-0,24 *	1,12	92	1,04	81	0,08	1,03	159	1,09	162	-0,06
Banheiros	1,48	67	1,54	81	-0,07	1,43	92	1,38	81	0,05	1,45	159	1,46	162	-0,01
Quartos	2,30	67	2,12	81	0,18	2,23	92	2,02	81	0,20 *	2,26	159	2,07	162	0,18 **
Computador	1,36	67	1,42	81	-0,06	1,48	92	1,53	81	-0,05	1,43	159	1,48	162	-0,05
Asfalto	0,91	67	0,86	81	0,05	0,82	92	0,89	81	-0,07	0,86	159	0,88	162	-0,02
Energia	0,97	67	0,98	81	-0,01	1,00	92	0,99	81	0,01	0,99	159	0,98	162	0,01

Apêndice D - Resultados via análise de regressão

Nessa análise, foi investigado o impacto das abordagens propostas em um contexto de regressão para controlar as variáveis que não se mostraram balanceadas nas suas respectivas comparações e possivelmente obter um efeito mais limpo das abordagens, ou seja:

$$R_{ijkm} = \beta_0 + \beta_1 * tratamento_{im} + \Phi * W_{im} + \varepsilon_{ijkm} \quad (5.1)$$

onde os subscritos são os mesmos apresentados anteriormente, com exceção do m adicionado agora que se refere ao tipo de tratamento do experimento (palestra, palestra + notificação, professor ou professor + notificação) e recebe valor 1 se o aluno i estava no grupo de tratamento ou 0 se estava no grupo de controle. O vetor W_{im} captura as variáveis que não se mostraram balanceadas no tratamento m para o indivíduo i .

Tabela 29 – Resultados da análise dos itens via regressão

	Palestra				Professor			
	Sentido	Sentido	Efeito		Sentido	Sentido	Efeito	
	DS	contrário	descontado		DS	contrário	descontado	
	N Itens	N Itens	X_l	RC	N Itens	N Itens	X_l	RC
BFI - E	0	2	-2	$(-\infty, -2]$ U $[3, +\infty)$	1	2	-1	$(-\infty, -2]$ U $[3, +\infty)$
BFI - A	1	1	0	$(-\infty, -2]$ U $[2, +\infty)$	0	2	-2	$(-\infty, -2]$ U $[2, +\infty)$
BFI - C	5	0	5	$(-\infty, -2]$ U $[2, +\infty)$	0	1	-1	$(-\infty, -2]$ U $[2, +\infty)$
BFI - N	1	2	-1	$(-\infty, -2]$ U $[3, +\infty)$	2	3	-1	$(-\infty, -2]$ U $[3, +\infty)$
BFI - O	1	0	1	$(-\infty, -3]$ U $[3, +\infty)$	4	1	3	$(-\infty, -3]$ U $[3, +\infty)$
NSLCS	7	1	6	$(-\infty, -4]$ U $[3, +\infty)$	2	4	-2	$(-\infty, -4]$ U $[3, +\infty)$
GS	2	1	1	$(-\infty, -3]$ U $[3, +\infty)$	0	0	0	$(-\infty, -3]$ U $[3, +\infty)$
SDQ - Em	0	1	-1	$(-\infty, -2]$ U $[2, +\infty)$	0	1	-1	$(-\infty, -2]$ U $[2, +\infty)$
SDQ - Cond	1	0	1	$(-\infty, -2]$ U $[2, +\infty)$	0	0	0	$(-\infty, -2]$ U $[2, +\infty)$
SDQ - Col	1	0	1	$(-\infty, -2]$ U $[2, +\infty)$	0	0	0	$(-\infty, -2]$ U $[2, +\infty)$
SDQ - Hip	0	0	0	$(-\infty, -2]$ U $[2, +\infty)$	1	0	1	$(-\infty, -2]$ U $[2, +\infty)$
SDQ - PS	2	0	2	$(-\infty, -2]$ U $[2, +\infty)$	1	1	0	$(-\infty, -2]$ U $[2, +\infty)$
RSES	2	0	2	$(-\infty, -3]$ U $[2, +\infty)$	4	0	4	$(-\infty, -3]$ U $[2, +\infty)$
CSE	4	0	4	$(-\infty, -3]$ U $[3, +\infty)$	0	1	-1	$(-\infty, -3]$ U $[3, +\infty)$
SEQC - Ac	3	0	3	$(-\infty, -2]$ U $[2, +\infty)$	1	0	1	$(-\infty, -2]$ U $[2, +\infty)$
SEQC - Sc	2	1	1	$(-\infty, -2]$ U $[2, +\infty)$	1	0	1	$(-\infty, -2]$ U $[2, +\infty)$
SEQC - Em	2	0	2	$(-\infty, -3]$ U $[2, +\infty)$	3	0	3	$(-\infty, -3]$ U $[2, +\infty)$

Tabela 30 – Resultados da análise dos itens via regressão - Notificação

	Palestra + Notificação				Professor + Notificação			
	Sentido	Sentido	Efeito		Sentido	Sentido	Efeito	
	DS	contrário	descontado		DS	contrário	descontado	
	N Itens	N Itens	Z_l	RC	N Itens	N Itens	Z_l	RC
BFI - E	2	1	-3	$(-\infty, -3] \cup [3, +\infty)$	1	1	-1	$(-\infty, -3] \cup [3, +\infty)$
BFI - A	1	1	0	$(-\infty, -3] \cup [3, +\infty)$	0	2	0	$(-\infty, -3] \cup [3, +\infty)$
BFI - C	1	4	8	$(-\infty, -3] \cup [3, +\infty)$	0	3	2	$(-\infty, -3] \cup [3, +\infty)$
BFI - N	1	0	-2	$(-\infty, -3] \cup [3, +\infty)$	0	0	-1	$(-\infty, -3] \cup [3, +\infty)$
BFI - O	1	1	1	$(-\infty, -4] \cup [3, +\infty)$	1	1	3	$(-\infty, -4] \cup [3, +\infty)$
NSLCS	1	5	10	$(-\infty, -3] \cup [3, +\infty)$	4	2	-4	$(-\infty, -3] \cup [3, +\infty)$
GS	0	1	2	$(-\infty, -2] \cup [2, +\infty)$	1	1	0	$(-\infty, -2] \cup [2, +\infty)$
SDQ - Em	1	0	-2	$(-\infty, -3] \cup [2, +\infty)$	0	3	2	$(-\infty, -3] \cup [2, +\infty)$
SDQ - Cond	1	0	0	$(-\infty, -3] \cup [3, +\infty)$	1	0	-1	$(-\infty, -3] \cup [3, +\infty)$
SDQ - Col	1	1	1	$(-\infty, -2] \cup [2, +\infty)$	0	0	0	$(-\infty, -2] \cup [2, +\infty)$
SDQ - Hip	0	0	0	$(-\infty, -3] \cup [3, +\infty)$	1	0	0	$(-\infty, -3] \cup [3, +\infty)$
SDQ - PS	1	0	1	$(-\infty, -3] \cup [3, +\infty)$	0	2	2	$(-\infty, -3] \cup [3, +\infty)$
RSES	2	0	0	$(-\infty, -3] \cup [3, +\infty)$	0	0	4	$(-\infty, -3] \cup [3, +\infty)$
CSE	1	0	3	$(-\infty, -3] \cup [3, +\infty)$	0	1	0	$(-\infty, -3] \cup [3, +\infty)$
SEQC - Ac	0	0	3	$(-\infty, -2] \cup [2, +\infty)$	0	1	2	$(-\infty, -2] \cup [2, +\infty)$
SEQC - Sc	2	0	-1	$(-\infty, -2] \cup [2, +\infty)$	1	1	1	$(-\infty, -2] \cup [2, +\infty)$
SEQC - Em	2	0	0	$(-\infty, -2] \cup [2, +\infty)$	0	0	3	$(-\infty, -2] \cup [2, +\infty)$

Tabela 31 – Resultados da análise dos constructos via regressão

	Palestra		Professor		Palestra + Notificação		Professor + Notificação			
	β_1	P-valor ¹	β_1	P-valor ¹	β_1	P-valor ¹	β_1	P-valor ¹		
BFI - E ²	-0,28	0,79	-0,39	0,86	0,35	0,16	-0,35	0,84		
BFI - A	0,11	0,38	-0,34	0,83	-0,04	0,55	-0,31	0,81		
BFI - C	0,78	0,03	**	-0,21	0,69	-0,55	0,92	*	-0,49	0,89
BFI - N ³	0,32	0,82	0,08	0,58	0,10	0,62	0,02	0,52		
BFI - O	0,39	0,16	0,11	0,39	0,12	0,38	-0,44	0,87		
NSLCS ³	-0,59	0,04	**	-0,01	0,49	0,15	0,73	-0,05	0,43	
GS	0,08	0,12	-0,01	0,53	-0,05	0,87	-0,05	0,83		
SDQ - Em ³	0,18	0,82	0,17	0,81	-0,10	0,29	0,38	0,94	*	
SDQ - Cond ³	-0,03	0,42	0,03	0,57	0,00	0,50	-0,08	0,35		
SDQ - Col ³	-0,01	0,48	-0,12	0,25	-0,11	0,25	-0,08	0,35		
SDQ - Hip ³	-0,01	0,48	-0,10	0,38	0,15	0,79	-0,23	0,18		
SDQ - PS	0,12	0,26	0,30	0,06	*	0,20	0,12	0,15	0,24	
SDQ ³	0,13	0,60	0,03	0,52	-0,07	0,44	-0,01	0,49		
RSES	0,02	0,48	0,53	0,09	*	0,32	0,20	-0,54	0,87	
CSE	0,04	0,18	-0,01	0,61	0,02	0,29	-0,04	0,72		
SEQC - Ac	0,63	0,11	0,22	0,34	0,15	0,35	-0,22	0,71		
SEQC - Sc	0,53	0,16	0,85	0,05	**	0,57	0,06	*	-0,08	0,58
SEQC - Em	0,57	0,16	0,91	0,05	*	0,44	0,15	0,05	0,45	
DS	0,08	0,32	0,03	0,42	-0,13	0,80	0,06	0,35		

¹: P-valor de teste unicaudal no sentido da desejabilidade social.

²: A dimensão extroversão é a menos consensual em relação ao sentido da desejabilidade social, principalmente em relação à monotonicidade do efeito. Considerou-se para a análise, maior extroversão relacionada a maior desejabilidade social.

³: Constructo em que há correlação negativa com desejabilidade social.

*, **, ***: estatisticamente significativa a 10%, 5% e 1% respectivamente.

*, **, ***: estatisticamente significativa a 10%, 5% e 1% respectivamente, mas no sentido oposto da desejabilidade social.

Apêndice E - *Heat maps* - Itens

A ordem dos constructos é sempre BFI-E, BFI-A, BFI-C, BFI-N, BFI-O, DS, SEQC-Ac, SEQC-Sc, SEQC-Em, NSLCS, GS, SDQ-Em, SDQ-Cond, SDQ-Hip, SDQ-Col, SDQ-PS, CSE e RSES. Após os itens que compõe cada constructo, foram incluídas as escalas em suas formas agregadas que estão delimitados com a linha pontilhada.

Figura 12 – *Heat map* dos itens - Grupo de controle

Figura 13 – *Heat map* dos itens - Grupo de tratamento: Palestra

Figura 14 – *Heat map* dos itens - Grupo de tratamento: Palestra + Notificação

Figura 15 – *Heat map* dos itens - Grupo de tratamento: Professor

Figura 16 – *Heat map* dos itens - Grupo de tratamento: Professor + Notificação

Apêndice F - Questionários

A seguir os questionários A e B nas suas versões com a notificação:

Caro aluno, obrigado pela sua participação!

Leia atentamente todas as orientações ao longo do questionário e marque as opções escolhidas na **folha de gabarito** com caneta preenchendo toda a área do quadrado conforme o exemplo abaixo:

Questão X: ☐ ☒ ☐ ☐ ☐

Esse questionário possui mecanismos para **detectar respostas inconsistentes**/mentirosas. Caso esse tipo de resposta seja encontrado, o questionário será invalidado. Por isso, **responda de forma sincera** todas as questões.

Parte I

Na parte I **não há respostas corretas**. Marque aquelas que mais se assemelham a você mesmo. **Cada bloco de perguntas possui alternativas diferentes** de respostas, fique atento às opções na folha de gabarito.

Responda às seguintes questões pensando em como você é, se sente ou se comporta na maioria das vezes .	
1	Sou falante.
2	Costumo achar defeitos nos outros.
3	Sou caprichoso(a) e detalhista nas tarefas escolares.
4	Sou meio triste.
5	Sou original, tenho ideias novas.
6	Sou reservado(a), fico mais na minha.
7	Não sou egoísta e gosto de ajudar os outros.
8	Sou meio desleixado(a), não tenho cuidado na hora de fazer as coisas.
9	Sou calmo(a) e controlo bem meu estresse.
10	Muitos assuntos despertam minha curiosidade.
11	Sou cheio de energia.
12	Começo bate-bocas com os outros.
13	Sou um(a) aluno(a) dedicado(a).
14	Sou meio tenso(a).
15	Gosto de pensar profundamente sobre as coisas.
16	Contagio os outros com meu entusiasmo.
17	Tenho facilidade em perdoar.
18	Costumo ser desorganizado(a).
19	Eu me preocupo demais com tudo.
20	Tenho uma imaginação bem ativa.
21	Costumo ser quieto(a).
22	Confio nas pessoas.
23	Costumo ser preguiçoso(a).
24	Sou emocionalmente estável, dificilmente fico irritado(a).
25	Sou inventivo(a).
26	Não tenho medo de expressar o que penso.
27	Posso ser indiferente, frio(a) e distante dos outros.
28	Não desisto até terminar a tarefa.
29	Fico de mau humor com facilidade.
30	Gosto de atividades artísticas.
31	Sou tímido(a), inibido(a).
32	Sou amável e legal com quase todo mundo.
33	Faço as tarefas bem e sem desperdício de tempo.
34	Mantenho a calma em situações difíceis.

Controle 1: A Controle 2: C Controle 3:

35	Gosto de ter uma rotina de tarefas.
36	Sou desinibido(a), sociável.
37	Sou grosseiro(a) com os outros.
38	Faço planos e os levo adiante.
39	Fico nervoso(a) com facilidade.
40	Gosto de refletir e brincar com minhas ideias.
41	Tenho poucos interesses artísticos.
42	Gosto de cooperar com os outros.
43	Eu desvio minha atenção com muita facilidade.
44	Conheço vários tipos de obras de arte, de música e/ou de literatura.

Avalie na escala abaixo **quanto você consegue**:

45	Estudar mesmo tendo outras coisas interessantes para fazer.
46	Estudar um texto para uma prova.
47	Terminar todo seu dever de casa.
48	Prestar atenção nas aulas.
49	Passar de ano sem dependência.
50	Deixar seus pais satisfeitos com seu desempenho na escola.
51	Ir bem em uma prova.
52	Expressar suas opiniões quando outros colegas discordam de você.
53	Se tornar amigo(a) de outras crianças/jovens.
54	Bater um papo com uma pessoa desconhecida.
55	Trabalhar em harmonia com seus colegas.
56	Dizer a outras crianças/jovens que eles estão fazendo algo que você não gosta.
57	Contar uma coisa engraçada para um grupo de colegas.
58	Permanecer amigo(a) de outras crianças/jovens.
59	Se reanimar quando algo desagradável acontece.
60	Se acalmar depois de ficar muito assustado(a).
61	Evitar ficar nervoso(a).
62	Controlar seus sentimentos.
63	Ter com você mesmo uma conversa de auto-motivação quando se sente para baixo.
64	Evitar pensamentos desagradáveis.
65	Não se preocupar com coisas que possam acontecer.

Leia com atenção e responda as perguntas abaixo.

66	Você acredita que a maioria dos problemas se resolve quando você os leva a sério?
67	Você é frequentemente acusado(a) por coisas que não foram sua culpa?
68	Você sente que muitas vezes não vale a pena se esforçar porque, de qualquer modo, as coisas nunca dão certo mesmo?
69	Você sente que os pais escutam o que os filhos têm a dizer?
70	Quando você recebe uma punição, você costuma achar injusta?
71	Você acha difícil fazer um(a) amigo(a) mudar de opinião?
72	Você sente que é quase impossível mudar a opinião que seus pais têm sobre as coisas?
73	Quando faz alguma coisa errada, você sente que existe muito pouco que você pode fazer para consertar?
74	Você acredita que as pessoas boas nos esportes já nasceram assim?
75	Você sente que a melhor maneira de lidar com os problemas é apenas não pensar neles?
76	Quando algum(a) colega decide te bater, você sente que pouco pode fazer para impedi-lo(a)?
77	Você já sentiu pessoas serem más contigo sem motivo algum?
78	Você acredita que pode mudar o que acontecerá com você amanhã pelo que faz hoje?

Controle 1: A Controle 2: C Controle 3:

79	Você acredita que quando coisas ruins vão acontecer, elas simplesmente acontecem, não importa o que você tente fazer para impedi-las?
80	Você sente que não adianta tentar fazer as coisas do seu jeito em casa?
81	Quando algum(a) colega decide ser seu/sua inimigo(a), você acredita que pode fazer pouca coisa para mudar a situação?
82	Você sente que tem pouca influência em casa para decidir sobre em que canal deve estar a TV?
83	Quando alguém de sua idade não gosta de você, você acredita que há pouco o que fazer para mudar a situação?
84	Você sente que é inútil se esforçar na escola porque a maioria dos outros alunos é mais inteligente que você?
85	Você é o tipo de pessoa que acredita que planejar o futuro faz as coisas darem mais certo?
86	Você sente que tem pouca influência nas decisões do cotidiano em sua casa?

Abaixo mostramos algumas frases que podem ou não ter a ver com você. Escolha uma das opções, na escala apresentada no gabarito, que melhor indique **sua opinião em relação a você mesmo**.

87	Novas ideias e novos projetos desviam minha atenção dos anteriores.
88	Contratempos (atrasos e dificuldades) não me desencorajam.
89	Eu me interessava muito pelas coisas por pouco tempo e depois me desinteressei.
90	Sou um(a) aluno(a) que se esforça.
91	Traço um objetivo, mas no meio do caminho acabo escolhendo outro para seguir.
92	Tenho dificuldade em manter minha atenção em atividades que demorem alguns meses para terminar.
93	Termino tudo que começo.
94	Sou um(a) aluno(a) cuidadoso(a) e dedicado(a).

Leia com atenção cada frase e marque se ela é falsa ou verdadeira em relação a você.

95	Às vezes eu me sinto ressentido(a) se as coisas não são do meu jeito.
96	Em algumas ocasiões, eu desisti de fazer algumas coisas porque duvidei da minha habilidade.
97	Houve ocasiões em que eu me rebelei contra pessoas com autoridade, mesmo sabendo que elas estavam certas.
98	Não importa com quem eu esteja falando, eu sou sempre um(a) bom/boa ouvinte.
99	Eu me lembro de ter fingido estar doente para conseguir algo.
100	Houve ocasiões nas quais levei vantagem sobre alguém.
101	Eu estou sempre disposto(a) a admitir quando cometo um erro.
102	Às vezes, eu tento acertar as contas em vez de perdoar e esquecer.
103	Eu sou sempre educado(a), mesmo com pessoas desagradáveis.
104	Eu nunca fiquei aborrecido(a) quando pessoas expressaram ideias muito diferentes das minhas.
105	Houve vezes em que senti inveja da boa sorte dos outros.
106	Eu fico às vezes irritado(a) com pessoas que me pedem favores.
107	Eu nunca disse algo que machucasse os sentimentos de alguém de propósito.

Controle 1: A Controle 2: C Controle 3:

Parte II

Leia as três questões abaixo com atenção e responda a alternativa correta na folha de gabarito:

- 1) Uma raquete e uma bola custam justas R\$1,10 reais. A raquete custa R\$1,00 a mais que a bola. Qual o preço da bola (resposta em centavos)?
- 2) Se 5 máquinas demoram 5 minutos para fazer 5 objetos, quanto tempo demora para 100 máquinas fazerem 100 objetos (resposta em minutos)?
- 3) Parte de um lago está coberto por flores. Todos os dias, as flores se multiplicam de forma que a área do lago coberta dobra de tamanho. São necessárias 6 dias para cobrir o lago inteiro. Quantos dias são necessários para cobrir metade do lago (resposta em dias)?

Parte III

Cada figura abaixo segue um padrão de preenchimento, mas em todas elas há um espaço vazio. Escolha, dentre as alternativas expostas abaixo de cada figura a que melhor preenche esse espaço e em seguida marque a opção na folha de gabarito:

1)

2)

3)

4)

Controle 1: A Controle 2: C Controle 3:

Parte IV

Por fim, responda às seguintes questões relacionadas a você, sua família e sua residência escolhendo apenas uma alternativa por questão e assinalando-a na folha de gabarito. Se você morar em mais de uma casa, responda sobre aquela na qual você passa mais tempo. Se houver dúvidas, pergunte ao aplicador:

1	Quantos anos você tem?	11 ou menos A	12 B	13 C	14 D	15 ou mais E
2	Que ano escolar você frequenta?	5 ano A	6 ano B	7 ano C	8 ano D	9 ano E
3	Qual seu sexo?				Masculino A	Feminino B
4	Como você se considera?	Amarelo A	Branco B	Indígena C	Negro D	Pardo E
5	Você mora com sua mãe ou com alguma mulher responsável por você? Se não mora, marque "NA" na próxima questão.				Sim A	Não B
6	Sua mãe ou mulher responsável por você sabe ler e escrever?	Sim A	Não B	Não sei C	NA D	
7	Você mora com seu pai ou com algum homem responsável por você? Se não mora, marque "NA" na próxima questão.				Sim A	Não B
8	Seu pai ou homem responsável por você sabe ler e escrever?	Sim A	Não B	Não sei C	NA D	
9	Você tem irmãos? Se não tem, marque "NA" nas duas próximas questões				Sim A	Não B
10	Você tem irmãos mais velhos?	Sim A			Não B	NA C
11	Você tem irmãos mais novos?	Sim A			Não B	NA C
12	Quantas pessoas moram na sua casa?	2 A	3 B	4 C	5 D	6 ou mais E
13	Possui empregada doméstica ou faxineira na sua casa?				Sim A	Não B
14	Quantos carros no total possuem as pessoas que moram com você?	Nenhum A	1 B	2 C	3 ou mais D	
15	Quantos banheiros tem na sua casa?	Nenhum A	1 B	2 C	3 ou mais D	
16	Quantos quartos tem na sua casa?	Nenhum A	1 B	2 C	3 ou mais D	
17	Quantos computadores ou notebooks tem na sua casa?	Nenhum A	1 B	2 C	3 ou mais D	
18	A rua onde você mora é asfaltada?	Sim A			Não B	Não sei C
19	Tem energia elétrica na sua casa?	Sim A			Não B	Não sei C

Caro aluno, obrigado pela sua participação!

Leia atentamente todas as orientações ao longo do questionário e marque as opções escolhidas na **folha de gabarito** com caneta preenchendo toda a área do quadrado conforme o exemplo abaixo:

Questão X: ☐ ☒ ☐ ☐ ☐

Esse questionário possui mecanismos para **detectar respostas inconsistentes/mentirosas**. Caso esse tipo de resposta seja encontrado, o questionário será invalidado. Por isso, **responda de forma sincera** todas as questões.

Parte I

Na parte I **não há respostas corretas**. Marque aquelas que mais se assemelham a você mesmo. **Cada bloco de perguntas possui alternativas diferentes** de respostas, fique atento às opções na folha de gabarito.

Responda às seguintes questões pensando em como você é, se sente ou se comporta na maioria das vezes .	
1	Sou falante.
2	Costumo achar defeitos nos outros.
3	Sou caprichoso(a) e detalhista nas tarefas escolares.
4	Sou meio triste.
5	Sou original, tenho ideias novas.
6	Sou reservado(a), fico mais na minha.
7	Não sou egoísta e gosto de ajudar os outros.
8	Sou meio desleixado(a), não tenho cuidado na hora de fazer as coisas.
9	Sou calmo(a) e controlo bem meu estresse.
10	Muitos assuntos despertam minha curiosidade.
11	Sou cheio de energia.
12	Começo bate-bocas com os outros.
13	Sou um(a) aluno(a) dedicado(a).
14	Sou meio tenso(a).
15	Gosto de pensar profundamente sobre as coisas.
16	Contagio os outros com meu entusiasmo.
17	Tenho facilidade em perdoar.
18	Costumo ser desorganizado(a).
19	Eu me preocupo demais com tudo.
20	Tenho uma imaginação bem ativa.
21	Costumo ser quieto(a).
22	Confio nas pessoas.
23	Costumo ser preguiçoso(a).
24	Sou emocionalmente estável, dificilmente fico irritado(a).
25	Sou inventivo(a).
26	Não tenho medo de expressar o que penso.
27	Posso ser indiferente, frio(a) e distante dos outros.
28	Não desisto até terminar a tarefa.
29	Fico de mau humor com facilidade.
30	Gosto de atividades artísticas.
31	Sou tímido(a), inibido(a).
32	Sou amável e legal com quase todo mundo.
33	Faço as tarefas bem e sem desperdício de tempo.
34	Mantenho a calma em situações difíceis.

Código 1: B

Código 2: C

Código 3:

35	Gosto de ter uma rotina de tarefas.
36	Sou desinibido(a), sociável.
37	Sou grosseiro(a) com os outros.
38	Faço planos e os levo adiante.
39	Fico nervoso(a) com facilidade.
40	Gosto de refletir e brincar com minhas ideias.
41	Tenho poucos interesses artísticos.
42	Gosto de cooperar com os outros.
43	Eu desvio minha atenção com muita facilidade.
44	Conheço vários tipos de obras de arte, de música e/ou de literatura.

Abaixo mostramos algumas características pessoais que podem ou não ter a ver com você. Por favor, escolha umas das opções, na folha de resposta, que melhor indique **sua opinião em relação a você mesmo**. Dê sua resposta baseado em como as coisas têm sido nos **últimos seis meses**.

45	Tento ser legal com outras pessoas. Eu me preocupo com os sentimentos deles/delas.
46	Sou agitado(a), não consigo ficar parado(a) por muito tempo.
47	Sinto muitas dores de cabeça, dores de barriga ou enjoos.
48	Costumo dividir/emprestar minhas coisas (comida, jogos/brinquedos, canetas, etc.).
49	Fico muito bravo(a) e costumo perder a paciência.
50	Costumo estar sozinho(a), brinco/jogo sozinho(a), fico na minha.
51	Costumo fazer o que me mandam.
52	Eu me preocupo muito com tudo.
53	Ofereço ajuda a quem fica chateado, magoado ou sentindo-se doente.
54	Fico o tempo todo inquieto(a) (balançando a perna, mexendo no cabelo, estalando os dedos).
55	Tenho pelo menos um(a) amigo/amiga de verdade.
56	Brigo muito. Consigo fazer com que outras pessoas façam o que eu quero.
57	Eu me sinto infeliz, desanimado(a) ou com vontade de chorar.
58	Meus colegas gostam de mim.
59	Eu me distraio com facilidade, acho difícil me concentrar.
60	Fico nervoso(a) com situações novas. Perco a confiança em mim com facilidade.
61	Sou legal com crianças mais novas.
62	Costumo ser acusado(a) de mentir ou enganar.
63	Meus colegas pegam no meu pé.
64	Costumo me oferecer para ajudar os outros (pais, professores, crianças).
65	Penso antes de fazer as coisas.
66	Pego coisas que não são minhas em casa, na escola ou em outros lugares.
67	Eu me dou melhor com adultos do que com meus colegas.
68	Tenho muitos medos, eu me assusto facilmente.
69	Termino as tarefas que começo. Consigo prestar atenção.

Abaixo mostramos algumas frases que podem ou não ter a ver com você. Por favor, escolha uma das opções, na folha de gabarito, que melhor indique **sua opinião em relação a você mesmo**.

70	Acredito que sou reconhecido como mereço pela vida.
71	Eu me sinto deprimido.
72	Quando eu tento fazer alguma coisa, eu consigo.
73	Quando eu falho me sinto péssimo.
74	Completo as tarefas com sucesso.
75	Não sinto que controlo minhas atividades escolares.
76	Sou satisfeito comigo mesmo
77	Tenho muitas dúvidas se sou competente.
78	Eu determino o que vai acontecer na minha vida.

Código 1: B

Código 2: C

Código 3:

79	Não acredito que controlo o sucesso que vou ter na minha carreira.
80	Sou capaz de enfrentar a maioria dos meus problemas.
81	As coisas podem parecer muito tristes e sem esperança para mim.

Abaixo mostramos algumas frases que podem ou não ter a ver com você. Por favor, escolha uma das opções, na folha de gabarito, que melhor indique **sua opinião em relação a você mesmo**.

82	Acredito que valho pelo menos tanto quanto meus/minhas colegas.
83	Acho que tenho várias qualidades.
84	No geral, eu tendo a achar que sou um fracasso.
85	Sou capaz de fazer outras coisas tão bem quanto meus/minhas colegas.
86	Não tenho do que me orgulhar.
87	Tento ter uma visão positiva de mim mesmo(a).
88	De maneira geral, estou satisfeito(a) comigo mesmo.
89	Eu queria ter mais amor próprio.
90	Tem situações em que me sinto inútil.
91	Acho que não sou bom/boa em nada.

Leia com atenção cada frase e marque se ela é falsa ou verdadeira em relação a você.

92	Às vezes eu me sinto ressentido(a) se as coisas não são do meu jeito.
93	Em algumas ocasiões, eu desisti de fazer algumas coisas porque duvidei da minha habilidade.
94	Houve ocasiões em que eu me rebelei contra pessoas com autoridade, mesmo sabendo que elas estavam certas.
95	Não importa com quem eu esteja falando, eu sou sempre um(a) bom/boa ouvinte.
96	Eu me lembro de ter fingido estar doente para conseguir algo.
97	Houve ocasiões nas quais levei vantagem sobre alguém.
98	Eu estou sempre disposto(a) a admitir quando cometo um erro.
99	Às vezes, eu tento acertar as contas em vez de perdoar e esquecer.
100	Eu sou sempre educado(a), mesmo com pessoas desagradáveis.
101	Eu nunca fiquei aborrecido(a) quando pessoas expressaram ideias muito diferentes das minhas.
102	Houve vezes em que senti inveja da boa sorte dos outros.
103	Eu fico às vezes irritado(a) com pessoas que me pedem favores.
104	Eu nunca disse algo que machucasse os sentimentos de alguém de propósito.

Parte II

Leia as três questões abaixo com atenção e responda a alternativa correta na folha de gabarito:

- 1) Uma raquete e uma bola custam justas R\$1,10 reais. A raquete custa R\$1,00 a mais que a bola. Qual o preço da bola (resposta em centavos)?
- 2) Se 5 máquinas demoram 5 minutos para fazer 5 objetos, quanto tempo demora para 100 máquinas fazerem 100 objetos (resposta em minutos)?
- 3) Parte de um lago está coberto por flores. Todos os dias, as flores se multiplicam de forma que a área do lago coberta dobra de tamanho. São necessárias 6 dias para cobrir o lago inteiro. Quantos dias são necessários para cobrir metade do lago (resposta em dias)?

Código 1: B

Código 2: C

Código 3:

Parte III

Cada figura abaixo segue um padrão de preenchimento, mas em todas elas há um espaço vazio. Escolha, dentre as alternativas expostas abaixo de cada figura a que melhor preenche esse espaço e em seguida marque a opção na folha de gabarito:

1)

2)

3)

4)

Parte IV

Por fim, responda às seguintes questões relacionadas a você, sua família e sua residência escolhendo apenas uma alternativa por questão e assinalando-a na folha de gabarito. Se você morar em mais de uma casa, responda sobre aquela na qual você passa mais tempo. Se houver dúvidas, pergunte ao aplicador:

1	Quantos anos você tem?	11 ou menos	12	13	14	15 ou mais
		A	B	C	D	E
2	Que ano escolar você frequenta?	5 ano	6 ano	7 ano	8 ano	9 ano
		A	B	C	D	E

Código 1: B

Código 2: C

Código 3:

3	Qual seu sexo?			Masculino	Feminino	
				A	B	
4	Como você se considera?	Amarelo	Branco	Indígena	Negro	Pardo
		A	B	C	D	E
5	Você mora com sua mãe ou com alguma mulher responsável por você? Se não mora, marque "NA" na próxima questão.			Sim	Não	
				A	B	
6	Sua mãe ou mulher responsável por você sabe ler e escrever?	Sim	Não	Não sei	NA	
		A	B	C	D	
7	Você mora com seu pai ou com algum homem responsável por você? Se não mora, marque "NA" na próxima questão.			Sim	Não	
				A	B	
8	Seu pai ou homem responsável por você sabe ler e escrever?	Sim	Não	Não sei	NA	
		A	B	C	D	
9	Você tem irmãos? Se não tem, marque "NA" nas duas próximas questões			Sim	Não	
				A	B	
10	Você tem irmãos mais velhos?	Sim	Não	NA		
		A	B	C		
11	Você tem irmãos mais novos?	Sim	Não	NA		
		A	B	C		
12	Quantas pessoas moram na sua casa?	2	3	4	5	6 ou mais
		A	B	C	D	E
13	Possui empregada doméstica ou faxineira na sua casa?			Sim	Não	
				A	B	
14	Quantos carros no total possuem as pessoas que moram com você?	Nenhum	1	2	3 ou mais	
		A	B	C	D	
15	Quantos banheiros tem na sua casa?	Nenhum	1	2	3 ou mais	
		A	B	C	D	
16	Quantos quartos tem na sua casa?	Nenhum	1	2	3 ou mais	
		A	B	C	D	
17	Quantos computadores ou notebooks tem na sua casa?	Nenhum	1	2	3 ou mais	
		A	B	C	D	
18	A rua onde você mora é asfaltada?	Sim	Não	Não sei		
		A	B	C		
19	Tem energia elétrica na sua casa?	Sim	Não	Não sei		
		A	B	C		