




Algumas dimensões ainda atuais da análise do comportamento aplicada

Some still-current dimensions of applied behavior analysis

 DONALD M. BAER¹
 MONTROSE M. WOLF¹
 TODD R. RISLEY²

¹UNIVERSITY OF KANSAS

²UNIVERSITY OF ALASKA

Resumo

Vinte anos atrás, uma nota antropológica descreveu as dimensões vigentes da análise do comportamento aplicada como prescrita e praticada em 1968: ela era, ou deveria se tornar, aplicada, comportamental, analítica, tecnológica, conceitual, eficaz e capaz de resultados apropriadamente generalizados. Uma nota antropológica semelhante, hoje, ainda encontra as mesmas dimensões prescritivas e, em maior medida, descritivas. Várias estratégias novas tornaram-se conhecidas; algumas no domínio da análise conceitual, algumas relacionadas ao *status* sociológico da disciplina e algumas sobre sua compreensão da natureza sistêmica necessária para qualquer disciplina aplicada que operará no domínio de comportamentos humanos importantes.

Palavras-chave: aplicação, disseminação, tecnologia, terminologia, história.

Abstract

Twenty years ago, an anthropological note described the current dimensions of applied behavior analysis as it was prescribed and practiced in 1968: It was, or ought to become, applied, behavioral, analytic, technological, conceptual, effective, and capable of appropriately generalized outcomes. A similar anthropological note today finds the same dimensions still prescriptive, and to an increasing extent, descriptive. Several new tactics have become evident, however, some in the realm of conceptual analysis, some in the sociological status of the discipline, and some in its understanding of the necessary systemic nature of any applied discipline that is to operate in the domain of important human behaviors.

Keywords: application, dissemination, technology, terminology, history.

Nota: referência do texto original, cuja permissão de tradução foi garantida pela editora Wiley: Baer, D. M., Wolf, M. M. & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 20(4), 313-327. <https://doi.org/10.1901/jaba.1987.20-313>.

Tradutores  João Eduardo Cattani Vilares e  Marcos Spector Azoubel.

✉ jeducattani@gmail.com

DOI: [HTTP://DX.DOI.ORG/10.18542/REBAC.V19I1.14945](http://dx.doi.org/10.18542/REBAC.V19I1.14945)

Vinte anos atrás, uma descrição antropológica do grupo que chama sua cultura de Análise do Comportamento Aplicada (ABA)¹ só podia começar pela descrição do contexto relevante (Baer et al., 1968): a existência e o poder da matriz disciplinar (Kuhn, 1970, p. 175) dentro da qual o comportamento de indivíduos era analisado experimentalmente. Aquela matriz era ela mesma o comportamento característico de um grupo mais inclusivo, mais antigo, que chamava sua cultura de Análise Experimental do Comportamento (AEC). As estratégias exemplares características (Kuhn, 1970, p. 189) do grupo AEC eram sua ênfase procedimental nas contingências de reforçamento, punição e discriminação de estímulo como variáveis ambientais analítico-comportamentais, sua confiança em delineamentos de sujeito único como os formatos de análise e comprovação e seu uso da caixa de Skinner como sua arena. Dentro de uma década, o grupo ABA excedeu em número o grupo AEC que originalmente o incluía, de tal forma que os debates inevitáveis sobre sua coesão e separabilidade conceituais atuais e desejáveis assumiram certa urgência sociológica (mas sem muita urgência, especialmente do ponto de vista analítico não comportamental. Em um congresso sobre análise do comportamento, um argumento de estado estacionário levou Nancy Datan a citar um provérbio árabe sobre a natureza do conflito: “Eu devo defender minha tribo contra o mundo, minha família contra a tribo, meus irmãos contra minha família e a mim contra meus irmãos”) (e.g., Baer, 1981; Deitz, 1978; Michael, 1980; Pierce & Epling, 1980). Ainda assim, o subgrupo ABA continuou a mostrar praticamente as mesmas estratégias que caracterizavam o grupo AEC, mas deslocou suas aplicações exclusivamente para o que o grupo chamava de “comportamentos socialmente relevantes” nas caixas de Skinner da vida real desses comportamentos.

Esse deslocamento de estratégia demandou um número considerável de novas táticas. No nível mais básico da matriz disciplinar, procedimentos de medição foram quase imediatamente abertos a medidas de resposta baseadas em intervalos contínuos; o registro não de ocorrências de respostas discretas, mas, em vez disso, de intervalos durante os quais a resposta ocorreu ao menos uma vez e permaneceu ocorrendo solucionou os problemas de outro modo não solucionáveis para registrar comportamentos da vida real com inícios e términos difíceis de definir (Baer, 1986, mas cf. Powell, 1984). Essas medições iniciais suplementaram e logo praticamente substituíram as medidas de taxa de resposta características do grupo AEC. Além disso, sete rótulos de classes de táticas foram propostos como controle de estímulo para condutas analítico-comportamentais apropriadas ao novo mundo de aplicação (dentro do qual a lógica analítico-comportamento é realmente difícil de defender – ironicamente, a parte do mundo que gosta de se considerar “real” usualmente prefere explicações mentalistas de seu próprio comportamento).

O controle de estímulo proposto para conduta analítico-comportamental no mundo da aplicação foram sete palavras-chave que sempre deveriam caracterizar um conjunto de injunções: *aplicada*, *comportamental*, *analítica*, *tecnológica*, *conceitual*, *efetiva* e *capaz* de resultados apropriadamente *generalizados*.

Atualmente, aqueles rótulos de táticas permanecem funcionais; eles ainda conotam as dimensões atuais do trabalho normalmente chamado análise do comportamento aplicada. Apesar disso, as táticas para as quais eles atuam como controles de estímulo mudaram em certa medida. (Se elas não tivessem mudado em alguma medida em duas décadas, poderíamos nos preocupar com a viabilidade de sua disciplina; se tivessem mudado demais, poderíamos nos perguntar se havia qualquer disciplina em sua viabilidade. Assim, faríamos bem em estimar, frequentemente, quão bem situados estamos entre esses dois extremos.)

Aplicada

Inicialmente, o significado de *aplicada* centrou-se em conceitos vagos de problemas sociais, interesse social e na importância imediata do comportamento ou de seus estímulos funcionais para quem se comporta. Vinte anos de experiência, especialmente com o que se costuma chamar de crítica social, começaram a esclarecer o que são problemas, interesses e importância social. À primeira vista, eles são apenas os comportamentos que atrapalham a pessoa chamada de sujeito ou cliente; mas mais frequentemente são também comportamentos de outras pessoas que não a chamada de cliente ou sujeito. Problemas sociais são aqueles comportamentos do sujeito ou cliente que resultam em oposição, às vezes pelo cliente, mas mais frequentemente por não clientes, suficiente para gerar algo chamado de uma solução ou, ao menos, um programa. (No mundo da aplicação, programas atrativos que não resolvem o problema ao qual eles são ostensivamente aplicados às vezes ainda assim são valiosos. Ao menos, eles resolvem o problema por vezes mais aversivo de não fazer nada com relação a esse problema. Além disso, muitas vezes eles resolvem um problema associado bastante importante: eles permitem que o cliente ou os não clientes

¹ Nota dos tradutores: optamos por usar essa sigla, derivada do inglês, por conta de seu uso difundido no país.

contrariados discutam o problema com um amigo compreensivo, fornecem a essas pessoas uma plataforma ou ambos. Talvez não exista algo que possa ser considerado um programa totalmente inefetivo. Mas quando programas não resolvem o problema-alvo, é típico – e funcional – não medir sua inefetividade para o que foram propostos, ainda assim pode ser esclarecedor medir sua validade social.)

Nesse sentido, problemas sociais são essencialmente os comportamentos de apresentar ou descrever problemas – de si próprio ou de outra pessoa. A apresentação de problemas se dá algumas vezes em larga escala, outras vezes em pequena escala. A apresentação em menor escala talvez seja aquela em que um cliente descreve um problema pessoal para um terapeuta; a questão é se o cliente consegue descrever bem o suficiente para assegurar a tentativa do terapeuta de solucioná-lo. Em comparação, algumas vezes uma sociedade inteira pode prever a aniquilação nuclear e o analfabetismo tecnológico; a questão nesses casos é se sua mídia consegue apresentar e descrever esse problema de forma suficientemente eficaz para assegurar o comportamento político que gerará a tentativa governamental de solucionar tal problema ou se o governo tentará resolver outro problema menor exatamente porque os que reclamam do problema pequeno são mais eficazes ao usar a mídia, fazer *lobby* ou apoiar campanhas financeiramente.

É evidente que a resposta do terapeuta é usualmente controlada não apenas pela promessa do cliente de pagamento de honorários, mas também pela concordância do terapeuta de que determinado problema merece uma solução – uma concordância por vezes implícita. Assim, a maioria dos terapeutas consideraria aplicar um programa de ensino autoinstrucional com o objetivo de melhorar a habilidade de arremesso de dardos de um cliente para os eventos sociais em seu bar preferido, mas não para melhorar a precisão de tiro de rifle do cliente para um assassinato que pretende cometer. De forma similar, a decisão do governo pode (nós esperamos) ser controlada não apenas pelo que irá reelegê-lo e pelos reforçadores que dão suporte a isso, mas também pela sua análise da sobrevivência e da prosperidade de sua sociedade.

As polaridades dessas duas decisões parecem ser, respectivamente, a apresentação do problema pelo cliente e sua disposição de pagar *versus* os valores do terapeuta (em outras palavras, as contingências históricas e atuais que controlam a concordância do terapeuta em programar as mudanças comportamentais necessárias) e a apresentação do problema pelos lobistas e sua disposição para apoiar campanhas *versus* a análise governamental da sobrevivência e da prosperidade da sociedade (em outras palavras, o histórico e as contingências atuais que controlam a concordância do governo em programar as mudanças comportamentais necessárias). Essas polaridades não mudaram muito em duas décadas (nem em dois milênios); o que é implacavelmente novo na disciplina é o reconhecimento explícito mais difundido de que todas essas polaridades são, elas mesmas, comportamentos de apresentadores e de espectadores; que os comportamentos de apresentar e de descrever problemas sempre existem num contínuo de suas efetividades para determinado espectador; e que sempre que uma agência apresente e descreva seus problemas efetivamente, sua efetividade pode fazer com que outra agência apresente esse problema como *seu* com a mesma efetividade – talvez até mais efetivamente –, e assim por diante, *ad infinitum*.

As últimas duas décadas não produziram uma análise pública melhor que essa sobre a efetividade da apresentação e da descrição de problemas (ainda que sua prática deliberada seja certamente uma das profissões mais antigas do mundo). Na melhor das hipóteses, elas nos mostraram que precisamos de análises: (a) da apresentação e da descrição de problemas de modo a atingir um uso efetivo da mídia, (b) do controle do comportamento das pessoas que podem funcionar como eleitores dos tomadores de decisões (i.e., fazer *lobby*), (c) da obtenção ou da possibilidade de obtenção de apoio de campanha, e (d) do reconhecimento de eventos chamados de *crises* como eventos ambientais nos quais esses repertórios serão mais efetivos. Essas análises são necessárias ao menos para compreender completamente o que nós mais frequentemente queremos dizer com *aplicada*. Nós nos referimos a toda forma de contracontrole que está tipicamente sob controle de estímulos de apresentação e de descrição de problemas. Isso nos deixa com uma questão programática muito abrangente: o que nós sabemos e o que podemos aprender sobre controle de estímulos efetivos que possa ser aplicado no domínio de apresentação de problemas? É nítido que algumas pessoas na nossa sociedade sabem bastante sobre isso. Se elas sabem, então nós podemos aprender. O comportamento crucial deve ser estabelecer a prioridade dessa área de pesquisa como essencial para nos tornarmos uma disciplina aplicada; claramente, as últimas décadas destacaram essa prioridade com uma urgência crescente.

Comportamental

Uma marca do sucesso da análise do comportamento aplicada nas duas últimas décadas é que seus praticantes, pesquisadores e teóricos se depararam com muitas oportunidades para se tornarem outra coisa em vez

de comportamental, geralmente na forma de se tornarem algo “mais” do que comportamental. Em particular, sua integração ocasional com terapia comportamental, educação, psicologia do desenvolvimento, psicolinguística e sociobiologia possibilitou que se envolvessem com construtos de ansiedade, atenção, inteligência, deficiências, espontaneidade, prontidão, períodos críticos, liberadores inatos, mecanismos de armazenamento e recuperação, esquemas e coisas similares. Alguns analistas do comportamento se envolveram com um ou mais desses construtos o suficiente para não serem mais comportamentais; outros simplesmente se inspiraram com esses construtos. A tarefa mais frutífera, no entanto, é reconhecer que cada um desses rótulos (e muitos outros como eles) frequentemente representa uma realidade comportamental ainda não analisada como tal. O ponto é que essas realidades comportamentais provavelmente não serão analisadas como tais dentro de suas disciplinas originais e, portanto, nunca se tornarão verdadeiramente aplicáveis lá, mas podem muito bem ser analisadas do ponto de vista analítico-comportamental, talvez com grande benefício para nós e para essas disciplinas e, portanto, para nossos papéis dentro dessas disciplinas.

Fazer isso não comprometerá nossa habilidade de discriminar uma disciplina comportamental de outra não comportamental: os vários padrões de comportamento profissional que constituem uma disciplina comportamental, minuciosamente descritos e analisados por Zuriff (1985), sempre podem ser discriminados dos consideravelmente mais variados padrões que constituem disciplinas não comportamentais, mesmo que ninguém mais exiba esses padrões de comportamento. (Em outras palavras, a análise de Zuriff é essencialmente filosófica em vez de antropológica.) Parece claro, no entanto, que o behaviorismo será uma abordagem minoritária, pelo menos para um futuro próximo desta cultura. Realmente, os livros didáticos comportamentais que explicam o princípio de Premack poderiam incluir em suas listas de reforçadores culturais o acesso ao uso e ao consumo de explicações internas, mentalistas, para o comportamento. Talvez a linguagem analítico-comportamental seja a solução para isso. Os últimos 20 anos nos mostraram várias vezes que nossas audiências respondem muito negativamente às explicações sistemáticas de nossos programas e a suas suposições subjacentes, mas muito positivamente aos programas como um todo – seus procedimentos e resultados –, desde que sejam deixados “inexplicados” por nós.

Hineline (1980) começou a analisar como nossa linguagem técnica afeta nossas audiências e como elas usam sua própria linguagem não técnica para explicar o comportamento. Às vezes, por exemplo, certos contextos evocam uma atribuição do comportamento a causas ambientais, mas mesmo esse tipo de atribuição e seu controle contextual podem, por sua vez, ser atribuídos a causas de uma “personalidade” interna e, em uma cultura linguística como a nossa, eles usualmente são (ver Hineline, 1980, p. 84). Talvez a análise do comportamento aplicada devesse considerar muito mais cuidadosamente e mais explicitamente as opções de linguagem que podem maximizar sua efetividade em nossa cultura: (a) encontrar caminhos de ensinar a nossa cultura a falar de uma forma analítico-comportamental (ou pelo menos a valorizar a linguagem analítico-comportamental); (b) desenvolver uma fala não analítico-comportamental para apresentação pública e ver se essa fala vai se provar tão útil para pesquisa e análise quanto a atual fala analítico-comportamental ou se as duas linguagens precisam ser mantidas; ou (c) deixar estar (nós representamos cerca de 2% da psicologia americana e estamos estáveis nesse patamar).

Parte do sucesso da análise do comportamento aplicada levou à sua aplicação exploratória nos consultórios. Nesses contextos, a observação direta do comportamento frequentemente é impraticável, então os profissionais recorrem a formas mais duvidosas de observação, por exemplo, autorrelatos e avaliações feitos por participantes-observadores, realizados frequentemente na forma de respostas a questionários, inventários, *checklists*, entrevistas, diários e afins. Ao lançar mão dessas medidas, considera-se mais seguro usar várias delas ao mesmo tempo (ver *Behavioral Assessment*², a partir de 1979). A tese de que um comportamento pode ser uma medida de outro comportamento parece comportamental, em um primeiro olhar; e sob análise, ela parece comportamental, mas é extraordinariamente arriscada, pois depende fortemente da escolha do “outro” comportamento.

Vinte anos de prática deram à análise do comportamento aplicada algo próximo a um método padrão de medição: a observação direta e o registro de comportamentos-alvo de um sujeito por um observador sob o controle de estímulo de uma categorização comportamental escrita. Obviamente, trata-se da medição de algum comportamento de uma pessoa por outro comportamento de outra pessoa. A força desse método específico está na capacidade de modificação do comportamento do observador por meio do treino direto e cuidadoso e na acessibilidade do comportamento do observador a avaliações de confiabilidade diretas e frequentes. Em especial quando essas avaliações de confiabilidade comparam o observador com o escritor da categorização, elas alcançam a

² Nota dos tradutores: o periódico se chama, atualmente, *Journal of Psychopathology and Behavioral Assessment* (<https://www.springer.com/journal/10862>).

validação necessária a qualquer análise comportamental baseada na observação: permitem a revisão empírica da categorização e, portanto, do controle de estímulos que essa categorização exerce sobre o comportamento de observar e registrar do observador, até que esse comportamento satisfaça o criador da categorização. Esse tipo de revisão é realizado por meio da reelaboração da categorização e da reaplicação do treino do observador até que os registros do observador concordem precisamente com aqueles do criador da categorização observando a mesma amostra do comportamento do sujeito. Assim, o criador da categorização *controla* o comportamento de observar e registrar do observador e, por princípio, pode avaliar e refinar esse controle com a frequência que o problema exigir. Nesse sentido, o criador da categorização é (ou deveria ser) a pessoa que define o comportamento do sujeito como sendo um problema (ou é responsável por essa pessoa), então garantir ao criador da categorização que todos e apenas os comportamentos corretos estão sendo registrados é a *única* abordagem para medição válida que faz sentido sistemático: medição válida é a medição daquele comportamento que fez com que o apresentador do problema o apresentasse (Baer, 1986). Claramente, esse argumento não se altera se o observador for substituído por um instrumento de registro. Esse é um argumento forte contra o uso de categorizações padronizadas: é pouco provável que uma categorização padronizada escrita com base na queixa sobre o comportamento de outra pessoa possa satisfazer tão bem quem agora se queixa quanto a categorização que esta pessoa que se queixa escreveria sobre a pessoa específica de quem ele se queixa.

Por outro lado, é arriscado assumir que o autorrelato do sujeito ou a avaliação de um observador-participante do comportamento-alvo do sujeito apresentaria uma confiabilidade semelhante à sua observação direta por um observador sob controle do comportamento do criador da categorização. O comportamento do observador pode ser controlado de uma maneira bem-entendida; os autorrelatos do sujeito e as avaliações do observador-participante usualmente não estão sob controle do profissional-pesquisador. Em princípio, os autorrelatos do sujeito e as avaliações do observador-participante podem ser controlados da mesma forma que uma observação e um registro de comportamentos padronizados do observador, mas sabemos relativamente pouco sobre como fazer isso e, embora muitas vezes possamos manter um controle quase exclusivo do comportamento relevante do observador, raramente podemos nos aproximar dessa exclusividade com o comportamento do sujeito ou de um observador-participante.

É evidente que autorrelatos do sujeito e avaliações do observador-participante podem ser estudados por si próprios como comportamentos a serem analisados, e não como substitutos para a observação direta do comportamento-alvo. Suas análises quase certamente produziram um conhecimento interessante do amplo universo do comportamento verbal e do pequeno universo da rotina profissional, mas, exceto por isso, elas pouco frequentemente pareceriam ter importância aplicada para além de documentar a já fortemente suspeitada invalidade de tais comportamentos como substitutos do comportamento-alvo. Dentro desse pequeno universo da rotina profissional, no entanto, é importante notar que o uso de tais medições – frequentemente chamadas de psicometria nas ciências sociais – tem certa validade social, especialmente para bolsas de pesquisas aplicadas: a inclusão da psicometria convencional em uma proposta de pesquisa aumenta a probabilidade de essa proposta ser aprovada e financiada quando os revisores da proposta não são analistas do comportamento (o que acontece quase sempre). É fato que qualquer escolha de Psicometria₁ *versus* Psicometria₂ inevitavelmente atrairá pelo menos algumas críticas de revisores, mas pelo menos será uma crítica no contexto da aprovação por jogar o jogo correto. Isso pode ser considerado de importância aplicada.

Portanto, a análise do comportamento aplicada costuma ser, e na maioria das vezes deve ser, o estudo do comportamento de um observador que foi colocado sob controle estrito do comportamento do sujeito. Às vezes, isso é exatamente o que se denota como avaliação comportamental. Mais frequentemente, diz respeito tanto ao estudo de como os sujeitos falam de seu próprio comportamento quanto de como outras pessoas falam do comportamento do sujeito, um tipo de fala que geralmente está sob um controle complexo, variado e altamente desconhecido, em que apenas um de seus componentes *pode* ser o comportamento-alvo do sujeito (vagamente chamado de verdadeiro).

Às vezes, apesar disso (e de forma crescente nas duas últimas décadas), a avaliação comportamental usa as formas de psicometria que são mais bem descritas como *amostras* do repertório-alvo, com destaque para os testes de QI e de desempenho. Os problemas com esses testes são muito parecidos com os dos autorrelatos e das avaliações de observadores-participantes: raramente sabemos se o contexto de teste controla aqueles comportamentos de maneira distinta do controle exercido no dia a dia e raramente sabemos se essas amostras são representativas do repertório desejado. O único jeito de se ter alguma certeza sobre esses fatos seria novamente recorrer à observação direta, mas esses testes representam (ou falham em representar) repertórios por vezes muito

abrangentes para permitir observação direta. Assim, eles são usados muitas vezes como a única alternativa praticável, apesar de suas incertezas (e às vezes são usados porque ainda possuem grande validade social nesta sociedade).

Essa tática não é uma novidade na análise do comportamento aplicada: na análise de acidentes, por exemplo, dificilmente podemos lidar com comportamentos acidentais diretamente, porque eles são muito infrequentes, então mudamos os comportamentos muito mais frequentes que supomos ser precursores de acidentes – nós não analisamos os acidentes, mas os comportamentos de risco. De modo similar, na análise da delinquência, nós dificilmente podemos alterar diretamente os atos delinquentes, novamente porque são pouco frequentes e porque costumam ser feitos de forma relativamente secreta, então também não mudamos esses atos, mas o que nós supomos serem os seus precursores nas várias esferas do controle social. Se as suposições implícitas nessas áreas de pesquisa não as desqualificam como exemplos de análise do comportamento aplicada, então a suposição análoga implícita no uso de, digamos, testes de desempenho também não precisam desqualificá-los automaticamente.

A questão aplicada levantada na maioria das vezes pode ser se as incertezas inerentes ao recorrer a tais medições são favoráveis ao *status quo* do conhecimento em cada área problemática, e a resposta, como a maioria das respostas, provavelmente estará sob controle contextual: às vezes, a incerteza é favorável ao *status quo*, às vezes não é.

Como resultado, o termo “avaliação comportamental”, uma nova categoria de eventos das duas últimas décadas, algumas vezes descreve táticas bastante pragmáticas e outras vezes, as táticas de medição menos válidas das disciplinas pseudocomportamentais muito antigas contra as quais a análise do comportamento se rebelou. Claramente, suas táticas podem incluir conceitos, técnicas e medidas de confiabilidade excepcionalmente elegantes e sofisticados (ver Cronbach et al., 1972) que algumas vezes são aplicáveis a uma observação direta (Hartmann, 1977); mas quando essas táticas medem o que desejamos analisar é problemático tanto de maneira analítica quanto pragmática. Em última análise, saber quando o fazem e quando não o fazem exigirá estudos muito difíceis com base na observação direta.

Analítica e Conceitual

Há 20 anos, *analítica* significava um delineamento experimental convincente e *conceitual* significava relevância para uma teoria ampla do comportamento. Os dois tópicos poderiam ser e eram frequentemente discutidos separadamente. Desde então, tornou-se crescentemente aversivo manter essa separação. Agora, a análise do comportamento aplicada costuma ser considerada uma disciplina analítica somente quando demonstra convincentemente como fazer mudanças comportamentais específicas e quando o seu método de mudança comportamental faz sentido sistemática e conceitualmente. Nos últimos 20 anos, nós às vezes demonstramos convincentemente que modificamos comportamentos como especificamos, mas por meio de métodos que não faziam sentido sistemático e conceitual – não estava claro *por que* esses métodos funcionavam. Esses casos nos permitem ver que algumas vezes nós fomos convincentemente aplicados e comportamentais, ainda assim, não suficientemente analíticos. De modo similar, algumas vezes modificamos comportamentos sem sequer termos uma demonstração convincente de como fizemos aquilo e, então, não sabíamos se nossos métodos faziam sentido sistemático e conceitual porque nós não sabíamos claramente qual haviam sido os métodos responsáveis; esses casos nos permitem ver como não ser uma disciplina, muito menos uma disciplina aplicada, comportamental ou analítica.

Agora, a teoria que define o sentido sistemático e conceitual para nós foi levada a ser não apenas sobre o comportamento, mas também sobre o comportamento de mudar comportamento: com maior frequência agora, nós podemos ver a nós mesmos como sujeitos de outra pessoa, não somente como um Experimentador (relembrar a discussão sobre contracontrole em *Aplicada*). Isso se encaixa bem com o aparente contextualismo que emerge gradualmente na análise do comportamento não aplicada. Uma avaliação adequada do contexto sempre implica que nós não meramente o estudamos ou manipulamos, mas também somos parte dele e, portanto, somos controlados por ele, até mesmo no nosso estudo e na manipulação dele.

A apreciação emergente do contexto como os eventos ambientais que devem ser mais bem compreendidos e manipulados em aplicações realmente efetivas flui facilmente da abordagem de campo de Kantor para o estudo do comportamento (Morris, 1982). Mas também flui com a mesma facilidade dos nossos recentes e crescentes conhecimento e manipulação do controle de estímulos e do controle de estímulos condicional (ver Sidman, 1986; e o número especial do Volume 6 de *Analysis and Intervention in Developmental Disabilities*, 1986). Esse desenvolvimento sugere fortemente que nós raramente encontraremos uma instância no controle de estímulos não modificada drasticamente por alguns (talvez por muitos) controles de estímulos condicionais. A relevância dessa tese

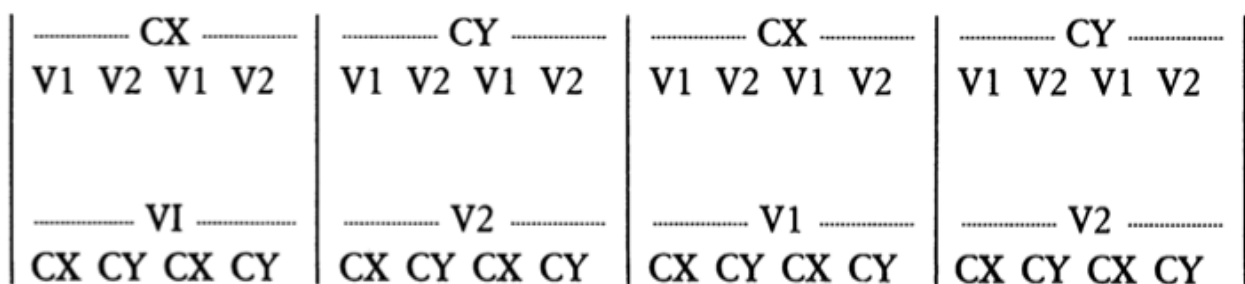
para a aplicação é urgente: ela começa a análise de generalidade da efetividade de cada intervenção, na medida em que nos incita a buscar as condições contextuais sob as quais a intervenção tem eficácia máxima e mínima.

Assim, a primeira lição aplicada do contextualismo é que sempre haverá tais condições; a segunda é que muitas delas devem ser esclarecidas como eventos de estímulo e resposta, porque isso raramente é autoevidente (cf. discussão de Wahler e Fox [1982] sobre a “insularidade” como uma condição limitante no treinamento dos pais); a terceira, mais difícil, mas mais pragmática, é que esclarecer os controles contextuais não é suficiente: se quisermos intervenções amplamente eficazes, teremos que manipular esses controles contextuais; em vez de parar com a avaliação de suas funções como fatores limitantes, teremos que aprender como programar contornando-as ou programando sua remoção.

O contextualismo também implica uma certa classe de delineamentos experimentais. As afirmações contextuais mais simples são da forma, o Comportamento B é controlado diferencialmente pela Variável V no Contexto 1 e no Contexto 2. Para observar isso de forma confiável, precisamos do controle experimental de pelo menos dois níveis da Variável V, digamos V1 e V2; e precisamos de controle experimental de pelo menos dois contextos de interesse, digamos, Contexto X e Contexto Y. A partir disso, precisamos ver como o Comportamento B se relaciona com V1 e V2 no Contexto X, e precisamos ver se esse controle é confiável. Em seguida, precisamos ver como o Comportamento B se relaciona com V1 e V2 no Contexto Y e precisamos ver se esse controle é confiável. Finalmente, precisamos ver essas duas relações (como B se relaciona com V1 e V2 no Contexto X e como B se relaciona com V1 e V2 no Contexto Y) repetidas vezes para que possamos ver se a diferença entre o que os Contextos X e Y fazem em como V1 e V2 controlam B é uma diferença confiável. Os delineamentos de reversão mais simples seriam parecidos com os dois a seguir (consulte diagrama), em que CX e CY são os Contextos X e Y. Ambos são delineamentos mínimos para o problema, ainda assim, cada um contém 16 condições para examinar a linha de base vigorando no comportamento em estudo. Seria melhor que o ritmo do delineamento fosse bastante rápido, sugerindo aquela variante do delineamento de reversão frequentemente chamada de delineamento de multielementos (e.g., Ulman & Sulzer-Azaroff, 1975). Delineamentos como esses podem ser encontrados na literatura da área, mas não com frequência. Na medida em que a análise do comportamento aplicada irá analisar, em vez de avaliar, a generalidade de suas intervenções, esses delineamentos e outros capazes do mesmo tipo de demonstração se mostrarão essenciais.

Figura 1.

Diagrama que representa as relações que precisariam ser avaliadas, de forma a esclarecer como os diferentes Contextos (CX e CY) influenciam o controle das Variáveis (V1 e V2) sobre o Comportamento B³



Nos últimos 20 anos, houve um desenvolvimento considerável de delineamentos de pesquisa. Nos primórdios, foi suficiente rotular apenas a reversão e aqueles com linha de base múltipla: o exame de um comportamento em condições experimentais repetidas e o exame de muitos comportamentos, às vezes com alguns em uma mesma condição experimental enquanto outros estão em uma condição experimental diferente. Essas são as duas estratégias de análise fundamentais, é claro; sua lógica é vista nos esquemas múltiplos, mistos e simultâneos que tantas vezes serviram como delineamentos experimentais na AEC. Naquele espaço, os esquemas tinham nomes, mas os delineamentos não: os pesquisadores simplesmente organizaram aquelas condições (frequentemente, esquemas) necessárias para responder às suas perguntas experimentais, refinaram as condições tão analiticamente quanto seu conhecimento de possíveis variáveis estranhas importantes permitia e fizeram tudo isso tão frequentemente quanto sua convicção exigia. O valor de seus delineamentos não estava em quaisquer nomes de

³ Nota dos tradutores: adicionamos um título à figura com o propósito de facilitar sua compreensão.

categorias que pudessem ser impostos a eles, mas na relação entre as condições que haviam arranjado e a questão proposta.

Agora, nomeamos tantos delineamentos que surgiram livros didáticos dedicados à sua taxonomia e às suas “regras”. A estratégia subjacente a esse desenvolvimento era provavelmente como aquela subjacente aos sete guias autoconscientes para a conduta analítico-comportamental apresentada em 1968 (aplicada, comportamental, analítica etc.): na aplicação, um bom delineamento muitas vezes se mostraria difícil de aplicar ou manter e o treinamento de pós-graduação em aplicação pode nem sempre penetrar as profundezas do assunto; a classificação pode ajudar a superar essas dificuldades. As questões agora são se de fato a classificação dos delineamentos de pesquisa em tipos e regras ajudou nesse propósito; se assim for, em que medida; e, finalmente, se essa extensão vale o custo. O custo pode ser principalmente o fato de que os pesquisadores aplicados transformam cada vez mais as perguntas para se adequarem aos delineamentos conhecidos e a suas regras, em vez de construir um delineamento que responda à pergunta original. Pode se mostrar valioso para a área relembrar seus delineamentos originais e sua lógica: um bom delineamento é aquele que responde à pergunta de forma convincente e, como tal, precisa ser construído sob controle da pergunta e, em seguida, testado por meio de argumentos nesse contexto (às vezes chamado “*thinking through*” [pensar com cuidado]), em vez de imitado de um livro didático. Por exemplo, um artigo publicado em anais de congresso avaliou um programa de treinamento de jovens para preencher formulários de candidatura de emprego de maneira mais eficaz. Os pesquisadores pediram a vários empregadores que lessem uma sequência de candidaturas, cada uma escrita por um aluno diferente, começando com alguns escritos antes do treinamento e terminando com outros escritos após o treinamento; a mudança das candidaturas pré para pós-treinamento ocorreu em pontos diferentes na sequência de cada empregador, em um aparente delineamento de linha de base múltipla entre os empregadores. O delineamento determinava que todas as candidaturas para estagiários, tanto pré quanto pós-treinamento, seriam lidas. Quase sem exceção, os empregadores disseram “Não” às candidaturas escritas antes do treinamento e “Sim” às candidaturas para estagiários redigidas após o treinamento. Esse delineamento é mencionado em um artigo de Mathews e Fawcett (1984), mas não é descrito em detalhes por causa da teimosia editorial. Muitos na audiência do congresso, talvez também o editor, ignoraram o fato de que esse delineamento mostrava claramente que o programa de treinamento era excepcionalmente eficaz e, em vez disso, argumentaram que não era um delineamento de linha de base múltipla “característico”. Talvez o ponto a ser destacado seja que delineamentos convincentes deveriam ser mais importantes do que delineamentos “característicos”.

Tecnológica

Vinte anos atrás, era urgente recomendar que uma nova área aspirando tanto à aplicação eficaz quanto ao *status* de ciência fosse tanto procedimental quanto explícita sobre isso. O objetivo era evitar a situação de muitas disciplinas clínicas, gerenciais e administrativas nas quais, uma vez encerradas as discussões sobre teoria e metas, os procedimentos a serem aplicados eram especificados sem ir além de “trabalhe com”. Em grande medida, isso aconteceu; artigos de periódicos e livros didáticos oferecem uma lista completa de seus procedimentos operacionais, de tal forma que um leitor tenha uma chance razoável de replicar a aplicação com os mesmos resultados. Mais do que isso, começaram a surgir coleções de procedimentos, de tal forma que os leitores agora podem escolher entre procedimentos alternativos voltados para o mesmo objetivo.

Ainda assim, três pontos merecem ser comentados:

1. Alguns procedimentos, como elogios ou ensino incidental, muitas vezes variam no que o pesquisador considera ser uma maneira desejavelmente natural de ocasião para ocasião. Essas topografias e suas sequências raramente são especificadas; fazê-lo com antecedência pode muitas vezes ser considerado antinatural e fazê-lo retrospectivamente (e.g., a partir de uma fita de vídeo) seria caro para editores e entediante para os leitores. A suposição subjacente é, naturalmente, que essas variações não fazem diferença para o resultado. Essa suposição raramente ou nunca é testada empiricamente. Seria bom para a disciplina se uma revisão daqui a 20 anos pudesse afirmar que a suposição se mostrou correta ou que foi incorreta tantas vezes que a prática corrente resolveu o problema, apesar dos custos. (O tédio dos leitores com esse detalhe teria se dissolvido na descoberta de que essas variações poderiam de fato fazer a diferença no resultado.)

2. Na aplicação, os procedimentos realizados por pessoas (que são a maioria dos procedimentos da análise do comportamento aplicada) geralmente são observados e registrados, assim como os comportamentos do sujeito. Esta prática documenta até que ponto os procedimentos especificados são realizados e também descreve quaisquer procedimentos não especificados que possam ocorrer. O processo provavelmente é reativo em muitas aplicações, o

que cria maior adesão aos procedimentos especificados do que poderia ser garantido de outra forma. Mas esses dados raramente são apresentados fora do grupo que conduz a aplicação, novamente, provavelmente por causa das despesas das editoras e do suposto tédio dos leitores: quando esses dados mostram que os procedimentos especificados estão sendo realizados bem o suficiente, não há problema; e quando eles mostram o contrário, a aplicação geralmente é paralisada até que uma melhor adesão ao procedimento seja obtida e então não há novamente nenhum problema. Esse argumento é provavelmente defensável com base no custo-benefício, mas seria melhor para a disciplina se sua revisão daqui a 20 anos pudesse afirmar que o debate relevante havia ocorrido publicamente. Esse debate é essencialmente uma questão para editores e revisores de periódicos e de livros didáticos: eles pedem os dados ou não; publicam esses dados quando fornecidos ou se posicionam contra fazê-lo. Assim, eles podem muito bem usar um de seus futuros simpósios de periódicos para considerar esta questão, que é principalmente uma questão de política editorial.

3. A disseminação é uma prática muito mais antiga que a análise do comportamento aplicada, mas, no domínio do comportamento, geralmente é muito menos tecnológica que a análise do comportamento aplicada. Mesmo assim, sua literatura e seus praticantes debatem (sem resolução) uma questão essencialmente tecnológica: quando um programa é disseminado, seus disseminadores devem exigir que os procedimentos sejam seguidos fielmente, não importa onde ou quando o programa seja utilizado? Ou seus usuários devem ser autorizados, e até mesmo incentivados, a modificar esses procedimentos para se adequarem às suas situações e contingências locais? (Podemos primeiro perguntar, funcionalmente, quando temos essa escolha. Ou seja, quando é necessário manter a fidelidade aos procedimentos originais disponíveis para nós, e quando não?) A fidelidade aos procedimentos originais é recomendada porque esses procedimentos foram estudados e são conhecidos por serem eficazes; suas variações e alternativas geralmente não foram estudadas, então nada pode ser dito sobre sua eficácia. Por outro lado, a flexibilidade na aplicação é recomendada com a premissa de que todo o programa se tornará aversivo para pessoas que não podem modificá-lo para se adequar à sua situação e às suas contingências e se um programa não for usado, não pode ser eficaz.

Ambos são argumentos tecnológicos; curiosamente, o contextualismo, a experiência e o bom senso parecem concordar que cada um deles provavelmente estará correto em certos contextos, mas não em outros. A investigação empírica desses contextos controladores, obviamente, é crucial para a disseminação futura em larga escala (que é certamente a essência de *aplicada*), assim como a investigação de quando temos essa escolha. Essa pesquisa não foi feita amplamente; mas, presumivelmente, agora que uma disciplina tão tecnológica quanto a análise do comportamento aplicada entrou no domínio da disseminação, é mais provável que seja feita, ainda que de forma custosa. A estratégia apropriada foi recomendada por Sidman (1960) há quase 30 anos, em um contexto diferente, mas relevante: um dos critérios da ciência relevante é explorar as condições de controle de qualquer fenômeno comportamental. Qual é a gama de variações dos procedimentos de um programa que ainda permite eficácia suficiente? Se ela for grande o suficiente, a aplicação flexível pode ser incentivada e a sobrevivência do programa em diversas configurações pode muito bem ser aprimorada. Se for estreita, a fidelidade será necessária ou o que sobreviver não será eficaz.

Será interessante ver se uma revisão da disciplina daqui a 20 anos será capaz de resumir alguns fatos sobre esses processos ou terá que relatar que a análise do comportamento aplicada ainda está entrando em suas aplicações em larga escala sob grande risco de fracasso.

Capaz de Resultados Apropriadamente Generalizados

Vinte anos atrás, a capacidade da disciplina de produzir resultados apropriadamente generalizados foi vista como crucial para sua sobrevivência: uma disciplina aplicada que tinha que fazer todas as várias topografias de suas mudanças de comportamento desejadas e tinha que vincular cada uma delas a um controle de estímulo apropriado, ao longo do tempo, era intrinsecamente impraticável. Hoje, o problema ainda é crucial, mas agora para sua eficácia máxima, em vez da sobrevivência da disciplina. Nos últimos 20 anos, mudamos o comportamento como especificado e mostramos controle experimental de sua generalização apropriada apenas com frequência suficiente para deixar claro que a disciplina é capaz de tais resultados. O que resta é a tarefa muito mais reconfortante (e muito maior) de explorar as condições que controlam a generalização adequada (i.e., controle de estímulo apropriado).

Felizmente, o problema é visto agora, de maneira geral, como algo que provavelmente pode ser resolvido por meio de programação adequada, e não por sorte; assim, uma boa quantidade de pesquisas tem sistematicamente examinado maneiras de ensinar desde o início para que os resultados adequadamente generalizados sejam estabelecidos. (Apesar disso, um número notável de estudos não compara seu ensino facilitador de generalização a qualquer ensino alternativo do mesmo comportamento-alvo que não facilite sua generalização, por isso não

aprendemos nada sobre o problema com esses estudos.) O problema está longe de ser resolvido; ainda não temos um sistema para combinar o método de promoção de generalização mais adequado com a mudança de comportamento em questão e não há nem certeza de que existe tal sistema a ser encontrado. Nossas categorizações de técnicas de promoção de generalização são claramente não analíticas; foram propostas (ver Stokes & Baer, 1977) da mesma forma que as dimensões atuais (aplicadas, comportamentais, analíticas etc.): com base no pressuposto de que a classificação evocará mais do comportamento profissional necessário, especialmente a pesquisa (Baer, 1982). Essa suposição provavelmente não pode ser testada empiricamente: dificilmente podemos realizar um experimento que compare o progresso de nossa disciplina em direção ao controle minucioso da generalização, com e sem essas classificações. Assim, permanece a obrigação de continuar o debate (ver Johnston, 1979).

Efetiva

A característica fundamental de qualquer disciplina aplicada deve ser efetividade; isso não é diferente para a análise do comportamento aplicada. No entanto, no domínio da modificação do comportamento, a efetividade fundamental pode ser sutil: às vezes, parece ser simplesmente o grau em que o comportamento-alvo foi modificado; mas frequentemente diz respeito ao grau em que outra coisa foi modificada, não o comportamento-alvo, e esse algo diferente quase invariavelmente é o contracontrole de alguém contra o comportamento original (ver a discussão anterior sobre *Aplicada*). Assim, por exemplo, se olharmos de perto, podemos descobrir que, em alguns casos, mudar as notas de um aluno de F para C satisfaz o aluno, a família do aluno e aquele segmento da sociedade que eventualmente lerá essas notas e reagirá a elas: se as notas forem F, esses agentes verão isso como um problema; se as notas forem C, eles não o farão. Em alguns outros casos, no entanto, podemos descobrir que as notas de um estudante podem precisar ser alteradas de B para A antes que esse estudante, a família desse estudante e o segmento da sociedade que eventualmente lerá as notas e reagirá a elas parem de reagir a elas como se fossem um problema. A variável marcadora que distingue esses dois casos pode frequentemente parecer ser a classe social, mas isso não é analítico (ver Baer, 1984, pp. 547-551) nem relevante para o objetivo, que é o fato de que mudar as notas não é eficaz por si só; parar e evitar as referências relevantes a essas notas como um problema é o verdadeiro critério da efetividade de nossa intervenção.

Quase todo estudo bem-sucedido de modificação do comportamento deve apresentar rotineiramente dois resultados: uma medida dos comportamentos alterados, é claro, e uma medida das apresentações e descrições do problema que pararam ou diminuíram como consequência. Apesar disso, muito poucos estudos fazem isso. Talvez os pesquisadores presumam que *eles* são os únicos detectores de problemas relevantes ou responsáveis pela detecção de problemas. De fato, isso às vezes pode ser verdade, mas é melhor que isso seja defensável e explicitamente defendido ou se torna arrogância (que pode não promover o *status* social da disciplina se for amplamente notada como tal). Por outro lado, a ausência da segunda medida mencionada pode representar uma fraqueza crucial em nossa efetividade atual. Nós podemos ter ensinado muitas habilidades sociais sem examinar se elas realmente favoreceram a vida social do sujeito; muitas habilidades de cordialidade sem examinar se alguém realmente as notou ou se importou com elas; muitas habilidades de segurança sem examinar se o sujeito esteve mais seguro depois delas; muitas habilidades linguísticas sem mensurar se o sujeito as utilizou para interagir diferentemente de antes; muitas habilidades para engajar-se em tarefas sem mensurar o valor real dessas tarefas; e, em geral, muitas habilidades de sobrevivência sem verificar subsequentemente a sobrevivência do sujeito. Algumas dessas medidas serão controversas de serem definidas e custosas para serem coletadas, é claro; mas pode ser verdade que a disciplina tenha evoluído para um ponto em que elas se tornam cruciais. (Crianças se tornam mais custosas à medida que crescem.)

Talvez essa prática se torne mais difundida na disciplina à medida que o cálculo da relação custo-benefício aumente de sua taxa atual próxima de zero – se assumirmos seriamente o lado “benefício”, em vez de assumirmos que a modificação do comportamento por si só é o benefício. Em face disso, avaliações de relação custo-benefício são a essência da efetividade e devem ser rotina em qualquer disciplina aplicada (e.g., Hill et al., 1987). Elas se mostraram problemáticas nesse ponto, talvez em parte porque a disciplina ainda está em sua fase de experimentos de pesquisa, em parte porque os benefícios comportamentais não são tão claramente definidos quanto a maioria dos benefícios de negócios e em parte porque os próprios conceitos e técnicas para avaliar a relação custo-benefício ainda não estão claramente estabelecidos.

Felizmente, pelo menos uma segunda medida de eficácia está começando a se tornar rotina: validade social (Kazdin, 1977; Wolf, 1978), que é a extensão em que os consumidores de uma intervenção gostam dela (i. e., gostam de seus objetivos, alvos, efeitos, procedimentos e equipe). O ponto da medida de validade social é prever (e, assim,

evitar) a rejeição de uma intervenção, especialmente quando ela for disseminada (ao atingir grande escala, ela pode ser menos tolerável para os consumidores do que durante as pesquisas em pequena escala). Se uma intervenção é socialmente inválida, dificilmente ela será efetiva, mesmo que altere minuciosamente o comportamento-alvo e apresente uma excelente relação custo-benefício; validação social não é suficiente para a efetividade, mas é necessária.

Infelizmente, a validade social às vezes é medida atualmente de maneiras muito rudimentares que podem frequentemente encontrar validade social onde ela não opera de fato. Talvez o problema seja que os pesquisadores estão em um contexto de desejar validade social, o que é sutilmente diferente e muito mais perigoso do que um contexto de procurar por qualquer fonte de invalidade social. Visto que a disciplina está agora se movendo em direção a uma disseminação em larga escala, medidas efetivas de validação social logo se tornarão essenciais para sua sobrevivência; ainda assim, esse aspecto da nossa técnica de medição tem recebido muito pouca investigação e desenvolvimento.

Talvez uma revisão daqui a 20 anos possa relatar um grande progresso nessa dimensão da efetividade. Se assim for, o problema não terá sido simples. Por exemplo, algumas medidas de validação social são deliberadamente e pragmaticamente enviesadas para resultados positivos, não com o objetivo de enganar seus usuários, mas de evitar alarde entre seus consumidores (e.g., conselhos governamentais) enquanto, ao mesmo tempo, alertam seus usuários (os pesquisadores-aplicadores) a detectar e remediar os problemas que devem estar subjacentes às pontuações que geralmente estão próximas de 7, mas no momento estão em 5 em uma escala de 7 pontos. Além disso, é plenamente possível que mesmo consultas sobre a validade social bastante inválidas sejam melhores que nenhuma consulta: dar aos consumidores qualquer oportunidade de expressar reclamações e descontentamentos que de outra forma passariam despercebidos pode salvar pelo menos alguns programas de repercussões fatais, pelo menos se o consumidor ofendido ficar motivado o suficiente com a simples existência do questionário de avaliação de validade social, que seria de outro modo inadequado, para escrever em suas margens ou conversar com os aplicadores.

Talvez igualmente significativo seja o recente desenvolvimento de técnicas de avaliação que investigam quais são os objetivos do consumidor antes de um programa ser delineado (Fawcett et al., 1982; Schriener & Fawcett, 1988⁴), de tal forma que esse programa tenha a chance de corresponder a essas preferências, fazendo o máximo para garantir que tenha alta validade social quando essa dimensão for eventualmente avaliada. Essa tecnologia, se perseguida com intensidade suficiente, pode se tornar parte da análise pragmática da validade social, especialmente porque encontra temas comuns emergindo de suas investigações sobre os objetivos de diferentes conjuntos de consumidores no que parecem ser situações-problema bastante diferentes (cf. Seekins et al., 1988⁵); assim, a análise de sua validade pode ser um dos alvos prioritários para pesquisas futuras.

Talvez a medida mais clara da efetividade de nossa disciplina seja o número crescente de aplicações inefetivas que temos testado nos últimos anos. Por bom senso ou boa sorte, começamos com casos dramáticos, perturbadores, mas essencialmente delimitados, e nossa eficácia com eles reforçou fortemente nossos comportamentos disciplinados. Se fosse de outra maneira, poderíamos não ser a reconhecida disciplina aplicada que somos hoje. Mas, por termos feito isso, é claro que estamos nos movendo para diferentes classes de problemas, em parte porque esses problemas estão aí, em parte porque eles são excepcionalmente importantes e em parte porque nós continuamos sendo uma disciplina aplicada que se baseia em pesquisa e porque pesquisa não deve ser muito repetitiva, então, na medida em que fizemos (ou pelo menos demonstramos) tudo o mais, esses problemas são o que resta abordar.

Mas os problemas de hoje não são tão delimitados quanto aqueles dos nossos primórdios. Eles são chamados de estilos de vida em reconhecimento de sua natureza sistêmica. As classes de comportamentos chamadas de delinquência, abuso de substâncias, segurança, exercício físico e dieta, por exemplo, representam complexas classes topográficas com complexas funções envolvendo muitos agentes de reforçamento/punição e controle de estímulo, todos eles interagindo para constituir e manter o sistema como tal. Portanto, a entrada em apenas um ponto de tais sistemas provavelmente produzirá apenas mudanças de comportamento limitadas e de curto prazo antes do inevitável contracontrole restaurar as posições anteriores desse sistema, conosco paralisados ou cooptados ineficazmente dentro dele (ver Wolf, Braukmann, & Ramp, 1987). O primeiro remédio é reconhecer: o

⁴ Nota dos tradutores: no artigo original, a obra ainda não havia sido publicada. Para facilitar o acesso de possíveis leitores interessados, optamos por apresentar os dados de publicação da referência.

⁵ Nota dos tradutores: no artigo original, a obra ainda não havia sido publicada. Para facilitar o acesso de possíveis leitores interessados, optamos por apresentar os dados de publicação da referência.

conceito de análise de sistemas é agora um componente importante de nossa efetividade e a pesquisa que nos mostrará como fazer isso melhor será excepcionalmente útil. O segundo remédio, seguindo qualquer análise que o primeiro permita, é a intervenção em todo o sistema: portanto, por exemplo, crianças obesas são tratadas não como simples interações terapeuta-cliente, mas dentro de seus sistemas de vida – pelo menos dentro de suas famílias (Brownell et al., 1983) ou, melhor ainda, dentro de suas famílias e seu sistema escolar (Brownell & Kaye, 1982). O terceiro remédio pode ser a discriminação entre aqueles problemas em que uma única intervenção de curta duração pode ser efetiva e aqueles problemas invariavelmente sistêmicos em que a presença crônica será necessária para manter a intervenção eficaz. Assim como a medicina reconhece que a apendicite precisa apenas de uma intervenção e que a diabetes precisa de tratamento para toda a vida e educa seus consumidores para a inevitabilidade desses tratamentos e seus custos, a análise do comportamento aplicada deveria começar sua validação e uso dessas mesmas duas categorias.

Talvez o remédio mais importante de todos, no entanto, será estabelecer um contexto apropriado para se responder às falhas. Os últimos 20 anos produziram uma taxa crescente delas; os próximos 20 anos quase certamente verão essa taxa se manter e, muito provavelmente, crescer ainda mais. Esse fato e essa probabilidade já foram interpretados como uma inadequação dos princípios analítico-comportamentais. Por exemplo, Reppucci e Saunders (1974) responderam a uma de suas falhas em uma instituição para delinquentes com um princípio amplamente generalizado:

Finalmente, há uma questão cuja resolução terá enormes consequências para a modificação do comportamento tal como a conhecemos e aplicamos hoje. A questão se inere ao fato⁶ de que os princípios da modificação do comportamento são insuficientes e muitas vezes inadequados para a compreensão dos ambientes naturais – sua estrutura, seus objetivos, sua tradição e seus vínculos de interseção. (p. 569)

Sua publicação no jornal *American Psychologist*, é claro, apresentou esse novo “fato” a potencialmente todos os membros da American Psychological Association (APA) (a maioria dos quais não sabia que é o único tipo de avaliação de modificação de comportamento que o jornal da associação publica e não se perguntaria se um “fato” como esse pode ser estabelecido a partir de uma falha na instalação e na manutenção de um programa em uma única instituição).

Vale a pena primeiro nos perguntarmos se falha tecnológica é o mesmo que falha teórica. Muito provavelmente, uma falha tecnológica é esperada e é de fato um importante evento no progresso de *qualquer* área aplicada, mesmo para aquelas cujas teorias subjacentes são totalmente válidas. Assim, a passagem do laboratório de física para a engenharia foi, e continuará a ser, marcada por elevadores emperrados, pontes caídas, aviões colididos e espaçonaves explodidas. Os engenheiros sabem disso; eles abandonam apenas seus projetos, não suas teorias, com cada um desses eventos. Petroski (1958), por exemplo, resume sua história da seguinte forma:

Acredito que o conceito de falha – falhas mecânicas e estruturais no contexto desta discussão – é central para compreender a engenharia, pois o projeto de engenharia tem como primeiro e principal objetivo evidenciar o fracasso. Assim, os desastres colossais que ocorrem são, no final das contas, falhas de projeto, mas as lições aprendidas com esses desastres podem contribuir mais para o avanço do conhecimento em engenharia do que as máquinas e estruturas bem-sucedidas que existem no mundo. De fato, falhas parecem ser inevitáveis no despertar de um sucesso prolongado, que encoraja margens de segurança mais estreitas. As falhas, por sua vez, levam a mais largas margens de segurança e, portanto, a novos períodos de sucesso. Compreender o que é engenharia e o que engenheiros fazem é compreender como falhas podem acontecer e como elas contribuem mais do que sucessos para o avanço da tecnologia. (p. xii)

Esse mesmo ponto é inerente ao progresso da medicina: toda morte é, em certo sentido, uma falha em nossos modelos atuais para manutenção da saúde, assim como cada ponte caída é uma falha no projeto para balancear carga e força.

A análise do comportamento aplicada deve lidar com fenômenos ao menos tão complexos quanto pontes sobrecarregadas e fisiologias tensionadas e, talvez, considerando os domínios de variáveis relevantes para o comportamento, mais complexos. Então, ela irá prosseguir de início com tantos, ou mais, delineamentos falhos quanto os que a área já tem; mas ela se beneficiará dessas falhas, como já tem feito, e precisará de tempo e repetição para fazê-lo, como também já tem feito.

Como sabemos que uma determinada falha reflete um delineamento ruim em vez de um princípio inadequado? Nós nunca sabemos isso; mas podemos buscar por um delineamento ruim imediatamente após cada falha e, se o

⁶ Nota dos tradutores: o texto original apresenta uma construção em inglês não usual (i.e., “the issue inheres in the fact”). Na tradução, optamos por manter construção semelhante.

encontrarmos, isso nos dará algo a testar na próxima intervenção muito mais prontamente do que se houver desespero com relação a nossos princípios. Por exemplo, Reppucci e Saunders fizeram o papel apenas de consultores externos em suas falhas. Na verdade, são os *princípios* da análise do comportamento (assim como uma experiência considerável) que sugerem o pouco potencial para alterar o comportamento de funcionários e gestores sobrecarregados e desinteressados com as geralmente poucas contingências disponíveis para os consultores (a menos que uma crise grave esteja em andamento). Liberman (1980), em resposta a um número similar de falhas, sugeriu não um novo princípio, mas meramente uma intervenção mais poderosa – que alguns de nós combinemos pesquisa com gestão:

Não podemos contar com a necessidade de responsabilidade e avaliação de programas por parte dos gestores como suporte para nossos programas comportamentais. ... Se quisermos que nosso trabalho viva para além de uma estante de biblioteca, nós teremos de partir para o domínio da grande política e nos aventurar como pesquisadores-gestores. (pp. 370-371)

Se nós buscarmos os programas comportamentais que se mantiveram por períodos de tempo impressionantes, encontraremos o padrão recomendado por Liberman com a mesma frequência com que encontramos períodos de tempo impressionantes: o próprio Liberman no Centro de Saúde Mental de Oxnard, McClannahan e Krantz no Instituto de Desenvolvimento Infantil de Princeton, Cataldo no Instituto Johns Hopkins Kennedy e Christian no Instituto May, por exemplo. Esses casos não são a prova de nada, nem pretendem ser; simplesmente vale a pena considerá-los como delineamentos que podem atrair uma prova e, se o fizerem, podem render benefícios para nossa disciplina.

O modelo Família-Ensino [*Teaching-Family model*] é outro exemplo de um delineamento um tanto diferente e aparentemente durável, tão velho quanto este periódico⁷. Ele criou 12 centros de treinamento regionais para intermediar 215 replicações do programa original de delinquência juvenil *Achievement Place* e não escreveu artigos em periódicos científicos, mas simples manuais de treinamento para seu uso. Os criadores desse programa também seguiram a prescrição de Liberman; eles acrescentaram à sua função de pesquisa aquelas de administrar os programas componentes e lutar suas batalhas políticas e de assegurar apoio financeiro consistente o suficiente para os necessários 10 anos de pesquisa de tentativa-e-erro-e-nova-tentativa-e-sucesso para compreender e implementar o controle de qualidade, o treinamento de equipe e os outros sistemas de suporte essenciais para sua sobrevivência e disseminação. De fato, mesmo 20 anos não foram suficientes para se observar a completude desse programa de pesquisa e o desenvolvimento do programa, mas as falhas agora são raras, apesar das oportunidades amplamente expandidas (Wolf et al., 1987).

O ponto é que falhas ensinam; o modelo Família-Ensino se desenvolveu a partir do modelo Ensino-Falha. Certamente nossos periódicos deveriam começar a publicar não apenas os sucessos de nossa área, mas também aquelas falhas que poderiam nos fazer enxergar a possibilidade de delineamentos melhores que elas.

Em resumo, a efetividade para o futuro provavelmente será construída com base principalmente em intervenções no âmbito dos sistemas e em falhas de alta qualidade, à medida que continuemos trazendo a teoria para delineamentos que resolvem problemas. Mas a base deve ser a teoria vigente, e não alguma substituição dela – a teoria atual tem funcionado suficientemente bem para não ser abandonada diante do que é visto mais parcimoniosamente como falha tecnológica e não teórica. Claramente, aumentar nossa efetividade não será fácil e não acontecerá rapidamente. Nós devemos esperar um longo período de pesquisa difícil, custoso, repetitivo e às vezes ineficaz nessas aplicações e entrar nessa pesquisa com nossas melhores habilidades sociais, porque podemos requerer a cooperação de uma quantidade extraordinariamente grande de pessoas, muitas vezes em cargos extraordinariamente evidentes. No entanto, mesmo com relativamente pouco trabalho de reação ao fracasso antes de nós, parece claro que podemos fazê-lo.

Parece claro que podemos fazer aquilo que ainda há a ser feito. O fato de que nós podemos é provavelmente a dimensão mais fundamental, mais importante e mais duradoura; e o fato de que nós conseguiremos é simplesmente lógico.

Referências

- Baer, D. M. (1981). A flight of behavior analysis. *The Behavior Analyst*, 4(2), 85-91. <https://doi.org/10.1007/BF03391857>
- Baer, D. M. (1982). The role of current pragmatics in the future analysis of generalization technology. In R. B. Stuart (Ed.), *Adherence, compliance, and generalization in behavioral medicine* (pp. 192-212). Brunner/Mazel.

⁷ Nota dos tradutores: os autores referem-se ao periódico em que o artigo original foi publicado, o *Journal of Applied Behavior Analysis* (JABA). Tal periódico é publicado desde 1968.

- Baer, D. M. (1984). Future directions? Or, is it useful to ask "Where did we go wrong?" before we go?. In R. F. Dangel, & R. A. Polster (Eds.). *Behavioral parent training: Where it came from and where it's at* (pp. 547-557). Guilford.
- Baer, D. M. (1986). In application, frequency is not the only estimate of the probability of behavioral units. In T. Thompson & M. D. Zeiler (Eds.). *Analysis and integration of behavioral units* (pp. 117-136). Lawrence Erlbaum Associates.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91-97. <https://doi.org/10.1901/jaba.1968.1-91>
- Brownell, K. D., & Kaye, F. S. (1982). A school-based behavior modification, nutrition education, and physical activity program for obese children. *The American Journal of Clinical Nutrition, 35*(2), 277-283. <https://doi.org/10.1093/ajcn/35.2.277>
- Brownell, K. D., Kelman, J. H., & Stunkard, A. J. (1983). Treatment of obese children with and without their mothers: Changes in weight and blood pressure. *Pediatrics, 71*(4), 515-523. <https://doi.org/10.1542/peds.71.4.515>
- Cronbach, L. J., Glaser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. John Wiley & Sons, Inc.
- Deitz, S. M. (1978). Current status of applied behavior analysis: Science versus technology. *American Psychologist, 33*(9), 805-14. <https://doi.org/10.1037/0003-066X.33.9.805>
- Fawcett, S. B., Seekins, T., Whang, P. L., Muiu, C., & de Balcazar, Y. S. (1982). Involving consumers in decision-making. *Social Policy, 13*(2), 36-41.
- Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability estimates. *Journal of Applied Behavior Analysis, 10*(1), 103-116. <https://doi.org/10.1901/jaba.1977.10-103>
- Hill, M. L., Banks, P. D., Handrich, R. R., Wehman, P. H., Hill, J. W., & Shafer, M. S. (1987). Benefit-cost analysis of supported competitive employment for persons with mental retardation. *Research in Developmental Disabilities, 8*(1), 71-89. [https://doi.org/10.1016/0891-4222\(87\)90041-2](https://doi.org/10.1016/0891-4222(87)90041-2)
- Hineline, P. N. (1980). The language of behavior analysis: Its community, its functions, and its limitations. *Behaviorism, 8*(1), 67-86.
- Johnston, J. M. (1979). On the relation between generalization and generality. *The Behavior Analyst, 2*(2), 1-16. <https://doi.org/10.1007/BF03391833>
- Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification, 1*(4), 427-452. <https://doi.org/10.1177/014544557714>
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2a ed.). University of Chicago Press.
- Lieberman, R. P. (1980). A review of Paul and Lentz's psychological treatment for chronic mental patients: Milieu versus social-learning programs. *Journal of Applied Behavior Analysis, 13*(2), 367-371. <https://doi.org/10.1901/jaba.1980.13-367>
- Mathews, R. M., & Fawcett, S. B. (1984). Building the capacities of job candidates through behavioral instruction. *Journal of Community Psychology, 12*(2), 123-129. [https://doi.org/10.1002/1520-6629\(198404\)12:2%3C123::AID-JCOP2290120204%3E3.0.CO;2-O](https://doi.org/10.1002/1520-6629(198404)12:2%3C123::AID-JCOP2290120204%3E3.0.CO;2-O)
- Michael, J. (1980). Flight from behavior analysis. *The Behavior Analyst, 3*(2), 1-21. <https://doi.org/10.1007/BF03391838>
- Morris, E. K. (1982). Some relationships between interbehavioral psychology and radical behaviorism. *Behaviorism, 10*(2), 187-216.
- Petroski, H. (1985). *To engineer is human: The role of failure in successful design*. Saint Martin's Press.
- Pierce, W. D., & Epling, W. F. (1980). What happened to analysis in applied behavior analysis? *The Behavior Analyst, 3*(1), 1-9. <https://doi.org/10.1007/BF03392373>
- Powell, J. (1984). On the misrepresentation of behavioral realities by a widely practiced direct observation procedure: Partial interval (one-zero) sampling. *Behavioral Assessment, 6*(3), 209-219. <https://doi.org/10.1007/BF01321463>
- Reppucci, N. D., & Saunders, J. T. (1974). Social psychology of behavior modification: Problems of implementation in natural settings. *American Psychologist, 29*(9), 649-660. <https://doi.org/10.1037/h0037630>
- Schriner, K. F., & Fawcett, S. B. (1988). Development and validation of a community concerns report method. *Journal of Community Psychology, 16*(3), 306-316. [https://doi.org/10.1002/1520-6629\(198807\)16:3<306::AID-JCOP2290160306>3.0.CO;2-5](https://doi.org/10.1002/1520-6629(198807)16:3<306::AID-JCOP2290160306>3.0.CO;2-5)
- Seekins, T., Mathews, R. M., Fawcett, S. B., Jones, M. L., & Budde, J. F. (1988). A market-oriented strategy for applied research in independent living rehabilitation. *Journal of Rehabilitation, 54*(1), 35-38.
- Sidman, M. (1960). *Tactics of scientific research*. Basic books.
- Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.). *Analysis and integration of behavioral units* (pp. 213-245). Lawrence Erlbaum Associates.

- Stoddard, L. T. (Ed.). (1986). Seção Especial. *Analysis and Intervention in Developmental Disabilities*, 6(1-2). [https://doi.org/10.1016/0270-4684\(86\)90002-9](https://doi.org/10.1016/0270-4684(86)90002-9)
- Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, 10(2), 349-367. <https://doi.org/10.1901/jaba.1977.10-349>
- Ulman, J. D., & Sulzer-Azaroff, B. (1975). Multi-element baseline design in educational research. In E. Ramp & G. Semb (Eds.), *Behavior analysis: Areas of research and application* (pp. 371-391). Prentice Hall.
- Wahler, R. G., & Fox, J. J. (1982). Response structure in deviant child-parent relationships: Implications for family therapy. In D. J. Bernstein (Ed.), *Response structure and organization. The 1981 Nebraska symposium on motivation* (pp. 1-46). University of Nebraska Press.
- Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11(2), 203-214. <https://doi.org/10.1901/jaba.1978.11-203>
- Wolf, M. M., Braukmann, C. J., & Ramp, K. A. (1987). Serious delinquent behavior as part of a significantly handicapping condition: Cures and supportive environments. *Journal of Applied Behavior Analysis*, 20(4), 347-359. <https://doi.org/10.1901/jaba.1987.20-347>
- Zuriff, G. (1985). *Behaviorism: A conceptual reconstruction*. Columbia University Press. <https://doi.org/10.7312/zuri90466>

Submetido em: 03/03/2023

Aceito em: 06/05/2023