

## Entrevista

## “Avaliação de Impacto e de Processo do Programa Criança Feliz – Desafios e aprendizados com a implementação do estudo”. Uma entrevista com Cesar Victora, Professor Emérito de Epidemiologia na Universidade Federal de Pelotas<sup>1</sup>

“Impact and Process Evaluation of the Criança Feliz Program – Challenges and lessons learned from the implementation of the study”. An interview with Cesar Victora, Emeritus Professor of Epidemiology at the Federal University of Pelotas

Eduardo Marino<sup>1\*</sup> 

<sup>1</sup>Manacá Avaliação e Aprendizagem, São Paulo, SP, Brasil

**COMO CITAR:** Marino, Eduardo (2022). “Avaliação de Impacto e de Processo do Programa Criança Feliz – Desafios e aprendizados com a implementação do estudo”. Uma entrevista com Cesar Victora, Professor Emérito de Epidemiologia na Universidade Federal de Pelotas. *Revista Brasileira de Avaliação*, 11(3), e110922. <https://doi.org/10.4322/rbaval202211009>

Eduardo Marino, branco, Consultor Sênior Associado, Manacá Avaliação e Aprendizagem, São Paulo, Brasil; responsável pela elaboração do roteiro, execução e revisão da entrevista.

### Resumo

O Programa Criança Feliz (Brasil, 2017) foi instituído pelo Decreto nº 8.869, de 5 de outubro de 2016, como o objetivo de promover o desenvolvimento integral das crianças na primeira infância, por meio de visitas domiciliares, considerando sua família e seu contexto de vida pelo então Ministério de Desenvolvimento Social e Agrário, atual Ministério da Cidadania. O Programa busca articular ações das políticas de assistência social, saúde, educação, cultura e direitos humanos, tendo como fundamento a Lei nº 13.257, de 8 de março de 2016 – conhecida como Marco Legal da Primeira Infância. O Programa foi instituído com a meta de atender 3 milhões de crianças de 0 a 36 meses beneficiárias do Programa Bolsa Família, além de 640 mil gestantes, com visitas previstas para o terço final da gravidez. Entendeu-se à época da sua estruturação que seria importante uma avaliação que acompanhasse uma amostra significativa de crianças e suas mães ou cuidadores desde o início do Programa, para mensurar o seu impacto. Para empreender o estudo avaliativo, foi convidado pesquisador Cesar Victora, epidemiologista, da Universidade Federal de Pelotas, que constituiu uma equipe de trabalho com representantes de seis outras Universidades Federais, além de um Comitê Técnico a apoiar no desenho da avaliação e acompanhar a implementação do estudo. A equipe optou por avaliar o impacto e a implementação do Programa com o uso das metodologias mais adequadas para cumprir essas duas finalidades. A avaliação teve início em 2018 e foi finalizada no início de 2022. Nesta entrevista com César Victora são explorados os desafios, os fatores facilitadores e as limitações de um estudo avaliativo dessa magnitude.

**Palavras-chave:** Avaliação de impacto. Avaliação de implementação. Parentalidade e desenvolvimento infantil.

### Abstract

The Criança Feliz Program was established by Decree No. 8,869, of October 5, 2016, with the objective of promoting the integral development of children in early childhood, through home visits, considering their family and their life context by the then Ministry of Social and Agrarian Development, currently the Ministry of Citizenship. The Program seeks to articulate social assistance, health, education, culture and human rights policies, based on Law No. 13,257, of March 8, 2016 – known as the Legal Framework for Early Childhood. The Program was established with the goal of serving 3 million children aged 0 to 36

A RBAVAL apoia os esforços relativos à visibilidade dos autores negros na produção científica. Assim, nossas publicações solicitam a autodeclaração de cor/etnia dos autores dos textos para tornar visível tal informação nos artigos.

**Recebido:** Abril 20, 2022

**Aceito:** Abril 20, 2022

**\*Autor correspondente:**

**Eduardo Marino**

**E-mail:** edumarino1@gmail.com

**Instituição parceira:** Fundação Maria Cecília Souto Vidigal



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

<sup>1</sup> Responsável pela elaboração do roteiro, execução e revisão da entrevista.



months who are beneficiaries of the Bolsa Família Program, in addition to 640,000 pregnant women, with visits scheduled for the final third of pregnancy. It was understood at the time of its structuring that it would be important to carry out an evaluation that followed a significant sample of children and their mothers or caregivers since the beginning of the Program, in order to measure its impact. To undertake the evaluative study, researcher Cesar Victora, an epidemiologist, from the Federal University of Pelotas, was invited, who formed a work team with representatives from six other Federal Universities, in addition to a Technical Committee to support the evaluation design and monitor the implementation of the evaluation study. The team chose to evaluate the impact and the implementation of the Program using the most appropriate methodologies to fulfill these two purposes. The evaluation began in 2018 and was completed in early 2022. In this interview with César Victora, the challenges, facilitating factors and limitations of an evaluative study of this magnitude are explored.

**Keywords:** Impact evaluation. Implementation evaluation. Parenting. Early child development.

**Revista: César, a proposta nesta entrevista é conhecermos mais sobre sua experiência na coordenação da avaliação de impacto e implementação do Programa Criança Feliz (PCF), que você liderou desde o início. Quais foram as escolhas metodológicas iniciais que vocês fizeram e por que vocês chegaram nesse desenho de avaliação?**

Cesar Victora: Quando fomos contatados pelo Ministério da Cidadania, que naquela época era o Ministério de Desenvolvimento, o MDS, eu aceitei o convite, mas pedi carta branca. Eu disse: “Nós queremos fazer uma avaliação na qual toda a parte metodológica é definida por Pelotas, pela nossa universidade”, e o ministro aceitou. Então eu gostaria de elencar aqui os aspectos principais do delineamento. O primeiro é que se trata de um estudo experimental incluindo crianças em 30 municípios de seis Estados com muitas crianças na clientela do bolsa família, e, portanto, muitas crianças elegíveis para o Programa Criança Feliz (PCF). Esse tipo de delineamento sorteia quem receberá ou não o programa.

Com uma amostra grande, o delineamento garante que os grupos sejam muito parecidos; e de fato foi o que observamos quando fizemos o estudo de linha de base, ainda em 2018, quando foram sorteados os dois grupos. Por sorte, o processo que chamamos de alocação aleatória resulta em grupos muito similares quando a amostra é grande. Então a ideia era pegar uma amostra grande, 3.000 crianças em 30 municípios, ou seja, uma média de 100 crianças por município. Nesse caso, era ético fazermos alocação aleatória, ou seja, sortear crianças para entrar e outras para não entrar no programa, porque havia mais clientela do que as vagas oferecidas pelo programa. Isso é o que chamamos de sobredemanda, pois os municípios tinham no mínimo quatro crianças para cada vaga.

Sendo assim, fizemos o sorteio de acordo com a loteria federal, o qual foi gravado no Youtube e está disponível até hoje. O processo foi completamente transparente e nós comunicamos imediatamente aos municípios. O sorteio foi baseado no Número de Inscrição Social (NIS) das crianças, no número do cadastro único, para não colocar o nome da criança no sorteio e para garantir sua anonimidade. O segundo aspecto importante é que o estudo foi uma avaliação independente. O ministério foi o apoiador financeiro majoritário da avaliação, mas toda a equipe de pesquisadores foi soberana em tomar as decisões junto com o comitê externo, do qual irei falar um pouco depois. Então fizemos o sorteio, selecionamos as crianças e fizemos um estudo muito ambicioso, estudo de três anos de duração.

O primeiro ano foi 2018 e terminou três anos depois, no início de 2022, passando por quatro fases que inicialmente definimos como T0, T1, T2 e T3. Recentemente acabamos a última fase, a T3, mas o T2 não pôde ser executado presencialmente em função da pandemia. Os estados que fizeram parte da avaliação foram Pará, Ceará, Pernambuco, Bahia, São Paulo e Goiás. Ficamos com três estados no Nordeste, um do Norte, um do Centro-Oeste e um do Sudeste. Nos três estados do Sul é pequena a clientela do Bolsa Família e a maioria dos estados já oferecia programas muito parecidos. No Rio Grande do Sul por exemplo, o Primeira Infância Melhor (PIM) tinha cobertura de quase todas as crianças elegíveis. Então optamos por não escolher um estado do sul, por não encontrar nenhum que fosse adequado para a avaliação.

Quanto à equipe de avaliação, convidamos seis pesquisadores experientes da área de saúde e desenvolvimento infantil, um em cada estado. Com uma única exceção, eram pessoas



com quem já havíamos trabalhado no passado em outros estudos multicêntricos. Criamos uma equipe central aqui na universidade de Pelotas envolvendo diversas especialidades, incluindo epidemiologistas com experiência em avaliação, estatísticos, psicólogos e pediatras. Envolvermos também um economista como consultor externo, pois parte das análises planejadas incluem métodos econométricos para complementar os métodos epidemiológicos que normalmente usamos em estudos de avaliação dentro da área de saúde. Tínhamos ainda um comitê externo de especialistas que nos assessorava em várias ocasiões. Por exemplo, para a randomização, nossa ideia inicial era alocar bairros ou setores censitários do IBGE, onde todas as crianças receberiam ou não a intervenção. Logo em seguida, descobrimos que o banco de dados do Ministério da Cidadania não permitia georreferenciar as famílias; portanto, mudamos o delineamento para alocação individual, sendo que o comitê referendou essa escolha.

Depois tivemos inúmeras discussões sobre os testes de desenvolvimento que iriam ser usados, optamos no primeiro momento usar o teste ASQ3 (ASQ, 2022), um teste amplamente utilizado e validado no Brasil para crianças de zero a seis anos de idade. Começamos a aplicar este teste anualmente, com a proposta de usar o teste Bayley (ScienceDirect, 2022), que é um teste mais complexo na última avaliação agora no final de 2021. No entanto, as crianças atendidas pelo Programa não eram tão precoces como gostaríamos, sendo poucas crianças recrutadas durante a gestação ou logo após o nascimento. Quando chegamos na última etapa, no T3, grande parte das crianças já tinham mais de 42 meses, que é a idade máxima para o teste Bayley. Esta foi a primeira razão de não optarmos por usar o Bayley. A segunda razão foi que o Bayley é um teste cujo treinamento é bem complexo, e com a pandemia nós não conseguiríamos fazer o treinamento presencial das entrevistadoras, pois seria preciso reunir todos em agosto ou setembro de 2021, quando a pandemia ainda estava forte. Então fizemos uma adaptação do ASQ3 e criamos o ASQ3 observado. Por exemplo, não perguntamos para a mãe se a criança sabe chutar uma bola, nós damos uma bola e pedimos para a criança chutar. Outro exemplo: não perguntamos se a criança consegue empilhar três ou quatro cubinhos, mas levamos os cubinhos e pedimos para a criança empilhar. Isso funcionou muito bem.

Nós usamos também o ASQ3 relatado pela mãe, porque às vezes ela diz: “Ele sabe empilhar”, mas a criança é tímida ou não consegue fazer na frente da entrevistadora. Então no fim temos dois testes, o ASQ3 observado, que é mais rigoroso, porque a criança tem que fazer aquela tarefa; e o ASQ3 relatado pela mãe, sendo os dois resultados parecidos, mas com escores cerca de 10% maiores para o teste relatado pela mãe.

Todas essas escolhas foram referendadas pelo comitê dos especialistas. O comitê tem pesquisadores brasileiros do Insper e da Universidade Federal do Rio de Janeiro e representantes do Banco Mundial, do Banco Interamericano do Desenvolvimento e das Nações Unidas, pessoas que são experts em desenvolvimento infantil. O comitê é bem diversificado e que nos ajudou muito. Esses dois exemplos de que falei, das escolhas dos testes e da randomização, foram ocasiões em que foi essencial termos essa contribuição.

**Sobre o financiamento, uma avaliação com essa robustez é uma avaliação que tem um orçamento importante. Vocês contaram com uma combinação de apoiadores, o principal financiador foi o Ministério da Cidadania, a SAGI (Secretaria de Avaliação e Gestão da Informação), era quem estava nesse papel de fazer essa gestão do investimento, além de toda a parte de contratação e acompanhamento dos trabalhos. Mas além do governo federal, vocês contaram com o apoio de duas fundações, a Fundação Maria Cecília Souto Vidigal e Itaú Social, também o próprio PNUD, que foi quem atuou como operador, apoio na operacionalização dos recursos. Como funcionou essa concertação de apoiadores para o estudo?**

Cesar Victora: É exatamente o que você falou, nós fizemos questão de que a equipe de Pelotas, que é a equipe central do estudo, não fosse financiada pelo Ministério da Cidadania porque achamos que isso criaria uma certa dependência e possivelmente afetaria o caráter de avaliação externa do estudo. Nós tivemos a felicidade de contar com o apoio da Fundação Maria Cecília Souto Vidigal e do Itaú Social, que financiaram o estudo de Pelotas via PNUD. Esta



verba financiou a equipe de sete pessoas que trabalharam durante três anos no projeto. Como mencionei, o PNUD agilizou o repasse de verbas das duas fundações para o nosso grupo. Já os estados foram contratados diretamente pelo Ministério da Cidadania através de Termos de Execução Descentralizada. Às vezes foi um pouco problemático, pois tem mais controle por causa da burocracia do governo federal e de auditorias do Tribunal de Contas da União (TCU). O que aconteceu em quase todos os estados é que o Ministério da Cidadania, por meio da SAGI, descentralizou recursos substanciais por meio de fundações ligadas às universidades. Apesar das barreiras burocráticas inerentes à administração federal, contamos com muita boa vontade, tanto da SAGI quanto das fundações e do PNUD. Foi importante também o apoio da Secretaria Nacional de Atenção à Primeira Infância (SNAPI), que implementa o programa via municípios, enquanto a SAGI avalia e monitora. Foi um esforço conjunto e muito frutífero entre as sete universidades e as quatro agências que nos financiaram.

**Agora, além da questão do impacto e da efetividade, vocês contemplaram também uma avaliação de implementação. Conta um pouco como é essa combinação, esse olhar para a implementação e para o impacto, como funcionou? O quanto isso foi um elemento que agregou valor a esse estudo?**

Cesar Victora: Sem dúvida, o nosso grupo tem bastante experiência em avaliações de impacto e sabemos que nunca se deve estudar impacto sem avaliar os processos, porque se não o fizermos é impossível interpretar os resultados. Então, desde o primeiro momento da avaliação, quando planejamos pesquisas de campo, visitas domiciliares e filmagens para avaliar o impacto, o processo de implementação foi avaliado. Nós coletávamos dados dos coordenadores municipais, dos dirigentes do Centro de Referência e Assistência Social (CRAS) e tínhamos uma série de sessões com as visitadoras, incluindo observação de visitas, um teste de conhecimento e grupos focais.

Pedíamos também para elas ranqueassem a sua satisfação com o programa, com o transporte oferecido, o salário, as condições de emprego e contrato etc. Isso nós fizemos em 2019 *in loco*, em 2020 por telefone, devido à pandemia, e em 2021 novamente *in loco*. Em 2020 não conseguimos fazer a observação de visitas, pois a maioria dos estados não estavam visitando, e também resolvemos não fazer a prova de conhecimento, porque por internet fica complicado. Fizemos então a prova em 2019 e 2021 diretamente *in loco*. De posse de todas essas informações, criamos mais de 30 variáveis e realizamos um procedimento de triar essas variáveis e avaliar como elas se correlacionavam entre si, selecionando as 10 variáveis mais informativas. Indicadores como: “Como foi a seleção de visitadoras? Foi por indicação de alguém ou por concurso?”, “Quantos carros tinham à disposição do programa no município conforme o seu tamanho?”, “Quanto era o salário das visitadoras?”, “Quem era o coordenador municipal? Ela já estava há tempo no cargo, com experiência, ou já havia uma alta rotatividade?”, “Qual a nota média das entrevistadoras?”. Usamos também as notas médias das provas e as escalas de satisfação das visitadoras, baseadas em 10 a 30 visitadoras em cada município. Usando uma técnica chamada análise de componentes principais, calculamos um escore de força da implementação para cada município. Isso revelou que alguns municípios implementam melhor que os outros

Como você mencionou, fizemos uma avaliação de efetividade. Este tipo de avaliação observa o que ocorre na realidade. Até agora, a maioria das avaliações existentes mostram que visitas domiciliares melhoram o desenvolvimento das crianças, mas estas são avaliações de eficácia que dispõem de uma equipe de pessoas treinadas por uma universidade ou grupo de pesquisa, as quais, elas mesmas, fazem a intervenção. Portanto, é um programa muito bem implementado, com muitas variáveis controladas, o que não corresponde à vida real.

**É um tipo de estudo quase laboratorial, com muitos controles, não é?**

Cesar Victora: O nosso estudo também é um laboratório, mas sem os controles de qualidade para quem implementa, ou seja, nosso laboratório é a população recebendo serviços de rotina. Por exemplo, perguntamos “Quantos meses de visitas presenciais houve no município?”.



Em três anos, ou 36 meses do estudo, houve municípios com apenas seis meses de visitas, porque não retomaram a visitação após a pandemia, enquanto um município parou apenas durante 2 meses, sendo a média igual a 22 meses dos 36 meses possíveis. Obviamente, esta informação precisa constar em nosso relatório final. Primeiro porque um município é diferente de outro, e segundo porque a pandemia afetou todos, pois não houve nenhum município que não tivesse uma suspensão pelo menos por alguns meses.

Em resumo, a avaliação nos permitiu identificar três grupos de municípios, com implementação forte, média e fraca. Com isto será possível verificar se, nos municípios onde a implementação foi forte, o impacto do programa foi maior do que em municípios onde a implementação foi fraca. Vamos então investigar, dentro da análise aleatorizada, dentro do delineamento experimental, se houve maior impacto do PCF nos municípios onde realmente parece que o programa funcionou melhor.

**Cesar, você já comentou um pouco sobre a pandemia e seu impacto. Eu queria que você trouxesse também quais foram as questões críticas, tanto para o estudo e para o próprio programa, e como é que vocês lidaram com isso: quais alternativas vocês tiveram que encontrar para prosseguir com o estudo?**

Cesar Victora: Realmente preciso confessar que em 40 anos de pesquisa, nunca fiz um estudo tão difícil na minha vida, porque além de toda a dificuldade na grande escala da pesquisa, com 30 municípios em seis estados diferentes e da duração de três anos da pesquisa, a pandemia caiu como uma bomba no meio do nosso estudo. Como mencionei, em média, os municípios ficaram 14 meses sem visita devido à pandemia. Durante este tempo, o Ministério da Cidadania desenvolveu estratégias para que as visitadoras utilizassem o *WhatsApp* e o *Facebook* para enviar mensagens de áudio e vídeos para as famílias. Desta forma muitos municípios mantiveram um contato mesmo durante a pandemia.

Obviamente, imaginamos que esse contato virtual seria menos eficaz que o contato presencial, onde a visitadora está lá e fala com a mãe, vê a criança etc. O Ministério da Cidadania fez o melhor que poderia ter sido feito, em nível central de Brasília. No nível municipal, o que observamos foi muita variabilidade, pois havia municípios que faziam o que o Ministério instruíra, mas outros o faziam em parte ou mesmo não faziam o que era recomendado. Como então nossa equipe não pôde fazer visitas, fizemos seis rodadas de telefonemas. Em cada estado, a universidade que estava envolvida no estudo ligava para as famílias perguntando se tinham recebido visitas ou contatos remotos (mensagens, vídeos etc.) e se haviam recebido livros infantis que a Fundação Itaú Social distribuiu. Também aplicamos alguns testes que eram passíveis de serem feitos por telefone, como o teste de parentalidade PAFAS (Sanders et al., 2014) e o SWYC (Tufts Children’s Hospital, 2022), que são testes muito simples de desenvolvimento infantil, 10 perguntas por faixa etária.

Um grande desafio é que nosso estudo foi planejado para recrutar 3.000 crianças no início, mas sempre recrutamos com uma certa “gordura”. Conseguimos visitar 3.342 crianças na primeira fase em 2019. Quando começamos a telefonar, conseguíamos 70% mais ou menos de contato, porque as pessoas mudam muito de telefone e de operadora, e ao longo do tempo esse número foi caindo, mas nunca caiu abaixo de 50% da amostra original. A grande vantagem é que agora na avaliação final de 2022, entrevistamos 80,0% das crianças originais. Localizar tantas crianças após três anos, para esse tipo de população, com pandemia e tudo, nos deixou muito felizes. Isso comprova o empenho das equipes estaduais em achar efetivamente essas crianças. Em resumo, a pandemia afetou o programa e a avaliação, através do cancelamento das visitas presenciais do PCF e de nossas entrevistas presenciais em 2020. Mesmo assim tivemos entrevistas em 2018, 2019 e 2021, além dos seis telefonemas. Isso gerou um excelente conjunto de dados aí para a análise.



**Esse alcance que vocês tiveram por telefone não foi baixo, dado todos estes desafios que o contexto trouxe, a pandemia, as mudanças frentes do número de telefone das famílias, 50%, 60% de respostas é um número muito bom. Você já comentou alguns dos obstáculos: há outros aspectos que te chamaram a atenção na implementação desse tipo de estudo nesse contexto pandêmico?**

Cesar Victora: Eu acho que obviamente o maior obstáculo foi a pandemia. O segundo obstáculo é que, como é um programa muito descentralizado, nós fizemos a randomização lá no início de 2018, quando fizemos o estudo de linha de base. Em cada contato com as famílias, perguntamos para as mães de crianças do grupo intervenção e do grupo de controle se elas estavam sendo visitadas e com que frequência. Nós observamos duas coisas. Primeiro, que na maioria dos municípios menos de metade das crianças do grupo intervenção visitadas receberam contatos remotos com regularidade. O PCF preconizava visitas semanais, mas as visitas, quando ocorriam, eram quinzenais ou mensais, pois os municípios não davam conta das visitas semanais. Em retrospecto, eu acho que um programa dessa escala com visita semanal é muito ambicioso. Esta foi uma lição que aprendemos, pois não creio que seja factível, pelo menos em um país tão grande assim, sob as condições da nossa rede de serviços públicos.

O segundo aspecto que também nos frustrou foi a aderência dos municípios à randomização por nós proposta. Como eu já comentei, em 2018 fizemos os sorteios das crianças que estariam com grupo que seriam atendidas pelo programa e as que ficariam no grupo controle; os sorteios foram filmados e estão disponíveis no *Youtube*, com a participação de servidores do Ministério da Cidadania, para garantir a lisura do processo. No mesmo momento imprimimos a lista de NIS. Fizemos a randomização de um município por vez. Por exemplo, quando terminou o estudo de linha de base em, digamos, Caucaia, no Ceará, nossos entrevistadores usaram tablets para transmitir os dados das crianças para a equipe central aqui em Pelotas. Realizávamos o sorteio imediatamente e enviávamos a lista ao Município e ao Ministério. Para nossa frustração, nem todos os municípios respeitaram a randomização. Por razões que não sei bem, talvez até políticas, algumas crianças recebiam o programa, outras não, e nosso delineamento original foi afetado pelo que chamamos de contaminação.

Durante as fases subsequentes da pesquisa, observamos que em média 10% das crianças do grupo controle foram visitadas contra 40% do grupo intervenção. Houve, portanto, implementação incompleta no grupo de intervenção, pois todas as crianças deveriam ser visitadas, enquanto 10% do grupo controle foi indevidamente visitado. Isso não invalida nosso estudo, mas prejudica o delineamento original. Estudos randomizados precisam ser analisados primariamente pôr o que chamamos de intenção de tratar; ou seja, se a criança foi sorteada por grupo PCF, ela fica no grupo PCF durante a análise de dados mesmo que não seja visitada todas as vezes, enquanto no grupo de controle, mesmo que seja visitada, a criança permanece no grupo controle. É claro que, se temos aderência de apenas 40% no grupo intervenção e aderência indevida de 10% no grupo controle, as análises por intenção de tratar subestimaram o verdadeiro impacto do programa. Portanto, essas análises devem ser complementadas por outros métodos em que comparamos quem realmente recebeu o programa com quem não recebeu. Esta não é uma análise tão rigorosa do ponto de vista de avaliação, pois quem recebeu o programa pode ser diferente de quem não recebeu, por exemplo, mães mais esclarecidas, que moram mais perto do CRAS, que recebem melhor as visitadoras, das quais as visitadoras gostam mais etc.

Todos estes vieses sugerem que as crianças que efetivamente receberam visitas podem ser diferentes das que não receberam. Estamos fazendo uma série de análises estatísticas para tentar controlar esses motivos, inclusive usando variáveis instrumentais, que é um enfoque bastante utilizado para avaliar programas com implementação incompleta. Estamos fazendo essas análises para ver se quando comparamos quem recebeu com quem não recebeu, aparece um impacto maior do que nas análises originais por intenção de tratar. Aqui temos novamente uma diferença importante com os estudos de eficácia, pois nestes o pesquisador garante que uma criança recebe as visitas e a outra não. A implementação incompleta e a contaminação representam enormes desafios que estamos tentando, se não contornar, porque é impossível contorná-los completamente, pelo menos através de análises adequadas.



**Quais foram os fatores facilitadores? O que ajudou vocês? Só do início já me vem várias coisas, acho que você contou com equipes muito competentes nas universidades, esse esforço de busca mesmo nesse contexto da pandemia, foi muito impressionante esse trabalho das equipes. Que outros elementos você traria como facilitadores?**

Cesar Victora: Eu acho que irei me repetir um pouco, mas na verdade foram as parcerias, entre as setes universidades, o Ministério da Cidadania, tanto a SAGI quanto a SNAPI, a Fundação Maria Cecília e a Fundação Itaú Social e o PNUD. Estas parcerias foram excelentes. Além disso, o financiamento foi plenamente adequado para cobrir o que precisávamos. Por exemplo, precisamos alugar carros, pagar mais de 100 pessoas envolvidas na pesquisa, comprar equipamentos etc. Agora, por exemplo, no final da pesquisa, precisamos comprar dezenas de kits para aplicar o teste ASQ3 observado: bolas para a criança chutar, cubinhos que era para a criança empilhar, papel e lápis para a criança escrever ou desenhar, roupinhas para ver se a criança sabia abotoar, e assim por diante. Então o orçamento foi perfeitamente adequado.

Outro aspecto facilitador foi a experiência dos grupos de pesquisa, tanto em Pelotas quanto nos estados. Os vários treinamentos de entrevistadores foram rigorosos, e a nossa equipe visitou várias vezes os estados para supervisionar a coleta de dados. Os treinamentos eram inicialmente presenciais, reunindo mais de 80 pessoas em Brasília em cada ocasião. Agora, no final de 2021, o treinamento precisou ser virtual, e após seu término observamos que foi necessário avaliar as equipes e retreinar alguns aspectos que ainda não estavam 100%. Assim, acho que entre fatores facilitadores se incluem as parcerias, o financiamento, o rigor no treinamento e na supervisão. E principalmente a competência dos coordenadores estaduais, todos pesquisadores com ampla experiência prévia em estudos de saúde e desenvolvimento infantil.

**Você comentou que, em seus 40 anos de experiência, nunca havia passado por um processo de tamanha complexidade. Como essa é uma revista dialoga com avaliadores e gestores de programas, com aqueles que contratam avaliações, que dicas ou aprendizados você deixaria como mensagens finais para esses grupos?**

Cesar Victora: Há alguns aspectos muito importantes. O primeiro é a necessidade de começar com um modelo conceitual, um modelo de impacto para o sucesso do programa. Precisamos que o município convidado para aderir ao programa aceite o convite, selecione um coordenador competente, selecione os visitantes, que estes sejam treinados rigorosamente, realizem as visitas preconizadas, se comuniquem e interajam bem com as famílias, que as mães ou cuidadoras aprendam os conteúdos ministrados, estimulem as crianças e finalmente que a criança se desenvolva melhor. Este modelo conceitual organiza o delineamento da pesquisa e identifica os tipos de dados a serem coletados.

O segundo aspecto é o planejamento para avaliar a implementação, ou seja, o processo, ao mesmo tempo que o impacto sobre as crianças. Usamos primariamente métodos quantitativos, que são a nossa especialidade como epidemiologistas, mas também os complementamos com métodos qualitativos como grupos focais, e entrevistas semiestruturadas com informantes-chaves nos municípios. O delineamento, portanto, lançou mão de uma mistura de métodos, o que é super importante.

Como mencionei foi também essencial ter parcerias, como tivemos com as universidades e os financiadores. O rigor científico e o delineamento experimental com alocação aleatória foram importantíssimos, mesmo que ao final tenha sido necessário lançar mão de análises não experimentais, como comparações entre quem recebeu ou não as visitas. E finalmente acho que a adaptabilidade foi essencial, pois uma avaliação desta magnitude necessita de correções de trajetória em seu decorrer. Claro que nunca esperávamos uma pandemia, mas é preciso estar preparado para os imprevistos e enfrentá-los.

Outro aspecto fundamental é a independência dos avaliadores, que têm autonomia para analisar os dados e publicá-los, sem ingerências ou pressões externas, o que ficou harmonicamente estabelecido com nossos parceiros do Ministério desde que o convite inicial



foi feito para o nosso grupo. Divulgar e compartilhar os resultados é um bem maior, um *public good*, pois estes podem e devem ajudar a melhorar programas no futuro, não apenas no Brasil mas também em outros lugares. Nós acreditamos, como disse Ricardo Paes de Barros, que é um grande especialista no assunto, que o nosso foi o maior estudo já feito sobre estimulação infantil no mundo, então acreditamos que tem muito a oferecer em termos das lições que nós aprendemos no manejo desse estudo tão complexo.

**Entrevistador: Esse é um ponto importante, porque há muita referência à estudos que foram feitos na década de 60, por exemplo, estudos até longitudinais, mas foram estudos muito reduzidos com amostras, o Perry Preschool (Schweinhart, 2022) por exemplo, teve uma amostra que, que com apenas 123 crianças, um estudo com 3.000 realmente tem muita robustez, muitos aprendizados. César, quero te agradecer muito por essa rica contribuição para a Revista Brasileira de Avaliação.**

### Fonte de financiamento

A pesquisa abordada na entrevista foi financiada pela Fundação Maria Cecília Souto Vidigal e contou com outros níveis de colaboração institucional, como relatado.

### Conflito de interesse

Não há

### Agradecimentos

Agradeço ao Professor Cesar Victora pela generosidade em compartilhar sua experiência e conhecimento.

### Referências

Ages & Stages Questionnaires – ASQ. (2022). How ASQ Works. Paul H. Brookes Publishing Co., Inc. Recuperado em 30 de março de 2022, de <https://agesandstages.com/about-asq/how-asq-works/>

Brasil. Ministério do Desenvolvimento Social e Agrário – MSDA. (2017). *Programa Criança Feliz. A Intersetorialidade na visita domiciliar*. Brasília: MSDA. Recuperado em 30 de março de 2022, de [https://www.mds.gov.br/webarquivos/publicacao/crianca\\_feliz/A\\_intersetorialidade\\_na\\_visita\\_domiciliar\\_2.pdf](https://www.mds.gov.br/webarquivos/publicacao/crianca_feliz/A_intersetorialidade_na_visita_domiciliar_2.pdf)

Sanders, Matthew R., Morawska, Alina, Haslam, Divna M., Filus, Ania, & Fletcher, Renee. (2014). Parenting and Family Adjustment Scales (PAFAS): Validation of a brief parent-report measure for use in assessment of parenting skills and family relationships. *Child Psychiatry and Human Development*, 45(3), 255-272. <http://dx.doi.org/10.1007/s10578-013-0397-3>. PMID:23955254.

Schweinhart, Lawrence J. (2022). *The High/Scope Perry Preschool Study Through Age 40: Summary, Conclusions, and Frequently Asked Questions*. HighScope. Recuperado em 30 de março de 2022, de <https://image.highscope.org/wp-content/uploads/2018/11/16053615/perry-preschool-summary-40.pdf>

ScienceDirect. (2022). Bayley Scales of Infant Development. Recuperado em 30 de março de 2022, de <https://www.sciencedirect.com/topics/medicine-and-dentistry/bayley-scales-of-infant-development>

Tufts Children’s Hospital. (2022). The Survey of Well-being of Young Children. Recuperado em 30 de março de 2022, de <https://www.tuftschildrenshospital.org/the-survey-of-wellbeing-of-young-children/overview>