

Teorias do Reforço na Tradição Conexionista

Na história da psicologia, ocupa lugar destacado a ideia de que o prazer e a dor, como conseqüências de nossos atos, são importantes fatores do comportamento. Esta ideia forma a base da teoria do hedonismo psicológico, desenvolvida por Jeremy Bentham e adotada por uma série de outros filósofos britânicos. De acordo com esse ponto de vista, todos nós fazemos coisas que nos dão prazer e evitamos aquelas que nos causam sofrimento. Entretanto, coube a Edward L. Thorndike (1874-1949) tornar esse enfoque um ponto central na psicologia da aprendizagem.

O CONEXIONISMO INICIAL DE THORNDIKE

Thorndike foi pioneiro em psicologia animal experimental. Em vez de basear-se em histórias de proezas inteligentes deste ou daquele animal, levou animais para o laboratório, colocou-os diante de problemas padronizados e efetuou cuidadosas observações sobre o modo como resolviam os problemas. Sua monografia, *Animal Intelligence* (Thorndike, 1898), é um dos clássicos de maior prestígio nesta área. Seu estudo mais amplamente citado foi o que realizou com gatos numa caixa-problema. Um gato faminto ficava preso numa jaula, com uma tentadora posta de peixe colocada do lado de fora. O gato podia abrir a portinhola puxando o laço de um cordão pendente no lado interno da jaula. Geralmente, um gato passava por um longo processo, que consistia em andar a esmo, arranhar as paredes da jaula, e outras respostas, até que chegava a puxar o laço do cordão e conseguia sair. Em provas sucessivas na jaula, os animais levavam tempos cada vez mais curtos para

puxar o cordão. Entretanto, essa melhora era muito gradual. Mesmo depois de algumas experiências de abrir a portinhola puxando o cordão, os animais em um determinado ensaio ainda levavam considerável tempo em outra forma de comportamento antes de puxar o cordão. Isto levou Thorndike a concluir que a aprendizagem do gato no sentido de puxar o cordão implicava não uma compreensão "inteligente" de uma relação entre puxar o cordão e abrir a portinhola, mas uma gradual "fixação" da conexão estímulo-resposta entre ver o cordão e puxá-lo.

Na época em que Thorndike publicou seus trabalhos, estes eram radicais em dois aspectos. Sua cuidadosa observação do comportamento animal sob condições controladas e seu interesse pelo gradual fortalecimento das ligações estímulo-resposta; constituíram a resposta de Thorndike à discussão a respeito de se os animais resolvem problemas pelo raciocínio ou por instinto. Nem uma coisa nem outra, disse Thorndike, mas sim pela gradual aprendizagem da resposta certa.

Em relação a Watson e Guthrie, porém, um outro ponto deve ser assinalado. Enquanto Watson e Guthrie eram teóricos da contigüidade, Thorndike foi um teórico do reforço. As leis da frequência e da recentidade de Watson e a lei básica da aprendizagem de Guthrie afirmam, ambas, que as ligações estímulo-resposta fortalecem-se simplesmente pelo fato de que a resposta ocorre em presença dos estímulos. Thorndike não rejeitou completamente este ponto de vista, que resumiu sob a forma da *lei do exercício*. Sua lei fundamental da aprendizagem, porém, foi a *lei do efeito*. Esta afirmava que a fixação das conexões estímulo-resposta dependia não simplesmente do fato de que o estímulo e a resposta ocorriam conjuntamente, mas dos efeitos que acompanhavam a resposta. Se um estímulo se acompanhava de uma resposta e, a seguir, de um fator de satisfação, a conexão estímulo-resposta era fortalecida. Se, porém, um estímulo se acompanhava de uma resposta e, depois, de um fator de irritação, a conexão estímulo-resposta se enfraquecia. Esses efeitos satisfatórios e irritantes das respostas determinavam se as conexões estímulo-resposta seriam fixadas ou eliminadas.

Os termos "fator de satisfação" e "fator de irritação" parecem demasiado subjetivos para uma teoria que se ocupava da fixação e eliminação mecânica de ligações estímulo-resposta. Essa linguagem é muito mais semelhante à dos filósofos hedonistas do que à dos psicólogos behavioristas. Thorndike foi efetivamente criticado pelos behavioristas por esse modo de falar a respeito da aprendizagem. Na realidade, porém, definiu estes termos de modo bastante objetivo (Thorndike, 1913):

Estado de coisas satisfatório significa aquele que o animal nada faz para evitar, frequentemente realizando coisas que o mantêm e renovam; estado de coisas irritante significa aquele que o animal nada faz para preservar, frequentemente realizando coisas que põem um fim a ele (Cap. 1).

Thorndike nada refere acerca dos sentimentos do animal, somente acerca daquilo que o animal faz. Com isso, adere ao interesse do behaviorismo por aquilo que os indivíduos fazem. Sua linguagem pode parecer subjetiva, mas a significação é tão objetiva como a de Watson. Trabalhando no auge do movimento behaviorista, Thorndike teve discordâncias com os seguidores extremados deste sistema, mas realmente ele e os demais se aproximavam muito quanto aos interesses e objetivos. Num sentido mais amplo do termo, Thorn-

dike certamente era behaviorista.

Numa fase posterior de sua carreira, Thorndike modificou a lei do efeito, tornando os fatores de satisfação muito mais importantes do que os fatores de irritação. A recompensa, afirmava ele, fortalece as conexões, mas a punição não as enfraquece diretamente. Se a punição, em algum caso, é eficaz para enfraquecer a tendência de fazer algo, isto assim ocorre porque produz comportamento variável, e, desse modo, dá a alguma resposta nova a possibilidade de ser recompensada. Essa posição, exceto quanto à ênfase na recompensa, parece muito semelhante à de Guthrie. Com tal modificação, a lei do efeito consistiu simplesmente na afirmação, atualmente bastante conhecida (mas absolutamente não conhecida na época em que foi apresentada por Thorndike) segundo a qual seqüências satisfatórias servem para reforçar ligações estímulo-resposta.

Thorndike era homem de interesses práticos, interessando-se especialmente pela psicologia da educação. Durante muitos anos, fez parte do corpo docente no Teachers College da Universidade de Columbia. Durante toda a sua vida profissional, seus estudos sobre a psicologia "pura" da aprendizagem, com temas relacionados tanto com animais como com o homem, misturaram-se aos estudos de psicologia aplicada à educação. No decorrer do tempo, educadores elogiaram e condenaram a ênfase que ele dava à especificidade na aprendizagem, e à fixação mecânica de conexões estímulo-resposta (na realidade, é a Thorndike que devemos o termo "conexionismo"). Para nossos fins, porém, estes aspectos do trabalho de Thorndike são muito semelhantes aos de Watson e Guthrie, para exigir maiores comentários neste capítulo. Thorndike foi tão pioneiro da psicologia objetiva quanto Watson; realmente, suas contribuições originais foram, com grande probabilidade, mais importantes que as de Watson. Entretanto, o que nos interessa assinalar aqui é que Thorndike incorporou à sua psicologia objetiva da aprendizagem a lei do efeito, tornando-se, desse modo, realmente o primeiro teórico do reforço.

A INTERPRETAÇÃO DA APRENDIZAGEM DE MILLER

Vejamos agora uma outra posição conexionista do reforço, que é a posição de Neal Miller (nascido em 1909); professor de psicologia em Yale, durante grande parte de sua carreira, passou a fazer parte da Universidade Rockefeller em 1966. O sistema de Miller representa, em grande escala, uma simplificação da teoria de Clark Hull, que veremos detalhadamente mais adiante. O exame das proposições de Miller antes das de Hull, embora inverta a seqüência cronológica natural, pode servir como introdução ao estudo do sistema de Hull, altamente técnico.

Grande parte da importância de Miller como teórico provém das aplicações de sua teoria aos temas da psicologia da personalidade, psicologia social e psicopatologia. Essas aplicações foram aperfeiçoadas conjuntamente por Miller e John Dollard (nascido em 1900). A experiência de Miller dentro da área da psicologia experimental e a experiência de Dollard em psicologia clínica e ciências sociais possibilitaram uma colaboração muito frutífera nestas áreas. A análise teórica da aprendizagem é basicamente de Miller, com grandes créditos em favor de Hull, mas as aplicações são principalmente de Dollard.

Os quatro elementos da aprendizagem

O conceito de *impulso* (*drive*) ocupa um lugar central na maneira como Miller interpreta a aprendizagem e a motivação. Um impulso (*drive*) é um estado de excitação do organismo, um estado que impele o indivíduo à ação. Para Miller, um impulso sempre envolve um estímulo intenso; além do mais, qualquer estímulo, se suficientemente intenso, atua como impulso. O estímulo de impulso pode ser ou interno ou externo. A dor é um exemplo de impulso produzido por um estímulo externo, enquanto a fome e a sede são impulsos produzidos por estímulos internos. Alguns impulsos são produzidos por estímulos advindos de respostas emocionais do indivíduo. Quando sentimos raiva ou medo, ocorrem em nosso organismo alterações fisiológicas. Alguns destas produzem intensos estímulos internos que são responsáveis pelos impulsos de raiva e medo. Qualquer que seja a fonte destes, um impulso desperta o indivíduo e mantém-no ativo. O impulso é, pois, a base da motivação.

Num bebê, ou num indivíduo colocado numa situação altamente não habitual, a atividade produzida pelo impulso tem pouca direção. Executam-se respostas diferentes de todos os tipos. Pode acontecer que uma dessas respostas sirva para reduzir a intensidade do impulso. Se o impulso é a fome, comer comida irá reduzi-lo. Se o impulso é dor provocada por choque elétrico, essa dor se reduz quando o indivíduo se afasta do condutor carregado de electricidade. Quando o impulso se reduz (o que, freqüentemente — mas não sempre — significa que é completamente removido), o indivíduo torna-se menos ativo. Assim, o que podemos dizer a respeito de impulso é que um estímulo intenso aumenta a atividade, e que a remoção do estímulo intenso diminui a atividade.

Até esse ponto, trata-se apenas de um enunciado a respeito de motivação, não a respeito de aprendizagem. Entretanto, a redução da intensidade de um impulso possui uma propriedade muito importante: reforça qualquer resposta que veio logo antes. Assim, qualquer resposta que sirva para reduzir o impulso é reforçada e, portanto, tende a ser aprendida. Conseqüentemente, o comportamento produzido pelo impulso numa situação habitual é bastante diferente do "ensaio e erro" numa situação não-habitual. O indivíduo, nesse caso, rapidamente executa a resposta aprendida que reduziu o impulso no passado e, deste modo, de imediato a reduz novamente. A redução do impulso, portanto, é a operação básica na aprendizagem.

Com isto, temos um impulso, uma resposta e uma resultante redução do impulso. É necessário acrescentar mais um elemento: o conjunto de estímulos que orienta a resposta. Não se aprende uma resposta simplesmente para reduzir determinado impulso sempre que ou onde quer que o impulso ocorra; a resposta depende de condições. Sob algumas condições, uma pessoa reduz a fome indo a um restaurante; sob outras condições, preparando uma refeição. Os estímulos que orientam a resposta e que determinam qual será a resposta são conhecidos como *índícios*. Por fim, a redução do impulso que põe término à seqüência da aprendizagem pode, naturalmente, ser chamada uma *recompensa*. Este é o quarto dos elementos da aprendizagem que Miller e Dollard tornaram famosos: *drive*, *cue*, *response* e *reward*, isto é, impulso, índice, resposta e recompensa. Na caixa-problema de Thorndike, por exemplo, o impulso era a fome (e a exploração, muito provavelmente), o índice mais importante era o laço do cordão, a resposta consistia em puxar o cordão, e a recompensa era a saída e a comida.

Essa posição, embora muito semelhante à de Thorndike, também guarda alguma semelhança com a de Guthrie. Um estímulo intenso (*drive*, impulso) age como estímulo mantenedor, nos termos de Guthrie. A remoção deste estímulo, porém, age conforme o que Thorndike denominou fator de satisfação e o que Miller chama de recompensa. Essa recompensa funciona como reforçador, aumentando a tendência no sentido de a resposta recompensada ocorrer novamente. Essa ênfase no reforço distingue ambos, Miller e Thorndike, de Watson e Guthrie, tornando aqueles representantes de uma diferente tradição na teoria da aprendizagem.

Imitação

Em seu primeiro livro, *Social Learning and Imitation*, Miller e Dollard (1941) expõem sua interpretação básica e, depois, passam a aplicá-la a uma série de situações complexas. Assinalam que grande parte do comportamento de aprendizagem do homem envolve imitação. Em muitíssimas situações, as pessoas solucionam problemas, não tentando uma resposta depois de outra até que uma dessas respostas seja recompensada, mas sim fazendo o que vêem uma outra pessoa fazer. De que modo esta conduta se coaduna com o modelo simples de aprendizagem apresentado por Miller e Dollard? A resposta é que a tendência a imitar é em si mesma aprendida. Quando um indivíduo executa uma resposta, frequentemente essa resposta é realizada em presença de indícios produzidos pelo comportamento de outrem. A resposta dessa pessoa pode ser ou semelhante ou diferente daquela de um outro indivíduo. Se a resposta é semelhante à de uma outra pessoa, e se vem acompanhada de redução do impulso, o indivíduo foi recompensado por usar os indícios oriundos de um outro indivíduo para modelar sua resposta em função daquela de outrem. Se a resposta é diferente e não se acompanha de uma redução do impulso, a tendência a comportar-se de maneira diferente daquela de outrem não é recompensada e começa a extinguir-se. Desse modo, o comportamento imitativo é recompensado e outro comportamento não o é, de forma que o indivíduo aprende a fazer aquilo que vê um outro indivíduo fazendo.

Miller e Dollard dão o exemplo de dois meninos, com idades de seis e três anos, participando de uma brincadeira com o pai. Para cada um deles, o pai escolheu um bombom, e, depois, os meninos tiveram que procurá-los. Em todos os lugares onde o menino mais velho dava buscas, o irmão menor "grudava" nele e também procurava. Quando o mais velho encontrou seu bombom e parou de procurar, o menor ficou sem ação. A única estratégia que conhecia nesse jogo consistia em imitar em todos os detalhes o irmão mais velho.

De que modo se formou esse padrão de imitação, tão frustrante nesse caso? Miller e Dollard remontam a um incidente anterior da vida desses irmãos. Certa vez, o menino mais velho ouviu os passos do pai e correu para cumprimentá-lo tão logo entrasse em casa vindo do trabalho. Aconteceu de o irmão menor estar correndo na mesma direção ao mesmo tempo. O pai presenteou ambos os meninos com doces. Visto que o menino menor não tinha aprendido a distinguir os passos do pai, não estava correndo para ir ao encontro dele. O fato de estar correndo era simplesmente uma coincidência. Não obstante, foi recompensado por ter corrido. Naturalmente, no passado, houve muitas ocasiões em que tinha corrido e não fora recompensado por isto. O indício que tornou esta vez diferente das outras foi o fato de ver o irmão

correndo. Assim, houve uma situação de aprendizagem para o menino menor; a fome era o impulso, a visão do irmão mais velho correndo era o indício, correr era a resposta, e o bombom, a recompensa. Miller e Dollard afirmavam que fora através desse incidente e de muitos outros semelhantes que o menino menor terminou por aprender a imitar o mais velho. Tendo sido recompensado muitas vezes por imitar o irmão, fazia-o até mesmo em situações tais como essa do brinco de esconder, em que tal conduta era inútil.

É fácil pensar nesse tipo de situações em que as pessoas são recompensadas por imitarem outras. Miller e Dollard não se satisfizeram, porém, simplesmente com apresentar esses exemplos. Passaram a realizar experiências em que tanto seres humanos como animais eram ensinados a imitar. Numa dessas experiências, o sujeito (uma criança de escola primária) conseguia obter doce de uma máquina, às vezes girando o cabo, às vezes pressionando-o para baixo. Se a criança fizesse um movimento errado, não conseguia doce nessa tentativa. Juntamente com cada criança, colocava-se uma outra pessoa, que também tinha uma oportunidade de operar a máquina logo antes que o sujeito o fizesse. Ora esta outra pessoa era uma outra criança, ora um adulto. Se o adulto girasse o cabo, girar o cabo também era a resposta certa para a criança. Se a outra criança girasse o cabo, porém, só a pressão do cabo poderia funcionar para o sujeito. Em outras palavras, nessa situação, as crianças eram recompensadas quando imitavam adultos, mas não eram recompensadas quando imitavam outras crianças. As crianças não apenas aprenderam a imitar o adulto e a não imitar a outra criança, como também generalizaram este comportamento a outros adultos e a outras crianças. Algumas outras crianças foram recompensadas por imitarem crianças, mas não por imitarem adultos, e também aprenderam. Este experimento ilustra não somente a aprendizagem da imitação, mas também a generalização da imitação a partir de uma pessoa para outra e a discriminação entre pessoas a serem imitadas e pessoas a não serem imitadas.

Miller e Dollard passam a aplicar esses princípios a uma série de situações sociais. Assinalam, por exemplo, que aprendemos a imitar mais as pessoas de grande prestígio do que aquelas de pouco prestígio. Uma criança provavelmente é recompensada por imitar "aquelas crianças boas (de classe média)" e não é recompensada por se comportar como "aquelas crianças malcriadas do outro lado do bairro (classe inferior)". Os resultados dessa aprendizagem, em que se imitam pessoas de alto prestígio, podem ser observados em meninas adolescentes, que procuram vestir-se e comportar-se da mesma forma que estrelas de cinema, e em homens adultos, que compram carros além de suas posses, para se manterem à altura de determinados figurões. Miller e Dollard também aplicam os princípios da imitação ao comportamento de multidões. Sugerem que as turbas consistem em pessoas que se imitam umas às outras e, desse modo, estimulam-se a praticar temeridades que poucas dessas pessoas seriam capazes de cometer na qualidade de indivíduos isolados. Este argumento é ilustrado com a horrível descrição de uma multidão num linchamento.

Medo e neurose

Em seu segundo livro, *Personality and Psychotherapy*, Dollard e Miller (1950) abordam a aprendizagem de características da personalidade e especialmente a aprendizagem e desaprendizagem de neuroses. Começam por assinalar três características da pessoa neurótica: que ela é infeliz por causa de seus conflitos, que estulta a respeito de certos aspectos de sua vida, e que

tem sintomas. E passam a explicar essas três características de acordo com a aprendizagem prévia do neurótico. O elemento básico nesse aprendizado é o impulso de medo, aprendido. Este impulso é a base do conflito, a origem da infelicidade e a causa da estultícia.

Que é um impulso aprendido? Dollard e Miller começam a exposição do tema com uma demonstração experimental em que empregam ratos. Coloca-se um rato numa caixa com dois compartimentos, um deles com paredes brancas e uma grade no piso, e o outro com paredes negras e soalho de madeira. O rato explora ambos os compartimentos e mostra pouca preferência entre um e outro. A seguir, é colocado no compartimento branco, aplica-se-lhe um forte choque elétrico através da grade do piso. A maioria dos ratos logo escapa do choque correndo para dentro do compartimento negro. Repete-se, várias vezes, essa sequência de choque no compartimento branco e fuga para o compartimento negro. Depois, coloca-se o rato no compartimento branco, agora sem choque. O animalzinho corre rapidamente para o compartimento negro. Visto que já não há mais o impulso da dor que motive esse comportamento de fuga, Dollard e Miller explicam-na com base no fato de que o rato aprendeu um impulso de medo, condicionado aos indícios do compartimento branco.

Qual a natureza do impulso aprendido? Como todos os impulsos, ele também, segundo Dollard e Miller, envolve estimulação intensa. Esses estímulos intensos são produzidos pelas próprias respostas do rato. Quando o rato recebeu choque pela primeira vez no compartimento branco, ele apresentou uma série de respostas ao choque, tais como aumentar a tensão dos músculos, aumentar a frequência cardíaca e outras respostas que se consideram indicações involuntárias de emoção. Estas, por sua vez, produziram intensa estimulação, que se somou ao impulso produzido pela dor do choque. Estes estímulos intensos produzidos pelas respostas emocionais formam um impulso de medo. Estas respostas emocionais tornaram-se condicionadas aos indícios do compartimento branco, devido ao fato de que ocorreram em presença desses indícios e foram acompanhadas de uma redução do impulso quando o rato escapou do choque. Agora, quando os indícios do compartimento branco são apresentados sem choque, produzem as respostas emocionais e, por sua vez, os estímulos do impulso de medo.

O medo como impulso pode ser demonstrado usando-se o medo como base de nova aprendizagem. Efetuou-se esta demonstração fechando-se a portinhola que une o compartimento branco ao compartimento negro e possibilitando ao rato abri-la somente girando uma rodinha na parede junto a ela. Os animais aprenderam a girar a rodinha e, depois, correr para dentro do compartimento negro, ainda que não houvesse choque no compartimento branco. Neste caso, o medo foi o impulso, a visão da rodinha foi o indício, girar a rodinha foi a resposta, e a eliminação do medo (escapando do compartimento branco) foi a recompensa. Depois, modificou-se a situação, de tal maneira que a portinhola se abria não quando se girasse a rodinha, mas apertando-se uma alavanca. Em consequência, a resposta de girar a rodinha extinguiu-se e adquiriu-se a resposta de pressionar a alavanca. Desse modo, o impulso de medo aprendido atuou como um impulso primário para motivar aprendizagem.

De que modo esse tipo de aprendizagem ocorre quando passamos dos ratos de experimentação para seres humanos neuróticos? Consideremos uma criança que é severamente punida por qualquer tipo de comportamento

auto-afirmativo. Sempre que tenta fazer as coisas à sua maneira, é submetida a dor e sofrimento, que produzem respostas emocionais, e estas, o impulso aprendido de medo. Estas respostas que produzem medo tornam-se condicionadas aos estímulos presentes na ocasião, inclusive aos indícios que provêm de sua própria auto-afirmação. Como resultado, qualquer comportamento auto-afirmativo passa a produzir medo, enquanto a submissão o reduz. Além disso, o medo não é meramente medo de comportamento auto-afirmativo manifesto, mas é medo também daqueles pensamentos e sentimentos de auto-afirmação concomitantes. Nesse caso, a criança sente medo da auto-afirmação da mesma forma como o rato sente medo do compartimento branco; e a submissão da criança (tanto em pensamentos como em ações) é uma forma de escapar de indícios provocadores de medo, da mesma forma que o é a conduta do rato quando foge para o compartimento negro para escapar do compartimento branco que provoca medo.

Entretanto, o problema da criança é pior que o do rato. Enquanto o rato é capaz de correr para o compartimento negro, seu medo é de curta duração e não destrutura sua vida. A criança, porém, frequentemente é colocada em situações nas quais o comportamento auto-afirmativo poderia proporcionar-lhe coisas de que necessita. Tais situações ocorrem ainda mais frequentemente depois que se torna um adulto. O fato de sentir medo de ser auto-afirmativo é uma grande desvantagem nessas situações. Vê-se num conflito entre o desejo de ter algo e o medo da conduta de auto-afirmação que poderia dar-lhe aquilo de que necessita. Esse conflito — no qual sai perdendo, não importa o que faça — é uma fonte de infelicidade. Se reconhecesse a grande diferença entre a situação atual e aquela em que foi castigado por causa da auto-afirmação, poderia ser capaz de aliviar o medo e resolver o conflito. Entretanto, essa evolução exige que reconheça seu problema e pense a respeito dele. Mas não consegue nem mesmo pensar, pois passou a sentir medo não só de se comportar de modo auto-afirmativo, mas também medo de dizer (aos demais ou a si próprio) que gostaria de ser auto-afirmativo. É este medo de dizer ou mesmo de pensar que gostaria de ser auto-afirmativo que torna estulto seu comportamento. Não consegue produzir as respostas de pensar, que o ajudariam a compreender e a resolver seus problemas. Há, porém, uma série de maneiras pelas quais pode obter algum alívio. Por exemplo, pode tornar-se tão dependente, talvez através de sintomas de uma aparente doença física, que as outras pessoas se sentirão obrigadas a cuidar dele e providenciar-lhe algumas das coisas de que necessita. Esta não seria uma estratégia adotada de modo deliberado, mas uma resposta aprendida através dos efeitos que reduzem o impulso. A "doença" poderia ser chamada sintoma. Entretanto, é apenas uma solução parcial ao problema colocado pelo medo e pelo conflito, interferindo na possibilidade de achar uma solução mais eficaz. Tal indivíduo tem todas as características do neurótico; é infeliz, vive em conflito, comporta-se de modo estulto em relação a seus problemas e tem sintomas.

Como se pode eliminar a neurose através da psicoterapia? Visto que o medo é a causa crucial, sua extinção é o elemento básico da cura. Se ao rato se proporciona experiência suficiente no compartimento branco sem choque, termina por extinguir-se a tendência de executar as respostas que produzem medo. Assim também, se o neurótico pode ser persuadido a emitir respostas de auto-afirmação (ou fazer qualquer outra coisa da qual tenha medo) sob condições nas quais não será punido, seu medo será extinto. Visto que os

indícios para o medo provém das próprias respostas do indivíduo, este gradualmente deve ser induzido a executar essas respostas, inicialmente numa forma muito débil e indireta, depois de modo mais direto e intenso. Assim, nas fases iniciais do tratamento, o neurótico pode dizer timidamente que, às vezes, pensa que poderia dar sugestões proveitosas ao chefe, enquanto, em fases posteriores, pode expressar violento desejo de xingá-lo com variados insultos. A medida que o paciente extingue o medo de manifestar-se de modo auto-afirmativo, torna-se mais capaz de pensar, de modo sensato, em seus conflitos. Visto que falar acerca de um ato e fazê-lo são um tanto semelhantes, o paciente, pela generalização do estímulo, passa a ter menos medo de um comportamento auto-afirmativo manifesto. Desse modo, através de um sistema teórico notadamente diferente, Dollard e Miller chegam a uma conclusão prática muito semelhante ao método do limiar de Guthrie para a eliminação de hábitos emocionais indesejáveis.

Durante os anos decorridos desde que começou a colaborar com Dollard, Miller prosseguiu dando muitas contribuições adicionais, principalmente na área da psicologia fisiológica. Também explorou a possibilidade de que o reforço possa não ser simplesmente uma questão de redução do impulso (Miller, 1963). Em 1975, pelas diversas contribuições à psicologia, Miller recebeu a Medalha de Ouro do ano conferida pela *American Psychological Foundation*. Sob alguns aspectos, não se faz justiça a Miller quando se descrevem mais suas primeiras idéias do que aquelas que expôs nos últimos anos; mas é sua teoria (para distingui-la dos achados experimentais) o que mais deve ser lembrado, pois suas idéias desempenharam um papel importante na história da teoria da aprendizagem.

A TEORIA DEDUTIVA DE HULL

A mais ambiciosa das teorias conexionistas da aprendizagem é a de Clark L. Hull (1884-1952). Professor em Yale, durante longos anos, Hull foi o mais influente teórico da aprendizagem, em sua época. Foi um behaviorista, mas muito mais avançado do que Watson na filosofia da ciência. Antes de se tornar psicólogo, sua formação foi na área da engenharia; algo da perspectiva do engenheiro é evidente no seu desejo de construir uma estrutura acabada, formal e exata da teoria psicológica. Nele podemos ver o pleno florescimento lógico da tradição conexionista do reforço.

O método dos postulados na construção da teoria

O conceito que Hull fazia da teoria ideal era uma estrutura lógica formada de postulados e teoremas, semelhante à da geometria de Euclides. Os postulados seriam afirmações a respeito de diversos aspectos do comportamento. Não seriam leis tomadas diretamente dos experimentos, mas enunciados mais gerais acerca dos processos básicos envolvidos. Da mesma forma que os postulados da geometria, não seriam provados por si mesmos, mas sim tomados como pontos de partida para comprovação. A partir desses postulados, poder-se-ia derivar uma grande variedade de outros enunciados, chamados teoremas. Poder-se-ia provar cada teorema raciocinando-se logicamente a partir de uma determinada combinação de postulados. Esses teoremas apareceriam sob a forma de leis do comportamento.

Até aqui, esse tipo de teoria é simplesmente uma criação lógica. Nada dissemos acerca da eventualidade de o enunciado ser verdadeiro ou falso,

apenas acerca do modo como esses enunciados se concatenam logicamente. Quando um teorema fica provado, o que isso significa é que, se os postulados são verdadeiros, o teorema também deve ser verdadeiro. Para que a teoria tenha algum valor como descrição do mundo real, é necessário comparar os teoremas com as leis reais do comportamento conforme são determinadas por experimentos. Em outras palavras, depois que o teórico determinou pela lógica que os teoremas decorrem dos postulados, deve, então, determinar, pela experimentação, se são verdadeiros. Se são verdadeiros, toda a teoria se confirma; se não o são, a teoria perde sustentação e necessita ser revisada.

Para uma pessoa não familiarizada com a filosofia da ciência, essa abordagem pode parecer retrógrada. Os teóricos começam com postulados que podem ou não ser verdadeiros. Depois, provam, pela lógica, que, se os postulados são verdadeiros, certos teoremas também devem ser verdadeiros. A seguir, determinam, por meio de experimentos, se cada um dos teoremas é realmente verdadeiro. Por fim, usam a verdade ou a falsidade dos teoremas para raciocinar indiretamente acerca da verdade ou falsidade dos postulados. No caso de um teorema revelar-se falso, sabem que pelo menos um dos postulados a partir dos quais o teorema foi provado também deve ser falso, pois logicamente levou a uma conclusão falsa. Em consequência, algum postulado (os teóricos podem ainda não saber qual deles) deve ser modificado de tal maneira que os teoremas que dele decorrem venham a ser verdadeiros. Se, em vez disso, todos os teoremas se revelam verdadeiros, esta verificação aumenta-lhes a confiança no sentido de que todos os postulados são verdadeiros. Entretanto, nunca podem ter a segurança de que os postulados são verdadeiros, pois, às vezes, postulados falsos podem levar a teoremas verdadeiros. A medida que um número cada vez maior de teoremas se revela verdadeiro, os teóricos adquirem maior confiança de que os postulados são verdadeiros, mas sempre há a possibilidade de que um teorema derivado dos postulados termine por se revelar falso.

Embora essa abordagem teórica possa parecer ao mesmo tempo complicada e estranha, ela é realmente bastante parecida com o que fazemos em muitas situações habituais. Consideremos uma professora que está preocupada com o baixo rendimento de determinado aluno, possuidor de um QI elevado. Ela suspeita de que a dificuldade do aluno possa decorrer de dois fatores: (1) medo de competição, e (2) tendência a reagir ao medo "bloqueando-se" e tornando-se incapaz de agir de maneira eficiente. Essas suposições são o equivalente dos postulados. Uma ou outra, ou ambas, podem ser ou verdadeiras ou falsas. Não podem ser testadas diretamente, pois o medo não é passível de observação direta. Entretanto, podem ser testadas indiretamente. Para testá-las, a professora deve, primeiramente, imaginar de que modo se comportaria em determinadas situações um aluno que tivesse esse medo e essa reação ao medo. É esse o processo pelo qual se deduzem teoremas a partir dos postulados. Assim, se os "postulados" da professora a respeito do aluno são verdadeiros, também deve ser verdadeiro que: (1) o aluno sair-se-á melhor nas perguntas da prova se estas lhe forem apresentadas mais de uma forma casual na conversação do que como parte de uma prova; (2) se, ao lhe ser dada a oportunidade de competir por um prêmio, o aluno a recusa; e (3) se o aluno aceita a oportunidade que se lhe dá de fazer uma palestra à classe sobre algum tema de seu interesse. A professora pode tentar qualquer uma dessas abordagens e verificar o que acontece. Estes, então, são três experimentos, cada um deles planejado para testar um dos

"teoremas" da professora. (Supomos que o planejamento seja correto.) Se em cada caso o aluno faz o que a professora predisse, a confiança dela em suas interpretações (isto é, em seus "postulados") aumentará. Ela não pode, porém, excluir a possibilidade de todo o comportamento do aluno se dever realmente a alguma outra combinação de fatores, e que ele termine por fazer algo que absolutamente não se ajuste à interpretação dela. Ela pode adquirir uma confiança cada vez maior de que sua interpretação é correta, mas disso nunca pode ter certeza absoluta.

No caso de virem a confirmar-se as duas primeiras predições, mas não a terceira, ela saberia que seus "postulados" não eram totalmente corretos. Poderia concluir, já que o aluno se recusou a falar à classe mesmo quando não estava em jogo nenhuma competição, que o medo dele não era simplesmente medo de competição, mas de alguma situação em que se sentia ameaçado por fracasso ou desaprovação. Nesse caso, ela poderia procurar situações novas nas quais pudesse testar esse "postulado" reformulado. Desse modo, a interpretação dela acerca do comportamento do aluno seria um processo autocorretivo de fazer e testar suposições.

Neste exemplo, o pensamento da professora e o pensamento de Hull ao construir uma teoria científica são, naturalmente, muito diferentes quanto aos objetivos e à forma lógica. Entretanto, estas são antes diferenças de grau de que de espécie. A professora estava interessada em interpretações muito estritas, envolvendo o comportamento de um único aluno, enquanto que Hull se interessava por interpretações muito mais amplas, envolvendo toda a gama do comportamento humano e animal. Além disso, Hull formulou por escrito seus postulados, de forma detalhada, e testou os teoremas em experimentos controlados, ao passo que a professora simplesmente desenvolveu palpites e testou-os de modo rudimentar. Não obstante, os dois casos basicamente são semelhantes o bastante para mostrar o tipo de raciocínio que Hull seguia ao edificar sua teoria.

Do que acima ficou exposto, deve tornar-se claro que Hull não considerava sua teoria um enunciado final a respeito da natureza da aprendizagem. Antes essa formulação tinha um caráter conjectural, sempre sujeita a revisão, para colocá-la a par dos novos dados e das novas idéias. Essa natureza conjectural era uma característica necessária do tipo de sistema que ele desejou construir. Um conjunto de postulados a partir dos quais se supõe que se derivam muitas leis do comportamento, usando-se a dedução lógica estrita, tem maiores possibilidades de ser errôneo do que um sistema informal que faz prognósticos em menor número ou menos exatos. As teorias de Watson e Guthrie, por exemplo, são suficientemente informais para deixar lugar a dúvidas consideráveis sobre o que elas prognosticariam numa determinada situação. Hull, em contraste, desejou criar uma teoria suficientemente específica, de modo que fosse fácil verificar quando era contrária às provas. E praticamente certo que uma teoria dessa natureza é errônea em alguns aspectos quando formulada pela primeira vez, motivo pelo qual deve estar aberta a modificações à medida que se lhe descobrem os erros.

O plano de Hull consistia em escrever três livros para expor sua teoria. O primeiro destinou-se a apresentar e explicar o sistema de postulados. Esse livro apareceu em 1943, sob o título *Principles of Behavior*. Embora, a partir de então, se tivessem publicado diversas revisões do sistema de postulados, esse livro permanece a pedra fundamental dos trabalhos teóricos de Hull. O segundo livro destinava-se a conter derivações detalhadas do comportamento

do indivíduo numa série de situações. Foi publicado em 1952, sob o título *A Behavior System*. Contém 133 teoremas, com as respectivas derivações dos postulados e a evidência que lhes dá suporte como leis do comportamento. Os temas de que tratam esses teoremas são muito diversos, como, por exemplo, aprendizagem de discriminação, deslocamento no espaço e aquisição de valores. Como seria de esperar, esse sistema de postulados e teoremas é muito mais complexo e técnico do que qualquer um dos outros sistemas que já discutimos.

Inválido durante os últimos anos de sua vida, Hull mal conseguiu terminar seu segundo livro importante. No prefácio, lamentou que nunca pudesse ter escrito o terceiro livro da série, que se destinava a apresentar derivações do comportamento em interações sociais. Quatro meses depois, faleceu. Embora essa trilogia permaneça, pois, incompleta, dá-nos um bom quadro tanto da envergadura dos conceitos de Hull como dos muitos aspectos em que o sistema real deixou de corresponder às suas expectativas. É possível que, se Hull tivesse vivido mais tempo e se tivesse tido melhores condições de saúde durante os últimos anos, muitos dos equívocos de sua teoria teriam sido corrigidos. Da maneira como ela é, devemos julgar ao mesmo tempo seus êxitos e seus fracassos pelas obras que nos deixou.

A análise em quatro estágios

Da mesma forma que a maioria dos teóricos, Hull desejou desenvolver um sistema para prever as variáveis dependentes do comportamento a partir das variáveis independentes. Reconhecendo a variedade de variáveis independentes e dependentes com as quais tinha de lidar, procurou simplificar a tarefa da predição introduzindo variáveis intervenientes. No Cap. 1, vimos de que maneira o emprego dessas variáveis intervenientes torna possível, com um pequeno número de conceitos, sintetizar uma grande quantidade de detalhes. Hull organizou suas variáveis intervenientes num esquema preditivo de quatro estágios. O primeiro estágio consistia nas variáveis independentes a partir das quais fazia a predição; a quarta fase, nas variáveis dependentes para as quais fazia a predição; e o segundo e terceiro estágios, nas variáveis intervenientes que estabelecem a conexão entre aquelas variáveis. Sabendo ele os valores (quantidades ou graus) das variáveis independentes, podia calcular os valores das variáveis intervenientes no estágio 2. Por sua vez, conseguia, a partir desses valores, calcular o valor da variável interveniente no estágio 3, e, a partir deste, por fim, prever os valores das variáveis dependentes. O esboço geral desse refinado esquema é-nos mostrado na Fig. 1, que serve de orientação à explicação.

As variáveis independentes compreendem todas aquelas que podem ser manipuladas diretamente por um experimentador. Podem ser sintetizadas com brevidade, pois são discutidas com maiores detalhes em conexão com as variáveis intervenientes. Algumas das variáveis independentes referem-se à estimulação que o aprendiz recebe no momento, tais como a intensidade do brilho de um sinal luminoso ou a intensidade do choque elétrico. Outras se referem a eventos imediatamente precedentes, tais como o número de horas desde a última refeição do sujeito ou a quantidade de esforço muscular recentemente desenvolvido. Ainda outras referem-se à experiência prévia na mesma situação de aprendizagem, como o número de vezes que o sujeito, anteriormente, executou a resposta a ser aprendida, ou a magnitude da recompensa recebida na última vez que a resposta foi executada. O número

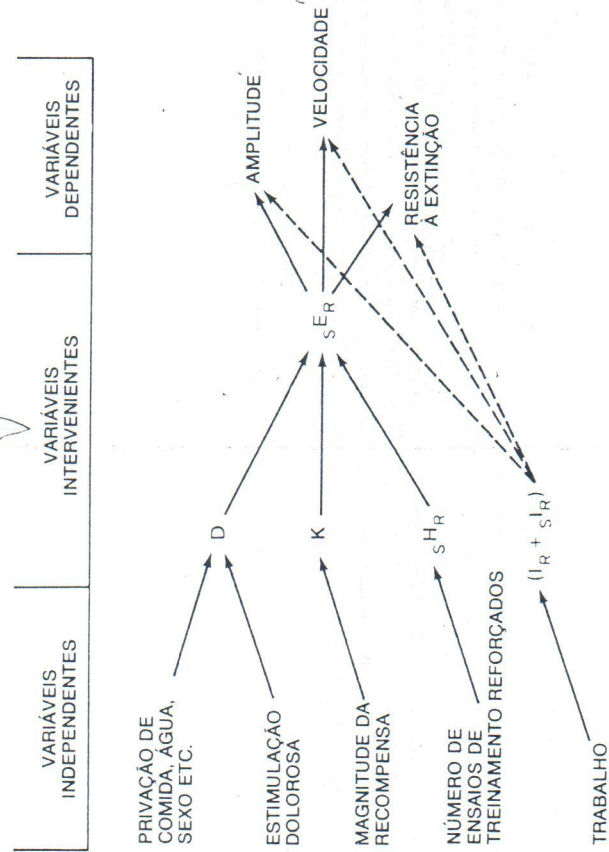


Fig. 1 Representação esquemática do sistema de Hull. As variáveis independentes influenciam as variáveis intervenientes, que, por sua vez, influenciam as variáveis dependentes. As setas em traço contínuo indicam que os aumentos na primeira variável tendem a produzir aumentos na segunda; as setas tracejadas indicam que os aumentos na primeira tendem a produzir diminuições na segunda. Observe que S_R não é um valor exato, mas, antes, a média de uma faixa de oscilação.

dessas variáveis independentes que se poderia enumerar é incalculável; para compreendermos Hull, o que importa é a maneira como ele as organizou em seu sistema.

O segundo estágio da análise introduz variáveis intervenientes. Essas são estados hipotéticos do organismo, que não se pode observar, mas que se supõe diretamente controlados pelas variáveis independentes. As duas mais marcantes, *força do hábito e impulso*, apenas quanto aos detalhes diferem dos correspondentes itens do sistema de Miller, constituído por impulso, início, resposta e recompensa. Força do hábito refere-se à força da conexão aprendida entre um indicio ou indícios e uma resposta, uma conexão formada mediante a prática reforçada. Tanto para Hull quanto para Miller, impulso é um estado de ativação do organismo, e a redução do impulso atua como recompensa.

Visto que Hull é um seguidor da teoria conexionista, a força do hábito é um conceito básico em seu sistema. É a força do vínculo que conecta um estímulo a uma resposta. Para indicar a natureza desta conexão, abrevia-se a força do hábito como S_H (pronuncia-se S H R), com o H significando hábito, e o S e o R subscritos, o estímulo e a resposta que o hábito conecta. Este hábito é uma conexão permanente, que pode aumentar de intensidade, mas não pode diminuir de intensidade. Toda aprendizagem a longo prazo implica a forma-

ção e o fortalecimento de hábitos. Toda vez que uma resposta ocorre em presença de um estímulo e esse evento é rapidamente seguido de reforço, aumenta a força do hábito dessa conexão estímulo-resposta. Dentro dessa linha de suposições (postulado 3, na versão final de seu sistema), Hull assemelha-se a todos os outros teóricos do reforço. A semelhança de Miller, também ele diz que todo reforço envolve uma redução na força de um estímulo de impulso. A rapidez com que a força do hábito se consolida com as sucessivas respostas reforçadas (dada por uma equação no postulado 4) segue a conhecida "lei dos retornos decrescentes", de modo que cada resposta reforçada sucessiva contribui menos para S_H do que a precedente. Por fim, atinge-se um ponto no qual respostas reforçadas adicionais contribuem muito pouco para a força do hábito.

O impulso é um estado temporário do organismo, produzido por privação de algo de que o organismo necessita ou por estimulação dolorosa. Há muitas condições de impulsos específicos diferentes, dos quais a fome, a sede e a dor são exemplos típicos. Os impulsos têm duas funções diferentes. Em primeiro lugar, cada condição de impulso, como fome ou sede, produz um estímulo de impulso característico. Este estímulo indica aquela necessidade particular que afeta o organismo. Uma rápida redução desse estímulo de impulso tem efeito de reforço. Em consequência, toda resposta que ocorre logo antes da redução de um estímulo de impulso tende a ser aprendida como uma resposta a quaisquer estímulos presentes. (Já verificamos esta relação quando examinamos o postulado referente à formação da força do hábito.) A outra função dos impulsos é uma função ativadora ou energizadora. Todas as condições de impulsos combinam-se para formar o nível total de impulsos (ou "drive") do organismo. Este impulso total (cuja abreviatura é D, de "drive") serve para elevar o nível de atividade do indivíduo. Pode-se constatar esse efeito ativador de D tanto num aumento do nível de atividade geral do organismo como num aumento do vigor com que se desempenham todos os hábitos aprendidos. Essa análise é essencialmente a mesma que a de Miller (o que não é surpreendente, pois Miller baseou-se em Hull), mas um tanto mais detalhada e enunciada de modo mais preciso.

Hull também se interessava pela magnitude da recompensa usada como reforçador. Todos nós conhecemos a lei segundo a qual as pessoas tendem a trabalhar com maior afinco por uma recompensa maior do que por uma recompensa menor. Em outras palavras, o nível de desempenho geralmente é mais elevado quando uma recompensa maior é dada após o aparecimento da resposta. De que modo essa lei é abordada pela teoria? Em seu primeiro livro, Hull (1943) abordou a magnitude da recompensa como um dos aspectos do reforço. Quanto maior a recompensa, maior a redução do impulso, e, portanto, maior o aumento da força do hábito. Entretanto, constatou-se que essa abordagem da questão não era satisfatória. Experimentos demonstraram que as modificações na magnitude da recompensa produziam modificações muito rápidas no nível de desempenho, muito mais rápidas do que a lenta formação da força do hábito poderia explicar. Hull teve particular dificuldade de explicar essa diferença quando a modificação se fazia de uma recompensa maior para uma recompensa menor. No caso de uma recompensa se tornar menor e o desempenho tornar-se mais deficiente, esse resultado parecia significar que a força do hábito se tinha tornado menor. Ora, a força do hábito envolvia uma ligação permanente, ligação que se tornava mais forte com a prática reforçada, mas que nunca se tornava mais débil. Essa situação constituía um

chamativo exemplo de discrepância entre teoria e experimento. Como é que Hull abordou esta questão?

A resposta de Hull, nas últimas versões de sua teoria, consistiu em introduzir uma terceira variável interveniente no segundo estágio da análise. Ao lado da força do hábito e do impulso, acrescentou a *motivação de incentivo* (cuja abreviatura é K). A magnitude da recompensa, na versão nova, afeta somente K, não a força do hábito. O nível de K depende da magnitude da recompensa nas experiências imediatamente precedentes. Quando a magnitude da recompensa aumenta, K diminui. Desse modo, a motivação de incentivo refere-se, como o nome dá a entender, ao efeito motivador do incentivo que é proporcionado para que se execute a resposta. A diferença entre motivação de incentivo e reforço é sutil, pois ambos dependem da recompensa, mas a diferença, mesmo assim, é importante. Para que um hábito aumente em força, as respostas devem ser acompanhadas de uma recompensa que reduza o estímulo de impulso. A magnitude da recompensa não faz diferença no que concerne à velocidade com que se forma a força do hábito; todos os reforços têm o mesmo valor. A magnitude da recompensa, porém, realmente afeta o nível da motivação de incentivo. Recompensas grandes correspondem a valores mais altos de K; recompensas menores, a valores menores. Quando o rato aprende a percorrer um caminho para obter alimento, constatamos um gradual aumento da velocidade com que percorre o caminho, aumento que resulta da prática reforçada. Entretanto, além desse lento aumento da velocidade, podemos aumentar ou diminuir a velocidade rapidamente modificando o tamanho da porção de comida que damos ao fim de cada percurso. Da mesma forma, tem-se a possibilidade de aumentar imediatamente a produtividade de um operário oferecendo-lhe melhor paga por produto executado (aumentando-se, com isso, K), enquanto levaria muito mais tempo para aumentar a produtividade treinando a pessoa em melhores métodos de trabalho (aumentando a força do hábito).

Estas três variáveis intervenientes, s_{H_R} , D e K, funcionam conjuntamente para produzir uma outra variável interveniente, que constitui a terceira fase ou estágio da análise. Essa variável chama-se *potencial de reação*; refere-se à tendência total de efetuar determinada resposta a determinado estímulo. É abreviada sob a forma s_{E_R} , de acordo com o mesmo princípio citado para s_{H_R} . É igual ao produto das três outras variáveis intervenientes; por outras palavras, $s_{E_R} = s_{H_R} \times D \times K$. Esta equação significa que a tendência a efetuar uma determinada resposta a um determinado estímulo depende de um hábito formado através da prática reforçada (s_{H_R}) e de dois fatores motivacionais, um dependente de um estado interno (D) e o outro dependente de um incentivo externo (K). Para fins práticos, poder-se-ia fazer referência a s_{H_R} para determinada resposta como "saber fazer", ao K como "saber o que se ganha fazendo" e ao D como "querer a coisa a ser ganha". Hull, porém, não costumava usar essa terminologia, pois a considerava demasiado vaga para a ciência do comportamento exata e objetiva que propugnava. Para Hull, s_{H_R} , D e K eram definidos pelas operações que as produziam (prática reforçada, para s_{H_R} ; privação ou estimulação dolorosa, para D; substância nutritiva, para K), e não por quaisquer nomes vagos ou banais que se pudessem escolher para designá-las.

O quarto e último estágio da análise de Hull é formado pelas variáveis dependentes, pelos aspectos do comportamento que realmente se podem

observar e medir. Hull relaciona três dessas variáveis ao potencial de reação (s_{E_R}): (1) a amplitude ou tamanho da resposta, (2) a velocidade da resposta, e (3) o número total de respostas que irá ocorrer, após a remoção do reforço, antes de completar-se a extinção. Na medida em que s_{E_R} aumenta, também aumentam a amplitude, a velocidade e o número de respostas até a extinção completa.

O que a Fig. 1 mostra, e o que discutimos, é apenas um esboço sumário do sistema de Hull (1952). A versão final de sua teoria tinha 17 postulados e 15 corolários (teoremas que decorrem imediatamente de um só postulado), que ele esperava pudessem cobrir toda a gama dos fenômenos do comportamento que se conhecem. Por exemplo, postulava a existência de *conexões* estímulo-resposta não aprendidas, mas não aprofundou a *pesquisa* desse tema, e raramente mencionou, em seus teoremas, essas conexões inatas. Discutiu o efeito de uma demora entre a resposta e o reforço, determinando um enfraquecimento na tendência à resposta. Entrou em grandes detalhes acerca da generalização e da maneira como se combinam forças de hábito generalizado (isto é, o que acontece quando a mesma resposta é aprendida em relação a dois estímulos diferentes, mas relacionados). Estes detalhes ultrapassam os objetivos deste livro.

Elaborações da teoria

Outros postulados e corolários, entretanto, são necessários para que se compreendam suas derivações. Um destes é a *oscilação*. A oscilação diz respeito à hipótese feita por Hull no sentido de que uma determinada quantidade de potencial de reação não é um valor exato, mas a média de uma distribuição, ao acaso, de valores. Quando se multiplicam em conjunto s_{H_R} e D e K para obter o valor de s_{E_R} , ainda não conseguimos prever com certeza qual será a amplitude ou a velocidade da resposta. O motivo é que a oscilação de s_{E_R} pode ter levado momentaneamente o potencial de excitação a valores acima ou abaixo do valor calculado. A oscilação do valor de s_{E_R} é rápida e ao acaso (segundo, aproximadamente, uma distribuição normal). O fato de Hull achar necessário introduzir este conceito mostra que seu sistema, com toda a sua complexidade, não podia esperar chegar a uma predição perfeita do comportamento. Teve-se de introduzir um elemento de acaso, uma variação imprevisível, para tornar o sistema realístico.

Quando duas respostas incompatíveis foram aprendidas em relação ao mesmo estímulo, de modo que dois s_{E_R} competem por expressar-se, pode-se usar o princípio da oscilação para prever qual dessas respostas irá ocorrer num determinado ensaio. As oscilações dos dois s_{E_R} não são sincrônicas, de modo que, quando um s_{E_R} está momentaneamente elevado, o outro pode estar elevado, médio ou baixo. Se as faixas de oscilação dos dois s_{E_R} se sobrepõem, cada uma das respostas ocorrerá em alguns ensaios. Se o s_{E_R} para uma resposta é consideravelmente mais forte do que o outro, de modo que suas faixas de oscilação se sobrepõem apenas ligeiramente, a mais forte ocorrerá na maior parte das vezes, e a mais fraca, apenas raramente. Se os dois são igualmente fortes, de modo que suas faixas de oscilação coincidem, cada uma das respostas irá ocorrer na metade das vezes. Todas essas situações envolvem conflito entre duas tendências de resposta: o grau de superposição das duas faixas de oscilação possibilita que se prediga o número de vezes que irá ocorrer cada uma das duas respostas competitivas.

Em segundo lugar, há o conceito de *inibição reativa*. Trata-se de uma

tendência a não repetir uma resposta que acabou de ser executada. O montante de inibição reativa que resulta de uma resposta depende do montante de esforço exigido para executar a resposta, de modo que a inibição reativa é aproximadamente o análogo da fadiga. Visto que esta tendência a não responder funciona contra e tende a cancelar os efeitos do potencial de reação (s_{ER}), subtrair-se-á de s_{ER} para se obter o potencial de reação líquido que se emprega ao prever a amplitude e a latência das respostas. A abreviatura de inibição reativa é s_{IR} . O fato de que não há um S subscrito indica que a inibição é contra toda produção dessa resposta, independentemente do estímulo que a produz. O montante total de I_R contra determinada resposta aumenta cada vez que a resposta ocorre, mas diminui com o passar do tempo. É por este motivo que as pessoas geralmente se saem melhor numa tarefa de aprendizagem que envolve grande quantidade de esforço se sua prática é *distribuída*, com breves períodos de prática separados por curtos períodos de repouso. Estes períodos de repouso proporcionam a oportunidade para que I_R se dissipe. Hull também emprega a inibição reativa para explicar por que ocorre a extinção. Uma vez removido o reforço, as respostas sucessivas já não formam mais s_{HR} , mas continuam a formar I_R . Como resultado, a inibição gradualmente ultrapassa s_{ER} , até que não ocorre nenhuma resposta mais. No entanto, depois de um período de tempo, I_R dissipa-se, e a resposta aparece novamente, ensejando a recuperação espontânea.

Há um derivado da inibição reativa que se chama *inibição condicionada*. A abreviatura desta é s_{IR} , que indica um tipo de hábito que conecta estímulo e resposta. Pode-se conceber inibição condicionada como hábito de não responder, produzido por inibição reativa. Podemos conceber I_R como impulso negativo, que pode ser reduzido pela não execução da resposta. Sempre que o indivíduo que esteve respondendo cessa de responder, esta "resposta de parar de responder" é reforçada pelo fato de que I_R começa a dissipar-se. O hábito de parar de responder, que é aprendido deste modo, é s_{IR} . Este s_{IR} explica, entre outras coisas, por que a recuperação espontânea não é completa. A I_R formada durante a extinção dissipa-se com o passar do tempo, mas a s_{IR} aprendida no processo permanece. Vale assinalar que Hull não aborda I_R e s_{IR} uniformemente como impulso e hábito, respectivamente. Em sua equação, simplesmente subtrai tanto I_R como s_{IR} de s_{ER} , de modo que, no final, $s_{ER} = s_{HR} \times D \times K - I_R - s_{IR}$.

Derivação dos teoremas

Dissemos que o objetivo de Hull consistia em derivar as leis do comportamento logicamente a partir de um sistema simples de postulados. Se considerarmos o número de conceitos e hipóteses que Hull agrupou em seus 17 postulados, o leitor poderá pensar que não há nada de simples nesse sistema. No entanto, sua simplicidade relativa torna-se mais evidente quando assinalamos que a *Behavior System* contém 133 teoremas, alguns com diversas partes, todos decorrendo dos 17 postulados. Além disso, esses 133 representam antes as limitações de tempo e energia de Hull do que as limitações da teoria — poderia ter havido muitos outros teoremas mais. Não há dúvida de que o sistema representa uma realização de grandes proporções.

Muitas das derivações de teoremas são bastante complexas, exigindo meticulosa elaboração de relações matemáticas. Para nossos fins, porém, o tratamento não-matemático de uma das derivações mais simples será suficiente para ilustrar a abordagem de Hull. Os teoremas 3 e 4 abordam um caso

especial de aprendizagem de discriminação. Em geral, a aprendizagem de discriminação exige que, dentre dois estímulos, o indivíduo aprenda a responder a um, mas não ao outro. A discriminação poderia envolver um pombo que aprende a dar bicadas numa chave vermelha, mas não numa chave verde, ou poderia envolver uma professora que aprende a ser firme com seus alunos, mas não com os administradores da escola. No caso especial considerado nos teoremas 3 e 4, a tendência a responder a um estímulo é muito mais forte no começo do treinamento do que a tendência a responder ao outro, mas somente as respostas ao último estímulo são reforçadas durante o treinamento. Aprender a discriminar significa que o indivíduo deve aprender a responder ao estímulo que, no início, tem um potencial de reação fraco e não responder ao outro estímulo, que inicialmente tem potencial de reação elevado. Os teoremas predizem a forma da curva de aprendizagem nessa situação.

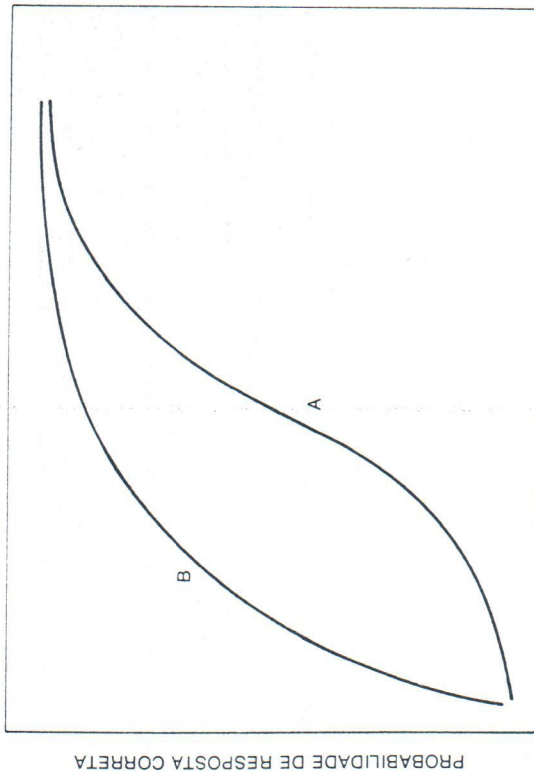
A derivação dos teoremas se faz da seguinte maneira: cada vez que a resposta certa ocorre, aumenta a força do hábito e, portanto, o potencial de reação. Cada vez que a resposta incorreta ocorre, forma-se uma inibição condicionada (Hull não toma conhecimento, nesse ponto, da inibição reativa, supondo distribuição da prática), e o potencial de reação total ($s_{ER} - s_{IR}$) para essa resposta fica diminuído. Entretanto, os dois estímulos são um tanto semelhantes (se não fossem, não haveria discriminação a ser aprendida), e ocorre generalização entre os dois. Quando o indivíduo é reforçado por ter respondido ao estímulo correto, esse reforço aumenta não apenas a força do hábito para a resposta a este estímulo, mas também (num grau muito menor) a força do hábito para a resposta ao outro estímulo. Assim também, quando uma resposta ao estímulo incorreto ocorre e não é reforçada, uma parte da inibição condicionada generaliza-se para a conexão estímulo-resposta correta reduzindo-se o potencial de reação total para essa conexão.

Poderia parecer, a um primeiro exame, que esses dois tipos de generalização se cancelariam mutuamente e não apresentariam diferenças. Entretanto, é necessário levar em consideração o número relativo de vezes que ocorrem respostas aos estímulos corretos e incorretos. No início da aprendizagem, visto que há muito mais s_{ER} para respostas ao estímulo incorreto, as respostas a esse estímulo serão muito mais frequentes. Por conseguinte, a inibição generalizada a partir de respostas incorretas não reforçadas enfraquecerá a tendência correta mais rapidamente do que as respostas corretas reforçadas conseguem fortalecer. Em consequência, enquanto s_{ER} pelas respostas ao estímulo incorreto se enfraquece rapidamente, s_{ER} pelas respostas ao estímulo correto também se enfraquece, embora muito mais lentamente. No entanto, visto que s_{ER} para o estímulo incorreto se reduz mais rapidamente, a tendência a responder ao estímulo correto torna-se *relativamente* mais forte, de modo que aumenta a probabilidade (isto é, a frequência relativa) da resposta ao estímulo correto, em vez de ao estímulo incorreto. Quanto mais frequentemente o sujeito responde ao estímulo correto, mais frequentemente recebe reforço por fazê-lo, e mais rapidamente aumenta seu s_{ER} por fazê-lo. Quanto menos vezes responde ao estímulo incorreto, menos vezes s_{IR} se generaliza a partir do estímulo incorreto para enfraquecer o s_{ER} ao estímulo correto. Por estes dois motivos, s_{ER} para o estímulo correto aumenta numa velocidade maior, à medida que a aprendizagem prossegue.

Quando a resposta ao estímulo correto se torna mais forte, três coisas acontecem: primeiro, a força do hábito para o estímulo correto começa a aproximar-se do limite superior, consoante o princípio dos retornos decres-

centes — desse modo, o potencial de reação para a resposta a esse estímulo aumenta cada vez menos rapidamente; segundo, s_{ER} generaliza-se cada vez mais a partir do estímulo correto para o estímulo incorreto, pois ocorre, agora, maior número de respostas ao estímulo correto do que ao estímulo incorreto; terceiro, como há menos respostas ao estímulo incorreto, s_{IR} forma-se menos rapidamente. Tanto o segundo como o terceiro desses eventos servem para evitar que se enfraqueça o potencial de reação para respostas ao estímulo incorreto. Visto que a tendência correta é aumentar de intensidade mais lentamente, e como a tendência incorreta não é mais decrescente, a probabilidade de responder ao estímulo correto agora aumenta mais lentamente do que antes.

Ao reunirmos todas estas considerações, podemos ver que a probabilidade de responder mais ao estímulo correto do que ao estímulo incorreto aumenta lentamente no início; depois, mais rapidamente e, logo após, de novo lentamente. Esta variação está representada graficamente na curva em forma de S, a curva de aprendizagem, na Fig. 2, que mostra a intensidade da resposta em cada fase da aprendizagem. Curvas de aprendizagem com esta forma são denominadas curvas sigmóides, nome que provém da letra grega sigma. Essa curva sigmóide é diferente da curva típica de aprendizagem, quando a resposta reforçada é originalmente mais forte; e também difere da curva teórica para a formação da força do hábito para uma resposta isolada, dada no postulado 4. Estas duas outras curvas elevam-se mais rapidamente no início e gradualmente declinam, à medida que a aprendizagem progride (Fig. 2).



ENSAIOS

Fig. 2 Duas formas de curvas de aprendizagem. A curva A é sigmóide, forma prevista por Hull para o caso em que a escolha errada tem s_{ER} maior no início do treinamento do que a escolha correta. A curva B tem aceleração negativa, forma prevista por Hull para o caso em que a escolha correta, inicialmente, é mais forte. A curva B assemelha-se à curva do crescimento para s_{ER} de uma só resposta.

Desse modo, a curva de aprendizagem sigmóide é prevista para a aprendizagem da discriminação apenas neste caso especial em que a conexão correta, no início, é mais fraca do que a conexão incorreta, sendo necessária uma linha de raciocínio bastante complicada para efetuar a predição. Entretanto, a predição é coerente com os experimentos. Numa série de diferentes tipos de aprendizagem, verificou-se que dão curvas sigmóides aquelas tarefas nas quais os sujeitos inicialmente se saem mal, enquanto aquelas nas quais os sujeitos, no início, se saem bem, dão curvas com elevações rápidas no início. Desse modo, a predição de Hull, por mais especializada e técnica que seja, encontra confirmação em dados experimentais.

Na derivação dos teoremas 3 e 4, Hull não apresentou simplesmente o raciocínio verbal aqui exposto. Também elaborou um exemplo numérico, calculando os valores reais do potencial de reação e da inibição condicionada e empregando-os para determinar, em cada estágio da prática, a probabilidade de uma resposta correta para um sujeito médio. (Por causa da oscilação, não poderia prever para um determinado sujeito em especial.) Assim, conseguiu dar não apenas a forma geral da curva de aprendizagem, mas as exatas probabilidades para os valores particulares de s_{ER} que usou em seu exemplo.

Virtudes e defeitos do sistema

O valor que Hull atribuiu a sua teoria não estava tanto nas particulares variáveis intervenientes incluídas nesse sistema como na sua rigorosa quantificação. Os postulados não afirmam simplesmente que certas variáveis são relacionadas; dão equações pelas quais uma pode ser calculada exatamente a partir da outra. Por exemplo, o postulado referente à formação de hábito contém a equação $s_{IR} = 1 - 10^{-0.0305N}$, na qual N é o número total de reforços. Usando esta equação, pode-se dizer qual o nível exato de força do hábito para determinado número de reforços. São apresentadas equações semelhantes que vinculam as outras variáveis intervenientes do segundo estágio às variáveis independentes, também vinculando o potencial de reação às variáveis dependentes. A equação que vincula as variáveis intervenientes do segundo estágio a s_{ER} , que envolve multiplicação simples, já foi apresentada.

Nos últimos anos de sua vida, Hull dedicou grande parte de seu trabalho aos problemas de quantificação. Desejava desenvolver uma escala de medida do potencial de reação que fosse aplicável a qualquer resposta. De que modo se pode fazer com que o s_{ER} , por pressionar uma alavanca, seja comparável ao s_{ER} por percorrer um caminho? A resposta de Hull (que terá significado apenas para os leitores que estudaram estatística) consistiu em empregar o desvio-padrão das oscilações de s_{ER} como unidade absoluta de medida. Este emprego do desvio-padrão como unidade de escala é uma idéia que Hull tirou dos especialistas em testeagem psicológica. Neste caso, podia dizer que, para uma determinada resposta, s_{ER} era duas unidades de desvio-padrão acima de zero, e esta afirmação teria o mesmo significado, independentemente de qual a resposta em discussão. Esta unidade também é conveniente para analisar a competição entre respostas, pois, conhecendo-se o s_{ER} para cada resposta em unidades de desvio-padrão, também se sabe automaticamente em que grau suas faixas de oscilação se sobrepõem e, pois, com que frequência cada uma delas ocorre. Hull considerava o desvio-padrão dessa unidade de escala uma das principais realizações de sua vida.

A avaliação da obra de Hull por outros tem sido, porém, bastante diferente da avaliação que ele próprio lhe fez. É justamente na questão do rigor quanti-

tativo que Hull é mais vulnerável a críticas. Em certo grau, o motivo é que suas tentativas de quantificação foram prematuras. Os valores exatos em suas equações, como o valor menos 0,0305 na equação referente à força do hábito, baseavam-se, de modo típico, nos resultados de um único experimento. Ele deu a sugestão, mas nunca a desenvolveu, de que os valores nessas equações poderiam variar de indivíduo para indivíduo. Assim, pessoas diferentes poderiam ter equações diferentes para o desenvolvimento da força do hábito, e o valor absoluto específico seria maior do que 0,0305 para aprendizes rápidos e menor para aprendizes lentos. Além disso, quando Hull passou a derivar teoremas a partir de seus postulados, freqüentemente empregou, em suas equações, valores diferentes dos valores dados nos postulados. Por todos estes motivos, parece claro que esses valores foram propostos mais como valores ilustrativos do que como valores literalmente corretos. O que Hull pretendeu apresentar em seus livros foi o esboço geral de uma rigorosa teoria, juntamente com algumas sugestões sobre a maneira de desenvolver o esboço e transformá-lo num sistema em grande escala, quantitativamente exato. Ao morrer, ainda estava muito longe de realmente haver obtido esse sistema.

Mesmo deixando de lado a questão dos valores exatos, quantitativos, o sistema não corresponde aos ideais que Hull estabeleceu para ele. Num sistema como este de Hull, deveria ser possível, tendo-se os valores das variáveis independentes, calcular os valores das variáveis intervenientes e, a partir destes, os valores das variáveis dependentes. Mesmo não se dispendo dos números exatos necessários para efetuar os cálculos, deveria ser possível explicar de que modo seriam feitos esses cálculos. Em muitos dos casos de Hull, porém, isso não é possível, porque determinado tópico ou é coberto por dois ou mais postulados contraditórios ou absolutamente não é coberto por exemplo deste tipo de problema é a definição de motivação de incentivo (K) quando o impulso é um estímulo doloroso, como o choque elétrico. Presumivelmente, nesse caso, K seria determinado pelo grau de redução da dor que a resposta produzisse, talvez especificado pela diferença em volts entre o choque que recebido antes da resposta e o choque menor recebido depois. Entretanto, visto que K é definido pelo peso da comida ou outro alimento consumido, é impossível aplicar as equações de Hull à situação em que a estimulação dolorosa é o impulso.

Vejamos em detalhe um caso em que se pode aplicar diversos postulados à mesma situação e se pode chegar a previsões conflitantes. Suponhamos que a pergunta seja: "Quantas respostas não reforçadas irão ocorrer antes da extinção completa, quando s_{ER} está num determinado nível ao fim da aquisição?" Há três maneiras de procurar a solução: em primeiro lugar, poderíamos empregar o postulado 16, que relaciona o potencial de reação com o número de ensaios não reforçados necessários para produzir extinção completa. Para usar este postulado, calcularíamos a força do hábito, o impulso e a motivação de incentivo ao fim da aquisição, e multiplicaríamos estes dados para obter o potencial de reação para a resposta no começo da extinção. Depois, empregariamos a equação do postulado 16 para calcular quantos ensaios até a extinção seriam produzidos por esse nível de s_{ER} . Esta abordagem seria a mais direta.

Entretanto, visto que Hull usou a inibição reativa e a inibição condicionada como explicações para a extinção, poderíamos, em vez disso, empregar essas variáveis intervenientes para calcular o número de ensaios até a extinção. Conhecendo o esquema geral do sistema de Hull, talvez conseguíssemos

calcular as quantidades de I_R e s_{IR} formadas por um determinado número de ensaios não-reforçados, que exigem determinado montante de esforço. Entretanto, quando consultamos o postulado 9, que aborda a inibição reativa, logo verificamos que não está formulado de maneira tal que possibilite estes cálculos. Há uma equação que relaciona a quantidade de inibição reativa com o número de ensaios não-reforçados, e uma outra equação que relaciona a quantidade de trabalho com o número de ensaios não-reforçados antes da extinção completa. Nenhuma dessas equações fornece todas as informações que deveríamos ter para fazer pleno uso das variáveis intervenientes da teoria. Poderíamos empregar uma ou outra dessas equações incompletas para calcular a resistência à extinção, e em ambos os casos obteríamos uma resposta diferente daquela obtida com o emprego do postulado 16.

Ainda uma terceira alternativa seria a de empregar o postulado (não-numerado, pois, em sua pressa, Hull deixou-o fora do conjunto de postulados e teve de acrescentá-lo, posteriormente, numa nota de rodapé), que aborda as modificações da motivação de incentivo (K). Não havendo recompensa, K rapidamente se aproximaria de zero. Visto que $s_{ER} = s_{IR} \times D \times K$, também s_{ER} se aproximaria rapidamente de zero. Dentro de alguns ensaios, s_{ER} seria tão pequeno que não ocorreria resposta alguma. Esta abordagem proporciona uma resposta ainda mais diferente das outras duas do que estas eram diferentes entre si. Enquanto ambas as outras abordagens pelo menos levavam à previsão de que a extinção seria gradual, esta abordagem prediz extinção muito rápida. Com isso, mais do que nunca nos assaltam dúvidas sobre o que é que a teoria de Hull realmente pode nos dizer a respeito do tempo que leva para que determinada resposta chegue à extinção.

O simples fato de que estas três formas de responder à questão podem nos dar três respostas quantitativamente diferentes não constitui em si mesmo um problema crucial, pois vimos que os valores numéricos particulares que Hull usou eram principalmente para exemplificação. Entretanto, a falta de consistência interna que esta ilustração revela constitui um defeito grave do sistema, provavelmente mais grave do que qualquer outra falha. Quando uma teoria faz previsões incorretas, ela pode ser modificada, conforme Hull pretendia modificar sua teoria. Quando uma teoria absolutamente não aborda determinada questão, podemos aceitar essa sua limitação de objetivo e esperar que, algum dia, ela possa expandir-se e incluir o tema não abordado. Entretanto, quando uma teoria é internamente inconsistente, ensejando previsões conflitantes acerca de determinada questão, seu valor como teoria rigorosa fica gravemente comprometido. Em certa medida, podemos atribuir a falha de Hull nesse aspecto às revisões em sua teoria. Por exemplo, quando, numa revisão posterior, a variável interveniente K foi acrescentada ao sistema, esta modificação exigia uma série de ajustes em outros postulados, mas alguns destes nunca foram efetuados. Correndo contra o tempo, durante os últimos anos de sua vida, Hull considerou mais importante mostrar o que se poderia fazer com partes de seu sistema do que fazer deste uma unidade completa e consistente. Mesmo assim, devemos reconhecer que o sistema teórico criado por Hull realmente fica muito aquém dos padrões que ele próprio estabeleceu para as teorias.

Por uma ironia do destino, parece que Hull fracassou onde ele mais desejava ter êxito, e teve êxito no aspecto que menos lhe interessava. Queria simultaneamente construir uma teoria dedutiva, a um só tempo ampla nos objetivos e rigorosa nos detalhes, e encorajar outras pessoas a continuarem o

mesmo tipo de trabalho teórico. Vimos em que medida ficou aquém desta sua primeira aspiração. Além disso, seu fracasso provavelmente desencorajou outros de tentarem essa tarefa. Muitos observadores assinalaram, na teoria da aprendizagem, uma tendência que se afasta do tipo de teorias destinadas a abranger tudo, como as que discutimos até agora, passando para teorias que visam a um objetivo menos ambicioso, destinado a explicar apenas certos tipos de aprendizagem. Tal tendência pode decorrer, em parte, do insucesso de Hull. Durante muitos anos, os psicólogos têm sonhado fazer por sua área de investigação o que Newton fez pela Física — desenvolver uma teoria que, a um só tempo, seria ampla em seus objetivos, precisa em suas aplicações e elegante na simplicidade, uma teoria que daria consistência a pontos obscuros na psicologia, formando um sistema soberano. Fazendo esforços tão ingentes por atingir tal objetivo e ainda assim ficando tão aquém dele, Hull convenceu muitos psicólogos de que ainda está muito distante o dia em que irá surgir essa teoria soberana em psicologia.

Mesmo em seu insucesso, Hull também obteve êxito marcante. Seus termos, seus interesses e sua maneira de formular questões psicológicas tornaram-se mais difundidos do que os de qualquer outro teórico de sua época. Mais experimentos provavelmente foram inspirados por seu trabalho do que pelo de qualquer outro teórico da aprendizagem. Suas interpretações do impulso, do reforço, da extinção e da generalização são pontos de partida básicos para discussão destes temas. Foi atacado, defendido e aperfeiçoado ao ponto de que, para muitas pessoas, "teoria da aprendizagem" e "teoria hulliana" se tornaram sinônimos. Somente duas décadas depois de sua morte, a importância de Hull na teoria da aprendizagem cedeu lugar a um outro construtor de sistemas, B. F. Skinner. É o sistema de Skinner que iremos abordar no próximo capítulo.