
ARQUITETURA DE PRODUTOS EDUCACIONAIS

A Metodologia APE

TRANSFORME SEU CONHECIMENTO EM PRODUTOS QUE ENSINAM DE VERDADE



JONATHAN MACHADO

JONATHAN MACHADO

Arquitetura de Produtos Educacionais

A Metodologia APE

Como transformar conhecimento especializado
em produtos que ensinam de verdade

1ª Edição

2026

Arquitetura de Produtos Educacionais

A Metodologia APE

Copyright © 2026 Jonathan Machado

Todos os direitos reservados.

1ª edição, 2026

Revisão: Jonathan Machado

Diagramação: Produção digital

Nenhuma parte deste livro pode ser reproduzida, armazenada ou transmitida por qualquer meio (eletrônico, mecânico, fotocópia, gravação ou outro) sem a permissão por escrito do autor.

Impresso no Brasil / Printed in Brazil

SUMÁRIO

- 1 O PROBLEMA MAIS ANTIGO DA HUMANIDADE: COMO TRANSFERIR O QUE SABE 1

PARTE I · FUNDAMENTOS EPISTEMOLOGICOS

- 2 O CONHECIMENTO QUE NÃO SE DIZ: MICHAEL POLANYI E A EPISTEMOLOGIA DO SILÊNCIO 13
- 3 O ADULTO QUE NÃO QUER SER TRATADO COMO CRIANÇA: MALCOLM KNOWLES E A ANDRAGOGIA 30

PARTE II · EVOLUCAO DA CIENCIA DA APRENDIZAGEM

- 4 COMEÇAR PELO FIM: WIGGINS, McTIGHE E O DESIGN REVERSO 52
- 5 A PIRÂMIDE QUE MUDA COMO SE ENSINA: WILLIAM GLASSER E A RETENÇÃO REAL 69
- 6 O QUE SIGNIFICA SABER: BENJAMIN BLOOM E A ARQUITETURA DA COMPETÊNCIA 87
- 7 A ENGENHARIA DA EXCELÊNCIA: ANDERS ERICSSON E A PRÁTICA DELIBERADA 105

- 8 A ARQUITETURA DO ENSINAR: ROBERT GAGNÉ E O DESIGN INSTRUCIONAL SISTEMÁTICO 122
- 9 A MENTE TEM LIMITES: JOHN SWELLER E A TEORIA DA CARGA COGNITIVA 140

PARTE III · CONSCIENCIA, PERSUASAO, DECISAO

- 10 O MAPA DA MENTE DO COMPRADOR: EUGENE SCHWARTZ E OS NÍVEIS DE CONSCIÊNCIA 159
- 11 A ARTE DE PERGUNTAR PARA VENDER: NEIL RACKHAM E O SPIN SELLING 170

PARTE IV · RETENCAO, VALOR, EXPANSAO

- 12 O CLIENTE COMO RESULTADO: LINCOLN MURPHY E O CUSTOMER SUCCESS 183
- 13 A MÉTRICA QUE REVELA TUDO: FRED REICHHELD E O NET PROMOTER SCORE 194

PARTE V · SÍNTESE EVOLUTIVA

- 14 A CONVERGÊNCIA: QUANDO DOZE PESQUISADORES CONSTRUEM A MESMA RESPOSTA 206
- 15 A TESE: ARQUITETURA DE PRODUTOS EDUCACIONAIS COMO NOVA DISCIPLINA 225
- 16 O EXPERT QUE ESCALA SEM PERDER O QUE É 241

1

O PROBLEMA MAIS ANTIGO DA HUMANIDADE: COMO TRANSFERIR O QUE SABE

Existe um tipo de perda que não aparece em nenhuma estatística. Não tem data. Não tem obituário. Não acumula no registro de nenhuma instituição. É a perda do conhecimento que morre com seu portador, não porque ele não queria transmitir, mas porque nunca soube como.

Todo ofício tem sua versão dessa história. O mestre que sabia exatamente quando o metal estava no ponto certo, não pelo termômetro, mas por uma tonalidade de cor que levou vinte anos para aprender a ler. O cirurgião cuja mão não tremia no momento que a de todos os outros tremeria, e que, perguntado sobre o motivo, encolhia os ombros. O negociador que fechava contratos que ninguém mais conseguia fechar, com um repertório de microgestos e pausas calculadas que nunca havia sistematizado porque nunca havia precisado. Esses homens e mulheres foram excelentes. E levaram consigo, ao se aposentar, o que mais valia.

A humanidade convive com esse problema desde antes de ter palavras para descrevê-lo.

As primeiras tentativas de resolver o problema da escala do saber foram orais. Durante milênios, o conhecimento viajava de boca a ouvido, de mestre a discípulo, numa cadeia frágil que dependia da longevidade de quem sabia, da proximidade física entre quem sabia e quem precisava saber, e da fidelidade de uma memória que, por natureza, transforma o que transmite. As escolas gregas representaram o primeiro experimento sistemático de escala: Platão fundou a Academia em 387 a.C. não como repositório de conteúdo, mas como ambiente de convivência intelectual onde o conhecimento circulava pela disputa, pelo diálogo, pela presença contínua de mentes em atrito produtivo ¹. Era um modelo que funcionava. E que não escalava.

O modelo medieval de transmissão levou esse princípio de proximidade ao extremo institucional. A relação mestre-aprendiz das guildas europeias medievais era, na prática, o único mecanismo disponível para transferir conhecimento de alta complexidade em áreas que iam da marcenaria à medicina ². O aprendiz vivia na casa do mestre, observava, imitava, errava sob supervisão, e gradualmente desenvolvia, por osmose prolongada, aquilo que o mestre possuía. O processo durava anos. Frequentemente, décadas. O gargalo era estrutural: um mestre, por mais generoso que fosse com seu tempo, conseguia formar poucos discípulos ao longo de uma vida. O conhecimento se propagava, mas devagar. Devagar demais para os ritmos que viriam depois.

A Revolução Industrial não resolveu o problema. Ela o reformulou de uma maneira que ainda hoje confundimos com solução.

A demanda por trabalhadores qualificados em escala industrial gerou, no século XIX, a escola de massa. Era uma resposta a um problema real:

¹ MARROU, Henri-Iréné. *Histoire de l'éducation dans l'Antiquité*. Éditions du Seuil, 1948.

² EPSTEIN, Steven A. *Wage Labor and Guilds in Medieval Europe*. University of North Carolina Press, 1991.

precisava-se de muitos operários com competências básicas uniformes, num tempo que as guildas não conseguiam suprir. A escola que emergiu desse contexto, com sua estrutura de turmas, séries, currículos padronizados e avaliações por memorização, foi uma solução engenhosa de manufatura aplicada à formação humana ³. Ela escala extraordinariamente. E, como toda solução que escala, ela simplifica. O que ela simplificou foi exatamente o que não poderia ser simplificado: o problema da transferência de conhecimento complexo, que não é homogêneo, não é passivo, não obedece ao ritmo de um calendário escolar, e não pode ser verificado por uma prova com gabarito.

O modelo permanece. Não porque funciona melhor, mas porque está em todo lugar.

O século XX trouxe a explosão do conhecimento especializado em dimensões sem precedente. A física quântica, a neurociência, o direito tributário, a engenharia de software, a cardiologia intervencionista, cada campo se fragmentou em subespecialidades que, por sua vez, se fragmentaram em sub-subespecialidades. O conhecimento total da humanidade, estimado em termos de publicações científicas, dobra a cada nove anos desde meados do século passado ⁴. O que isso significa, na prática, é que há mais saber disponível do que em qualquer momento anterior da história, e que transferir esse saber tornou-se, paradoxalmente, mais difícil.

Michael Polanyi, químico húngaro que se tornou um dos maiores filósofos do conhecimento do século XX, formulou o problema em 1966 com uma frase que ainda hoje ressoa como o diagnóstico mais preciso já feito sobre o tema: “sabemos mais do que conseguimos dizer” ⁵. O conhecimento que define a competência real de um expert não é o conhecimento

³ TYACK, David; CUBAN, Larry. *Tinkering Toward Utopia: A Century of Public School Reform*. Harvard University Press, 1995.

⁴ BORNMANN, Lutz; MUTZ, Rüdiger. Growth Rates of Modern Science: A Bibliometric Analysis Based on the Number of Publications and Cited References. *Journal of the Association for Information Science and Technology*, v. 66, n. 11, p. 2215–2222, 2015.

⁵ POLANYI, Michael. *The Tacit Dimension*. Doubleday, 1966.

que ele consegue articular em palavras, em manuais, em aulas. É o conhecimento que ele acumulou em anos de prática repetida e de que não tem consciência explícita: o senso de timing, a calibração perceptual, o julgamento situacional que opera antes que a mente verbal tenha tempo de processar. Polanyi chamou isso de conhecimento tácito. E o que ele demonstrou foi que, quanto mais alto o nível de expertise, maior a proporção do que é tácito em relação ao que é explícito. Os experts mais brilhantes de qualquer campo são, com frequência, os que menos conseguem explicar o que fazem.

O paradoxo contemporâneo tem essa forma precisa: o expert que mais tem a transmitir é o que tem mais dificuldade de transmiti-lo.

Não por falta de inteligência. Não por falta de motivação. Por falta de estrutura de externalização. Por ausência de um protocolo que converta o que está operando tacitamente no repertório de um especialista em algo que possa ser transferido com consistência para quem ainda não possui esse repertório. A pergunta que a humanidade nunca respondeu adequadamente não é “como aprendemos”. É “como o que um expert sabe chega, com integridade, a quem ainda não sabe”.

A internet prometeu resolver esse problema. Foi uma promessa honesta, baseada num diagnóstico real.

Se o obstáculo histórico à transferência de conhecimento era a escassez de acesso, então democratizar o acesso deveria resolver o problema. O YouTube, lançado em 2005, e as plataformas de educação online que o seguiram tornaram possível que qualquer pessoa com conexão à internet assistisse a aulas de Harvard, de Stanford, de qualquer universidade que decidisse disponibilizar seu conteúdo. A promessa era de democratização radical: o conhecimento que antes custava dezenas de milhares

de dólares em anuidades universitárias agora poderia ser acessado gratuitamente por qualquer um, em qualquer lugar.

O que aconteceu foi diferente do prometido.

Os MOOCs, os Massive Open Online Courses que se proliferaram a partir de 2012, são o experimento em maior escala já realizado sobre o que ocorre quando o acesso a conteúdo é tornado universalmente disponível. Os dados são consistentes: a taxa média de conclusão de MOOCs oferecidos pelas melhores universidades do mundo oscila entre 12% e 15% ⁶. Em alguns casos documentados, cai abaixo de 5%. Isso significa que, em média, entre 85% e 88% das pessoas que se inscrevem num curso online gratuito da Harvard Extension School ou do MIT OpenCourseWare nunca chegam ao fim. E “chegar ao fim” não é a mesma coisa que adquirir a competência que o curso prometia.

A interpretação mais imediata desse dado é sobre motivação: as pessoas se inscrevem mas não têm disciplina. Essa interpretação é conveniente e incorreta. Ou, mais precisamente, é uma explicação que localiza o problema no lugar errado.

O problema não é motivação do aprendiz. É arquitetura do produto.

A diferença entre conteúdo e conhecimento transferível não é uma distinção filosófica abstrata. É uma distinção operacional com consequências diretas sobre o que acontece no processo de aprendizagem. Conteúdo é o que o expert sabe, organizado na sequência em que aprendeu ou na sequência que considera lógica a partir do ponto de vista de quem já sabe. Conhecimento transferível é o que o aprendiz precisa ser capaz de fazer ao final do processo, construído a partir da distância entre onde esse aprendiz está e onde precisa chegar, com uma arquitetura que respeite como mentes que ainda não dominam o tema processam informação nova. O primeiro design é centrado no expert. O segundo, no

⁶ JORDAN, Katy. Massive Open Online Course Completion Rates Revisited: Assessment, Length and Attrition. *International Review of Research in Open and Distributed Learning*, v. 16, n. 3, p. 341–358, 2015.

resultado do aprendiz.

Quase toda a indústria de educação online produz o primeiro. Quase todos os cursos digitais são registros de conhecimento, não sistemas de transferência. São o expert falando para uma câmera sobre o que sabe, na ordem em que faz sentido para quem já sabe, sem mecanismo de verificação de progressão, sem ciclos de prática estruturada, sem feedback que permita ao aprendiz saber se está no caminho certo. E depois, quando a taxa de conclusão é de 12%, a conclusão é que os alunos não têm disciplina.

Não é sobre disciplina. É sobre design.

Benjamin Bloom demonstrou, num estudo de 1984 que produziu o dado mais citado da história da educação, que a diferença entre tutoria individual e ensino coletivo convencional é de dois desvios padrão de performance ⁷. O aluno mediano com acesso a um tutor individual supera 98% dos alunos num sistema coletivo convencional. Bloom chamou isso de Efeito 2 Sigma. A pergunta que ele colocou ao campo foi: como escalar o que só a tutoria individual entrega? Não obteve resposta suficientemente precisa. O que a indústria de educação online fez, quarenta anos depois, foi construir o equivalente digital do ensino coletivo convencional: mais barato, mais acessível, igualmente incapaz de escalar o que a tutoria individual entrega.

O que torna a tutoria individual eficaz não é, afinal, a presença física do tutor. É o feedback imediato e específico sobre a performance do aprendiz, a adaptação constante do nível de dificuldade ao estado atual de desenvolvimento, e a exigência de que o aprendiz pratique, não apenas consuma. Anders Ericsson passou décadas estudando como experts de alto nível em campos tão distintos quanto xadrez, violino e medicina desenvolveram suas capacidades ⁸. O que ele encontrou foi que a variá-

⁷ BLOOM, Benjamin S. The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, v. 13, n. 6, p. 4–16, 1984.

⁸ ERICSSON, K. Anders; KRAMPE, Ralf T.; TESCH-RÖMER, Clemens. The Role of Deliberate Practice in the Acqui-

vel que diferenciava experts de praticantes medianos não era o tempo total de prática, mas a estrutura da prática: feedback imediato, objetivos específicos, esforço cognitivo intenso em zonas de dificuldade adjacente à competência atual. Sem essa estrutura, décadas de experiência produzem pouco crescimento de competência real.

A maioria dos cursos digitais não tem essa estrutura. Tem vídeos. Tem PDFs. Tem, no melhor dos casos, provas de múltipla escolha que verificam memória e não competência. A plataforma é sofisticada; o design de aprendizagem é o mesmo da aula expositiva que não funcionava em 1962.

O NPS médio da indústria de educação online é 22, num setor onde empresas de alto desempenho operam acima de 50 e os casos excepcionais ultrapassam 70. Um NPS de 22 significa que a proporção de pessoas que recomendariam ativamente um produto educacional para alguém que conhecem é marginalmente maior do que a proporção das que ativamente desaconselhariam. Não é fracasso catastrófico. É mediocridade sistêmica: uma indústria inteira que ainda não resolveu seu problema fundamental de entrega.

O mercado global de e-learning deve atingir US\$ 375 bilhões até 2026⁹. É um mercado que cresceu de forma extraordinária e que, pela evidência do NPS médio, ainda não possui o que mais importa: um modelo sistemático de transferência de conhecimento que funcione. A maioria dos produtos educacionais existentes é boa o suficiente para ser vendida. Não é boa o suficiente para ser indicada por quem fez. Essa diferença é a definição operacional de um mercado que ainda não foi arquitetado.

Este livro é o argumento de que a transferência de conhecimento é

sition of Expert Performance. *Psychological Review*, v. 100, n. 3, p. 363–406, 1993.

⁹ GLOBAL MARKET INSIGHTS. *E-Learning Market Size, Share and Industry Analysis Report*. Global Market Insights, 2022.

uma disciplina de engenharia, não de pedagogia.

A distinção não é semântica. Pedagogia, no sentido corrente do termo, é o conjunto de práticas e reflexões sobre como ensinar. Pressupõe um professor, um aluno, um conteúdo, e uma relação de transmissão entre os três. A metáfora subjacente é a transmissão: quem sabe passa para quem não sabe. Engenharia, no sentido em que este livro usa o termo, é o design sistemático de sistemas que produzem resultados especificados com consistência, independentemente de quem os opera. Aqui a metáfora não é transmissão: é transformação. A diferença entre as duas não é apenas de linguagem. É de quem carrega a responsabilidade pelo resultado.

Num modelo pedagógico convencional, a responsabilidade pelo aprendizado é, em última instância, do aprendiz. Ele não aprendeu porque não se dedicou, porque não tinha o perfil, porque não era o momento certo. Num modelo de engenharia de transferência, a responsabilidade pelo resultado está no design do sistema. Se o sistema não produz o resultado especificado, o sistema precisa ser redesenhado. Não o aprendiz.

Essa inversão tem consequências práticas que percorrem cada decisão de design de um produto educacional: quais resultados o produto especifica como entregáveis; como esses resultados são verificados antes de o produto ser considerado completo; qual é a arquitetura de progressão que leva o aprendiz de onde está até onde precisa chegar; quais mecanismos de feedback informam ao aprendiz e ao designer se o sistema está funcionando; e como o produto itera quando os dados indicam que não está.

A tese central deste livro: experts falham na transferência de conhecimento não por falta de saber, mas por falta de estrutura de entrega. Domínio de conteúdo e capacidade de transferir esse conteúdo são competências distintas. Não apenas distintas: frequentemente inversas. O

expert que mais domina um campo é o que tem maior proporção de conhecimento tácito, o que mais opera por intuição calibrada e julgamento automatizado, e portanto o que mais precisa de um protocolo externo de externalização. Sem esse protocolo, o que o expert produz quando tenta ensinar é uma projeção de como ele aprendeu, não uma arquitetura de como outro perfil de pessoa, com outro ponto de partida, precisaria aprender.

Ao longo de seis décadas, pesquisadores em campos que raramente dialogavam produziram, sem coordenação entre si, um conjunto de descobertas que apontam para a mesma estrutura fundamental. Polanyi, trabalhando em filosofia do conhecimento, identificou que a competência real é majoritariamente tácita e resiste à articulação verbal direta ⁵. Malcolm Knowles, trabalhando em educação de adultos, demonstrou que adultos aprendem de forma radicalmente diferente de crianças: precisam de relevância imediata, orientação por problemas reais, e autonomia sobre o processo ¹⁰. Grant Wiggins e Jay McTighe, trabalhando em reforma educacional, propuseram inverter a lógica de design instrucional: definir primeiro o resultado esperado, depois a evidência de que foi alcançado, só então o conteúdo que levará até lá ¹¹. John Sweller, trabalhando em psicologia cognitiva, mapeou os limites da memória de trabalho e demonstrou que design instrucional que ignora esses limites produz sobrecarga que impede a aprendizagem, independentemente da qualidade do conteúdo ¹². Ericsson demonstrou que a competência real se desenvolve apenas sob condições específicas de prática estruturada que a maioria dos cursos não oferece ⁸.

Nenhum desses pesquisadores estava tentando construir uma metodologia de design de produtos educacionais. Cada um estava respon-

¹⁰ KNOWLES, Malcolm S. *The Adult Learner: A Neglected Species*. Gulf Publishing, 1973.

¹¹ WIGGINS, Grant; MCTIGHE, Jay. *Understanding by Design*. Association for Supervision and Curriculum Development, 1998.

¹² SWELLER, John. Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, v. 12, n. 2, p. 257–285, 1988.

dendo às perguntas do seu próprio campo. E, no entanto, quando se mapeia o que cada um descobriu sobre como o conhecimento se forma, se organiza e se transfere, emerge uma estrutura coerente que nenhum deles articulou sozinho.

A essa estrutura, este livro dá um nome: Arquitetura de Produtos Educacionais.

O framework que organiza essa estrutura é: INPUT → GATEKEEPER → BLACK BOX → OUTPUT → FEEDBACK. Não é uma metáfora. É um modelo funcional que descreve os componentes de qualquer sistema de transferência de conhecimento que funciona. O INPUT é o que entra no sistema: o perfil do aprendiz, seu nível atual de competência, sua necessidade específica de transformação. O GATEKEEPER é o filtro que define se o produto é o instrumento correto para aquele INPUT específico. A BLACK BOX é o conjunto de processos, atividades, ciclos de prática e mecanismos de feedback que produzem a transformação. O OUTPUT é o resultado especificado: o que o aprendiz deve ser capaz de fazer ao final que não conseguia antes. O FEEDBACK é o sistema que verifica se o OUTPUT foi produzido e informa ao design o que precisa ser ajustado.

Quando um desses componentes está ausente ou mal projetado, o produto educacional falha. Não por azar. Por arquitetura deficiente.

O que este livro oferece é a ciência por trás de cada um desses componentes, reconstituída a partir das contribuições dos pesquisadores que, sem saber, estavam construindo partes de um mesmo edifício, e os princípios operacionais que permitem a um expert com domínio profundo em qualquer campo construir um produto educacional que entregue o que promete, com consistência, a aprendizes com perfis distintos, em escala.

Não existe fórmula. Existe estrutura. E a diferença entre as duas é

a diferença entre um truque que funciona uma vez e um sistema que funciona sempre.

O leitor que chega até aqui com a expectativa de um manual de criação de cursos digitais vai encontrar algo diferente. Este livro não ensina a montar um produto educacional. Ensina a pensar como arquiteto de sistemas de transferência de conhecimento, a distinguir entre o que um expert sabe e o que um aprendiz precisa ser capaz de fazer, a projetar a distância entre esses dois pontos como um sistema de engenharia, não como uma sequência de aulas.

É um livro para quem já chegou em algum lugar relevante e quer que o que construiu sobreviva à própria presença. Para quem percebe que seu conhecimento, acumulado ao custo de anos de prática e de erros caros, tem valor para outros mas não sabe como estruturar essa transferência sem perder o que a torna valiosa. Para quem já tentou criar um produto educacional, falhou, e suspeitou que o problema não era o conteúdo.

O problema nunca foi o conteúdo.

PARTE I

Fundamentos Epistemologicos

2

O CONHECIMENTO QUE NÃO SE DIZ: MICHAEL POLANYI E A EPISTEMOLOGIA DO SILÊNCIO

“We can know more than we can tell.” _ “Sabemos mais do que conseguimos dizer.”. Michael Polanyi, *The Tacit Dimension (A Dimensão Tácita)*, 1966 ¹

2.1. CONTEXTO HISTÓRICO

A Europa que emergiu de 1945 não era apenas um continente destruído fisicamente. Era um continente destruído epistemologicamente. A confiança no racionalismo científico como fundamento suficiente para a civilização havia produzido, entre outras coisas, a câmara de gás e a bomba de fragmentação. O positivismo lógico do Círculo de Viena, que pretendia

¹ POLANYI, Michael. *The Tacit Dimension*. Doubleday, 1966. Reeditado por University of Chicago Press, 2009.

reduzir todo conhecimento válido a proposições verificáveis empiricamente, havia dominado a filosofia ocidental desde os anos 1920: a ciência como método universal, a objetividade como único critério de verdade, o sujeito como variável a ser eliminada do processo de conhecimento. Esse programa intelectual não sobreviveu intacto ao que a Europa fez com ele.

O colapso não foi imediato. Mas as fissuras eram visíveis para quem olhasse com atenção.

Karl Popper havia publicado *Logik der Forschung* (“A Lógica da Pesquisa Científica”) em 1934², propondo o falsificacionismo como critério de demarcação entre ciência e pseudociência, o que já desestabilizava o verificacionismo vienense. Thomas Kuhn, em 1962, viria demonstrar que a ciência não avança por acumulação de fatos verificados, mas por rupturas paradigmáticas, num processo carregado de sociologia, psicologia e poder³. Entre esses dois movimentos, Michael Polanyi construiu sua crítica mais silenciosa e, sob certos aspectos, mais radical: a de que o próprio ato de conhecer, em qualquer domínio, envolve uma dimensão pessoal, incorporada e irredutível que o projeto positivista simplesmente ignorava.

Polanyi não era filósofo de formação. Era químico. Nasceu em Budapeste em 1891, numa família judia de intelectuais, e fez sua carreira inicial em química física, trabalhando com adsorção e cinética química em Berlim durante os anos 1920 e início dos 1930. Quando o nazismo tornou insustentável sua permanência na Alemanha, migrou para Manchester, onde se tornou professor de química física. Foi lá que sua trajetória intelectual tomou a virada que o tornaria, retrospectivamente, uma das figuras mais influentes da epistemologia do século XX: precisamente por não ter começado como epistemólogo.

² POPPER, Karl R. *Logik der Forschung*. Julius Springer, 1934. Tradução inglesa: *The Logic of Scientific Discovery*. Hutchinson, 1959.

³ KUHN, Thomas S. *The Structure of Scientific Revolutions*. University of Chicago Press, 1962. 3. ed., 1996.

O que o levou à filosofia foi um problema prático. Como cientista, Polanyi havia observado repetidamente algo que as teorias do conhecimento dominantes não conseguiam explicar: experts em qualquer domínio sabem muito mais do que conseguem articular verbalmente. Não por falha de vocabulário. Não por preguiça intelectual. Por algo estrutural na natureza do conhecimento adquirido em profundidade. Esse hiato entre o que se sabe fazer e o que se consegue dizer sobre o que se sabe fazer foi a pedra sobre a qual Polanyi construiu uma epistemologia inteira.

A disputa intelectual dos anos 1950 e 1960 na qual ele se inseriu opunha, de um lado, o projeto positivista em suas variantes ainda sobreviventes e, de outro, uma série de esforços para devolver ao sujeito cognoscente seu papel constitutivo no conhecimento. Polanyi situou-se nesse segundo campo com uma peculiaridade: diferentemente dos fenomenólogos que partiram de Husserl ⁴ e dos existencialistas que partiram de Heidegger ⁵, ele construiu seu argumento a partir da prática científica concreta, daquilo que cientistas efetivamente fazem quando investigam e descobrem, não daquilo que dizem fazer quando descrevem seu método para os leigos.

Personal Knowledge (“Conhecimento Pessoal”), publicado em 1958 ⁶, é o resultado desse esforço. É um livro denso, demorado, por vezes difícil, que percorre física, química, biologia, matemática, percepção, linguagem e ética para sustentar uma tese única: todo conhecimento é pessoal, no sentido de que envolve o comprometimento ativo de uma pessoa que conhece, e esse comprometimento não pode ser eliminado sem eliminar o próprio conhecimento. *The Tacit Dimension* (“A Dimensão Tácita”), de 1966, é a síntese mais acessível desse programa, condensando em três conferências (as Gifford Lectures de 1964, série de renome internacional

⁴ HUSSERL, Edmund. *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*. Max Niemeyer, 1913. Tradução inglesa: *Ideas: General Introduction to Pure Phenomenology*. George Allen & Unwin, 1931.

⁵ HEIDEGGER, Martin. *Sein und Zeit*. Max Niemeyer, 1927. Tradução inglesa: *Being and Time*. Harper & Row, 1962.

⁶ POLANYI, Michael. *Personal Knowledge: Towards a Post-Critical Philosophy*. University of Chicago Press, 1958.

dedicada à filosofia da religião e do conhecimento) o argumento central que o tornaria canônico.

2.2. O PROBLEMA QUE POLANYI ESTAVA RESOLVENDO

A formulação do problema exige precisão. Não se trata simplesmente de observar que “algumas coisas são difíceis de explicar”. Esse seria um problema pedagógico, solúvel por melhores professores ou materiais didáticos mais elaborados. O problema que Polanyi identificou é de outra ordem: o conhecimento que os experts possuem não é, em sua maior parte, conhecimento de proposições que eles simplesmente não verbalizaram ainda. É conhecimento de um tipo diferente, que não assume forma proposicional, que não pode ser reduzido a declarações verificáveis sem perder o que o torna eficaz.

A observação inaugural, relatada em diversas variações ao longo de sua obra, é a do reconhecimento de rostos. Somos capazes de reconhecer um rosto entre milhares, com velocidade e precisão extraordinárias, mesmo em condições degradadas de iluminação ou a partir de ângulos não-canônicos. Mas quando tentamos descrever como fazemos isso, o que produzimos é generalidades vagas (“formato oval”, “olhos espaçados”, “nariz levemente achatado”) que, se tomadas como especificações, permitiriam construir inúmeros rostos completamente diferentes do que estávamos reconhecendo. As características que articulamos não são as que usamos para reconhecer. O processo efetivo está em algum lugar que o relato verbal não acessa.

O mesmo vale para habilidades motoras complexas. Ninguém aprende a andar de bicicleta lendo um manual de física que explica o princípio do momentum angular (a quantidade de movimento de rotação que mantém

a bicicleta estável) e a necessidade de inclinar o guidão na direção contrária à queda iminente. Não porque a física esteja errada: ela está correta. Mas dominar essa física proposicional não produz equilíbrio. Equilíbrio é produzido por algo diferente, que não passa pela proposição, que se instala no corpo por repetição e que, uma vez instalado, opera sem acesso consciente ao conjunto de ajustes microscópicos que o sustenta.

Polanyi estendeu essa observação, que poderia parecer limitada a habilidades motoras, a toda a cognição humana de alta competência. Um médico experiente que examina um raio-X e reconhece imediatamente uma anomalia que estudantes de medicina treinados demoram minutos para encontrar, ou não encontram, não está aplicando conscientemente uma lista de critérios diagnósticos. Está percebendo um padrão que anos de exposição calibraram seu sistema perceptivo a detectar. O conjunto de critérios que ele poderia articular não reproduz esse processo: é, na melhor das hipóteses, uma racionalização retrospectiva parcial de algo que aconteceu de outra forma.

A relevância filosófica vai além da descrição de casos. O projeto positivista pressupunha que o conhecimento científico genuíno era público, replicável e transmissível por meio de proposições. Polanyi demonstrou que a prática científica efetiva depende de um substrato de habilidades perceptivas, julgamentos tácitos e heurísticas adquiridas que não são transmitidas por proposições, que se adquirem por convívio com praticantes experientes e que constituem, paradoxalmente, a condição para que as proposições explícitas façam sentido. Um estudante de física que aprende as equações sem o contexto de laboratório onde essas equações têm referentes concretos possui algo, mas não o que um físico possui quando usa as mesmas equações.

O problema tem duas dimensões que Polanyi nunca separou completamente. De um lado, uma dimensão epistemológica: qual é a estrutura

do conhecimento humano, considerando que ele não se reduz ao que pode ser dito? De outro, uma dimensão que é, no fundo, uma pergunta sobre pedagogia: como se transmite o que não se diz? É essa segunda questão que tornará Polanyi indispensável para a Arquitetura de Produtos Educacionais, e é ela que a seção seguinte examina em sua estrutura mais precisa.

2.3. IDEIA CENTRAL

2.3.1. A Frase e o Que Ela Nega

“*We can know more than we can tell*” (sabemos mais do que conseguimos dizer). A frase é deceptivamente simples. Sua radicalidade só aparece quando se percebe o que ela nega.

Ela nega que o ideal do conhecimento seja o conhecimento explicitável. Nega que o objetivo da aprendizagem seja sempre produzir alguém capaz de articular verbalmente o que sabe. Nega que a diferença entre o expert e o iniciante seja, em última análise, quantitativa (o expert sabe mais proposições) em vez de qualitativa (o expert tem um tipo diferente de relação cognitiva com o domínio). E nega, consequentemente, que a transmissão de conhecimento se reduza à transmissão de informação.

O que Polanyi propõe em lugar disso é uma teoria da estrutura do conhecimento que distingue dois polos em toda relação cognitiva: o polo *focal* (aquilo sobre o qual a atenção está diretamente dirigida) e o polo *subsidiário* (aquilo que está presente na consciência apenas como suporte funcional, sem ser o objeto direto de atenção) ¹. Quando um cirurgião segura o bisturi, sua atenção focal está no tecido que está incidindo: a incisão, a profundidade, a reação dos tecidos circundantes. Mas seus dedos, o cabo do instrumento, a pressão que exerce, a posição do pulso,

tudo isso está presente em sua consciência de forma subsidiária. Ele não ignora seus dedos; usa-os como extensões perceptivas. Se, por algum motivo, sua atenção focal se deslocar para os próprios dedos e a sensação de pressão no cabo, a habilidade cirúrgica se degrada. O expert opera *com* a ferramenta, não *sobre* ela.

Essa estrutura focal-subsiária é a que Polanyi usa para explicar o conhecimento tácito. O que é subsidiário num ato de cognição não é inconsciente no sentido psicanalítico: o cirurgião pode, se questionado, descrever a sensação em seus dedos. Mas enquanto opera, esse nível de consciência funciona como suporte invisível para o nível focal. É esse suporte que não se transmite por proposições: o iniciante que leu todas as descrições disponíveis de como segurar o bisturi ainda não tem o que o cirurgião tem, porque o que o cirurgião tem não é uma representação verbal, mas uma calibração sensoriomotora que se produziu por meio de repetição orientada.

2.3.2. As *Dimensões do Conhecimento Tácito*

A distinção focal-subsiária abre para uma questão mais detalhada: o tácito não é monolítico. Polanyi não desenvolveu uma taxonomia formal, mas sua obra sugere ao menos três dimensões que analistas posteriores [7, 8] tornaram explícitas, e que têm implicações distintas para quem projeta sistemas de transferência de conhecimento.

A primeira pode ser chamada de dimensão *funcional*: o conhecimento tácito que se manifesta na capacidade de realizar operações sem acesso consciente aos seus mecanismos, como andar de bicicleta, reconhecer rostos ou falar a língua materna sem conjugar conscientemente os verbos. Essa é a dimensão mais óbvia e a que Polanyi ilustra com mais exemplos concretos.

A segunda é mais sutil: o tácito como estrutura de percepção e rele-

vância. O expert não apenas sabe fazer mais coisas; ele percebe o campo de problemas de forma diferente do iniciante. Onde o iniciante vê uma situação caótica de variáveis, o expert percebe padrões e hierarquias de relevância. Essa diferença perceptiva não é simplesmente produto de memória de mais casos: é uma reorganização de como o campo se apresenta, que guia o que merece atenção e o que pode ser ignorado. Aprender uma disciplina é, em grande parte, aprender a perceber como os praticantes dessa disciplina percebem, e isso não se transmite por declaração direta.

A terceira dimensão é aquela que Polanyi articula a partir da distinção do filósofo britânico Gilbert Ryle ⁷ entre conhecimento *de como* (saber fazer) e conhecimento *de que* (saber que algo é o caso), aprofundando-a: a diferença entre saber que determinado processo químico ocorre sob certas condições e saber conduzir experimentos químicos. Os dois tipos de conhecimento são distintos, adquiridos por processos distintos e não redutíveis um ao outro. Um estudante pode conhecer todos os fatos sobre reações de oxidação sem saber conduzir uma titulação, e um técnico de laboratório altamente competente pode conduzir titulações com precisão sem conseguir explicar a estequiometria subjacente.

A implicação que percorre as três dimensões é sempre a mesma: se o conhecimento de competência genuína tem uma dimensão tácita irreduzível, nenhuma sequência de proposições, por mais bem estruturada, é suficiente para transmiti-la. A transmissão exige algo mais: tempo, prática, retorno corretivo, convívio com praticantes, exposição a casos com complexidade crescente. Isso não é intuição educacional; é uma conclusão que decorre da estrutura do conhecimento, e é exatamente o que os desenvolvimentos posteriores das ideias de Polanyi tentaram operacionalizar.

⁷ RYLE, Gilbert. *The Concept of Mind*. Hutchinson, 1949.

2.4. EVOLUÇÃO POSTERIOR

2.4.1. *Nonaka e Takeuchi: Do Indivíduo à Organização*

O salto mais consequente na aplicação das ideias de Polanyi foi dado pelos pesquisadores japoneses Ikujiro Nonaka e Hirotaka Takeuchi em *The Knowledge-Creating Company* (“A Empresa Criadora de Conhecimento”), publicado em 1995⁸. A proposta era transpor o problema epistemológico para o nível organizacional: como as empresas, em particular as japonesas que estavam dominando mercados globais nos anos 1980 e 1990, criavam e sustentavam vantagem competitiva pelo uso eficaz de conhecimento?

A resposta passou pela distinção polyaniana entre tácito e explícito, reconfigurada como dois eixos de um processo dinâmico de conversão. O modelo SECI (acrônimo formado pelas iniciais de seus quatro movimentos) descreve quatro formas de conversão do conhecimento: *Socialização* (tácito para tácito, via observação e convívio direto); *Externalização* (tácito para explícito, via articulação, metáforas e modelos); *Combinação* (explícito para explícito, via documentação e sistematização); e *Internalização* (explícito para tácito, via prática que automatiza o que foi explicitado). Para os propósitos deste livro, a categoria central é a externalização: o processo pelo qual conhecimento tácito é transformado em formas articuláveis, ainda que aproximadas e incompletas. Nonaka e Takeuchi identificaram que as empresas japonesas criavam condições específicas para essa conversão por meio de metáforas, analogias, protótipos e ambientes físicos e culturais de criação colaborativa (o que chamavam de *ba*, termo japonês que designa um espaço compartilhado de emergência de relações e conhecimento). A externalização não produz conhecimento explícito completo: produz aproximações que funcionam como andai-

⁸ NONAKA, Ikujiro; TAKEUCHI, Hirotaka. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press, 1995.

mes para a transmissão de algo que, em última instância, precisará ser reinternalizado pela prática.

O modelo SECI foi amplamente adotado e, igualmente, amplamente criticado. Tsoukas ⁹ argumentou que Nonaka e Takeuchi haviam reificado o conhecimento tácito, tratando-o como substância extraível e transferível, perdendo a dimensão relacional e performática que Polanyi havia enfatizado. A crítica é pertinente, e sua relevância para os limites das ferramentas de externalização da APE será retomada na seção 7.

2.4.2. Davenport, Prusak e Terra: Gestão do Conhecimento como Campo

Thomas Davenport e Laurence Prusak, em *Working Knowledge* (“Conhecimento em Ação”) de 1998 ¹⁰, sistematizaram o que havia se tornado um campo emergente e identificaram as condições organizacionais que dificultam ou facilitam o fluxo de conhecimento dentro e entre empresas. Sua análise das “fricções” que impedem o conhecimento de circular (a tendência ao acúmulo e retenção individual de informação, a falta de confiança entre áreas, a ausência de linguagem compartilhada, a dificuldade de codificação) é um mapa das falhas que a APE foi projetada para endereçar no contexto de experts individuais. Um ponto central, consistente com Polanyi: o conhecimento mais valioso tende a ser o mais difícil de gerenciar, porque sua profundidade e eficácia dependem da dimensão tácita que resiste à documentação. Sistemas de gestão de conhecimento que se limitam à captura de documentos e procedimentos explícitos tendem a capturar exatamente a parte menos determinante da competência organizacional.

⁹ TSOUKAS, Haridimos. Do We Really Understand Tacit Knowledge? In: EASTERBY-SMITH, Mark; LYLES, Marjorie A. (Orgs.). *The Blackwell Handbook of Organizational Learning and Knowledge Management*. Blackwell, 2003. p. 410–427.

¹⁰ DAVENPORT, Thomas H.; PRUSAK, Laurence. *Working Knowledge: How Organizations Manage What They Know*. Harvard Business School Press, 1998.

No contexto brasileiro, a contribuição mais sistemática foi a de José Cláudio Terra, cujo *Gestão do Conhecimento: O Grande Desafio Empresarial*¹¹, publicado em 2000 com base em pesquisa empírica com empresas brasileiras, documentou como o conhecimento tácito se manifesta nas especificidades do ambiente empresarial do país. Terra identificou que as práticas informais de transmissão (o aprendizado “pegando junto”, a mentoria não estruturada, os rituais de socialização profissional) desempenhavam um papel crítico que estratégias formais de documentação não conseguiam substituir.

2.4.3. Os Críticos e os Refinamentos

O projeto de Polanyi não ficou sem contestação. Harry Collins¹², em seu programa de estudos sobre expertise, argumentou que a distinção entre tácito e explícito é menos nítida do que Polanyi sugeria e que muitas formas de conhecimento que parecem tácitas são, na verdade, conhecimento explícito que simplesmente não foi articulado, seja por falta de incentivo, seja por limitações culturais. A tipologia de Collins entre “fraco tácito” (articulável mas não articulado), “médio tácito” (articulável com dificuldade considerável) e “forte tácito” (genuinamente irreduzível à articulação) foi um refinamento útil que evita o risco de usar “conhecimento tácito” como categoria vaga para qualquer coisa difícil de descrever.

Hubert e Stuart Dreyfus¹³ construíram sobre Polanyi complicando-o produtivamente: seu modelo de aquisição de expertise em cinco estágios (novato, iniciante avançado, competente, proficiente e expert) sugere que a dimensão tácita cresce progressivamente com o nível de expertise. O novato opera com regras explícitas que o expert transcendeu,

¹¹ TERRA, José Cláudio Cyrineu. *Gestão do Conhecimento: O Grande Desafio Empresarial*. Negócio Editora, 2000.

¹² COLLINS, Harry. *Tacit and Explicit Knowledge*. University of Chicago Press, 2010.

¹³ DREYFUS, Hubert L.; DREYFUS, Stuart E. *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. Free Press, 1986.

não porque as esqueceu, mas porque os padrões de reconhecimento que desenvolveu permitem respostas mais rápidas e adaptáveis do que a aplicação deliberada de regras poderia produzir. Isso tem implicações diretas para o design instrucional: o que se ensina a um iniciante não pode ser simplesmente o que o expert faz, porque o expert opera em um nível de integração cognitiva que pressupõe a trajetória inteira até ali.

2.5. IMPACTO REAL NO MUNDO

2.5.1. *Gestão do Conhecimento como Indústria: Promessa e Limite*

O impacto mais direto e mensurável das ideias de Polanyi, mediadas por Nonaka e Takeuchi e por Davenport e Prusak, foi a criação de um campo inteiro de prática gerencial. A gestão do conhecimento tornou-se, nos anos 1990 e 2000, uma das áreas de maior investimento corporativo em tecnologia e consultoria: estimativas do mercado global apontavam para mais de 1 trilhão de dólares investidos entre 1995 e 2010¹⁴, com crescimento sustentado.

Os resultados foram mistos, por razões que Polanyi poderia ter previsto. A maior parte do investimento foi em sistemas de gerenciamento de documentos, bancos de dados de melhores práticas e portais corporativos de informação, todos endereçando o conhecimento explícito que já estava articulado. O conhecimento mais valioso, o que permanecia tácito, ficava sistematicamente fora do alcance dessas ferramentas. Organizações que investiram em práticas de transmissão tácita (mentoria estruturada, comunidades de prática, rotação supervisionada de cargos com acompanhamento direto de profissionais experientes) reportaram resultados significativamente melhores. Essas práticas eram mais caras,

¹⁴ DALKIR, Kimiz. *Knowledge Management in Theory and Practice*. Elsevier, 2005.

mais lentas e mais difíceis de escalar do que sistemas de informação.

2.5.2. O Problema da Sucessão

O que se passou a chamar de “risco de conhecimento crítico” é a versão mais concreta e mais cara do problema de Polanyi: o que acontece quando especialistas com décadas de expertise saem da organização. Estudos do APQC (American Productivity and Quality Center, centro norte-americano de referência em produtividade e gestão do conhecimento) ¹⁵ indicam que organizações perdem sistematicamente entre 30% e 70% do conhecimento efetivo de posições-chave nos processos de sucessão, mesmo quando há documentação formal extensiva. A documentação não captura a dimensão tácita: o saber reconhecer problemas antes de se tornarem crises, o repertório de julgamentos situacionais que só os anos de prática acumulam.

Hospitais enfrentam esse problema de forma especialmente aguda na transição de residências médicas. Estudos documentaram aumentos mensuráveis de erros nos períodos de troca de equipes, quando o conhecimento tácito que médicos experientes haviam desenvolvido sobre casos específicos de pacientes não estava disponível para seus substitutos ¹⁶. O relatório escrito não capturava o que o médico sabia sobre como aquele paciente respondia a determinados tratamentos. Forças armadas de vários países investiram substancialmente em programas de captura de conhecimento tácito de veteranos, em domínios como reconhecimento de ameaças e tomada de decisão sob pressão, onde a dimensão tácita é especialmente pronunciada e especialmente difícil de simular em treinamento formal ¹⁷. Escritórios de advocacia e firmas de consul-

¹⁵ AMERICAN PRODUCTIVITY AND QUALITY CENTER (APQC). *Retaining Valuable Knowledge: Proactive Strategies to Deal with a Shifting Workforce*. APQC Publications, 2002.

¹⁶ PHILIBERT, Ingrid et al. New Requirements for Resident Duty Hours. *JAMA: The Journal of the American Medical Association*, v. 288, n. 9, p. 1112–1114, 2002.

¹⁷ KLEIN, Gary A. et al. The Recognition-Primed Decision (RPD) Model: Looking Back, Looking Forward. In: ZSAMBOK, Caroline E.; KLEIN, Gary A. (Orgs.). *Naturalistic Decision Making*. Lawrence Erlbaum Associates, 1997.

toria enfrentam o problema em sua versão mais diretamente comercial: o valor que esses negócios entregam é, em grande parte, o julgamento experiente de seus sócios seniores, que não está em nenhum manual e que a estrutura piramidal tradicional de pessoal transmite por convívio, em escala estruturalmente limitada.

2.5.3. Na Educação Formal

A recepção das ideias de Polanyi na pedagogia formal foi mais lenta do que na gestão empresarial, mas produziu consequências duradouras. A aprendizagem baseada em problemas (problem-based learning ou PBL, método em que os alunos aprendem a partir do contato direto com problemas reais ou simulados, em vez de absorver teoria antes de aplicá-la), desenvolvida originalmente na Escola de Medicina de McMaster nos anos 1960¹⁸, é parcialmente fundamentada na premissa de que a competência médica tem uma dimensão tácita que só se desenvolve pelo contato com casos reais, não pela assimilação de teoria previamente organizada. O sucesso do modelo, hoje adotado em centenas de escolas de medicina no mundo, é uma evidência indireta da validade da premissa.

Os estágios supervisionados e os programas de aprendizado formal junto a mestres do ofício, em profissões como arquitetura, engenharia e direito, são institucionalizações do mesmo princípio: o reconhecimento de que a formação profissional exige convívio direto com praticantes experientes, algo que a instrução formal, por mais rigorosa, não substitui. Essas práticas existiam muito antes de Polanyi as ter teorizado; ele ofereceu o quadro conceitual que permite entender por que elas funcionam. E por que alternativas puramente proposicionais não as substituem.

¹⁸ p. 285–292.
BARROWS, Howard S.; TAMBLYN, Robyn M. *Problem-Based Learning: An Approach to Medical Education*. Springer, 1980.

2.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

Polanyi fornece o diagnóstico; os autores que se conectam a ele fornecem, cada um a seu modo, instrumentos de resposta. A relação não é de simples extensão: há tensões produtivas que precisam ser reconhecidas para que a integração seja honesta.

2.6.1. *Polanyi + Ericsson: o Tácito é Criado pela Prática Deliberada*

A relação entre o conhecimento tácito de Polanyi e a prática deliberada de Ericsson ¹⁹ é de complementaridade precisa, embora os dois autores pertençam a tradições completamente distintas (filosofia e psicologia cognitiva experimental) e raramente se citem diretamente. Polanyi descreve *o que* o conhecimento de expertise é em sua estrutura: subsidiário, incorporado, resistente à verbalização, organizado em torno de percepções de padrão que o iniciante ainda não possui. Ericsson descreve *como* esse estado é alcançado: pela prática focada em aspectos específicos do desempenho que estão abaixo do nível de automatização atual do praticante, com retorno corretivo imediato e incremento progressivo de dificuldade ²¹. O resultado de anos de prática deliberada é precisamente o conhecimento tácito que Polanyi descreve: uma reorganização profunda do sistema cognitivo e perceptivo que habilita o praticante a perceber e responder ao campo de problemas de formas que o iniciante não consegue.

A tensão é real: Polanyi enfatiza a irredutibilidade do tácito (não há atalho, não há descrição que substitua a prática), enquanto Ericsson identifica que a prática deliberada não é qualquer prática, mas prática estruturada, o que abre a possibilidade de acelerar a aquisição quando as condições são bem projetadas. Para a APE, essa tensão é constitutiva: o

¹⁹ ERICSSON, K. Anders; POOL, Robert. *Peak: Secrets from the New Science of Expertise*. Houghton Mifflin Harcourt, 2016.

objetivo não é encontrar um atalho para o tácito, mas criar as condições de prática que tornam a trajetória de aquisição mais eficaz.

2.6.2. Polanyi + Klein: o Problema Tácito como Ponto de Partida da Tomada de Decisão Naturalística

Gary Klein desenvolveu a Naturalistic Decision Making (Tomada de Decisão Naturalística, ou TDN), campo que investiga como pessoas experientes decidem em condições de pressão e incerteza real: bombeiros, comandantes militares, enfermeiras de Unidade de Terapia Intensiva²⁰. O que encontrou foi consistente com Polanyi: experts não geram opções e as avaliam segundo critérios explícitos. Reconhecem situações como instâncias de padrões familiares e geram diretamente cursos de ação calibrados pela experiência. O *Critical Decision Method* (Método de Decisão Crítica, ou MDC) que Klein desenvolveu para extrair esse conhecimento é, em certo sentido, a operacionalização metodológica do problema de Polanyi: reconhece que perguntar diretamente ao expert “como você decide?” produz racionalização, não descrição fiel do processo. Por isso usa técnicas indiretas: reconstrução retrospectiva de casos críticos, questionamento de alternativas não escolhidas, identificação de pistas perceptivas que orientaram a decisão. É o método de acessar o polo subsidiário sem exigir que o expert o torne focal, o que, como Polanyi havia demonstrado, degrada o desempenho em vez de descrevê-lo.

A tensão aqui é metodológica: Klein é mais otimista sobre a possibilidade de externalização. O MDC assume que é possível, por entrevista estruturada, trazer à superfície explícita uma porção significativa do tácito do expert. Polanyi seria mais cético: o que se externaliza são aproximações e reconstruções, não o processo real, que permanece em seu núcleo resistente. Para a APE, ambas as posições são instrumentalmente corre-

²⁰ KLEIN, Gary A. *Sources of Power: How People Make Decisions*. MIT Press, 1998.

tas: a externalização via MDC é melhor do que nenhuma externalização, e o resultado ainda é incompleto.

2.6.3. Polanyi + Gagné: a Condição que Precede o Design Instrucional

Robert Gagné, ao desenvolver os Nove Eventos de Instrução e a taxonomia de tipos de resultados de aprendizagem ²¹, criou o quadro mais influente do design instrucional (planejamento sistemático de experiências de aprendizagem) do século XX. O pressuposto implícito de Gagné é que os objetivos de aprendizagem podem ser especificados com precisão suficiente para orientar o design de sequências instrucionais. Isso funciona bem para conhecimento declarativo e habilidades procedurais relativamente simples.

O problema aparece quando o objetivo é transmitir competências de expert em domínios complexos: cirurgia, diagnóstico clínico, criação de produtos, gestão de crises. Nesses casos, aplicar os Nove Eventos pressupõe que o conteúdo a ser ensinado já foi identificado e explicitado, o que é precisamente o que o problema de Polanyi coloca em questão. Gagné sem Polanyi produz design instrucional para o que o expert consegue articular, não para o que constitui o núcleo de sua competência. A sequência correta é: primeiro extrair (Polanyi via Klein), depois estruturar (Gagné). Inverter essa ordem não simplifica o processo: esconde o problema mais custoso até o momento em que ele se torna irreparável.

²¹ GAGNÉ, Robert M. *The Conditions of Learning and Theory of Instruction*. 4. ed. Holt, Rinehart and Winston, 1985.

3

O ADULTO QUE NÃO QUER SER TRATADO COMO CRIANÇA: MALCOLM KNOWLES E A ANDRAGOGIA

“A andragogia é a arte e a ciência de ajudar adultos a aprender.” Malcolm Knowles, *The Modern Practice of Adult Education*, 1970¹

3.1. CONTEXTO HISTÓRICO

Há campos científicos que nascem de grandes rupturas filosóficas. A andragogia não nasceu assim. Nasceu de uma constatação ordinária, repetida em salas de treinamento corporativo por todo o pós-guerra norte-americano, que ninguém havia sistematizado porque parecia óbvia demais para merecer teoria: adultos aprendem de modo distinto das crianças, e quando são tratados como crianças, param de aprender.

¹ KNOWLES, Malcolm S. *The Modern Practice of Adult Education: Andragogy Versus Pedagogy*. Association Press, 1970. Revisado como: *The Modern Practice of Adult Education: From Pedagogy to Andragogy*. Cambridge Book Company, 1980.

O cenário intelectual que precede Knowles é o da pedagogia como disciplina total. Na primeira metade do século XX, não havia ciência da educação de adultos. Havia pedagogia, palavra que em grego designa literalmente “condução da criança” (do grego *paidós*, criança, e *agogos*, aquele que conduz), e a pressuposição tácita de que seus princípios se aplicavam universalmente a qualquer ser humano em situação de aprendizagem. O adulto que voltava para a escola, que participava de treinamentos corporativos ou que buscava requalificação profissional recebia, invariavelmente, o mesmo tratamento: currículo predefinido, progressão linear por disciplinas, avaliação por desempenho em testes, autoridade centrada no instrutor. O modelo era o da escola primária, escalado para corpos adultos.

Um antecedente que Knowles sempre reconheceu merece ser nomeado antes de qualquer coisa. O pedagogo e educador social dinamarquês-americano Eduard Lindeman, em 1926, publicou *The Meaning of Adult Education* (“O Significado da Educação de Adultos”) ², obra na qual afirmava que a experiência acumulada pelo adulto era o recurso mais rico da aprendizagem e que a educação de adultos deveria partir da situação, não da matéria. Lindeman não chegou a construir um sistema. Nomeou o impasse com precisão suficiente para que, três décadas depois, Knowles o retomasse.

O contexto histórico do pós-guerra americano é determinante. O GI Bill (Servicemen’s Readjustment Act de 1944, lei federal que financiava educação superior para veteranos de guerra) havia enviado milhões de ex-combatentes para as universidades norte-americanas, criando uma população de estudantes adultos em escala sem precedente histórico. Esses homens chegavam com experiências de combate, responsabilidades familiares, urgências financeiras e uma impaciência radical com qual-

² LINDEMAN, Eduard C. *The Meaning of Adult Education*. New Republic, 1926. Reeditado por Oklahoma Research Center for Continuing Professional and Higher Education, 1989.

quer aprendizagem que não tivesse aplicação imediata e visível. As universidades não sabiam o que fazer com eles. O modelo pedagógico tradicional os infantilizava; as taxas de abandono eram expressivas; os relatos de insatisfação, documentados em avaliações institucionais da época, apontavam para o mesmo diagnóstico: o adulto veterano se recusava a ser tratado como aluno do ensino médio.

Malcolm Shepherd Knowles nasceu em 1913 em Montana e construiu sua trajetória profissional no cruzamento entre trabalho social, administração educacional e pesquisa em comportamento organizacional. Passou pela Boston YMCA, pela National Education Association, e chegou ao Boston University para a carreira acadêmica que o tornaria referência. Foi professor na North Carolina State University de 1974 a 1979. Sua formação não foi em psicologia cognitiva, tampouco em filosofia da educação: foi em administração e serviço social, o que explica por que sua abordagem sempre privilegiou a dimensão prática, institucional e aplicada sobre a especulação teórica.

The Modern Practice of Adult Education (“A Prática Moderna da Educação de Adultos”), publicado em 1970 ¹, foi o primeiro livro em que Knowles usou o termo andragogia em inglês, tomando-o de empréstimo do pedagogo alemão Alexander Kapp, que havia cunhado a palavra em 1833, e do pedagogo iugoslavo Dusan Savicevic, que o empregava na Europa oriental nos anos 1960. A história do termo é ela mesma reveladora: a ideia de que adultos precisam de uma teoria de aprendizagem própria existia dispersa em vários contextos culturais sem nunca ter encontrado formulação sistemática em língua inglesa. Knowles fez exatamente isso. Não inventou o problema; construiu a ciência. E o problema que ele precisava resolver era mais caro e mais concreto do que qualquer disputa filosófica: era o fracasso sistemático do treinamento corporativo norte-americano a afetar dezenas de milhares de funcionários que aprendiam

pouco, aplicavam menos ainda, e eram depois acusados de “resistência à mudança”.

3.2. O PROBLEMA QUE KNOWLES ESTAVA RESOLVENDO

Nos anos 1950 e 1960, o treinamento corporativo nos Estados Unidos havia se tornado uma indústria considerável, e a falha que o definia tinha números; não impressão de instrutores insatisfeitos, mas dado repetível. Alimentada pela expansão das grandes corporações e pela necessidade de qualificar rapidamente uma força de trabalho em transformação acelerada, essa indústria reunia empresas como IBM, General Motors e AT&T, que investiam somas crescentes em programas internos de desenvolvimento: instrutores externos, centros de treinamento, materiais didáticos. Os resultados eram, consistentemente, decepcionantes. Pesquisas de efetividade da época, revisadas por Kirkpatrick ³ e outros pesquisadores do campo, apontavam taxas de transferência da aprendizagem para o ambiente de trabalho abaixo de 20%: menos de um em cinco participantes aplicava o que havia aprendido.

O diagnóstico prevalente na época era de que os funcionários “não queriam aprender” ou “resistiam à mudança”. Era um diagnóstico conveniente porque deslocava a responsabilidade do programa para o participante. Knowles propôs a inversão: e se o problema não fosse o aluno adulto, mas o método pedagógico aplicado a ele?

A observação central que sustenta toda sua teoria é simples e difícil de refutar: o adulto que chega a um programa de aprendizagem não é uma tábula rasa. É um sujeito com décadas de experiência acumulada, com uma identidade profissional construída, com demandas que o pressionam no presente e com um autoconceito que o define como agente

³ KIRKPATRICK, Donald L. *Evaluating Training Programs: The Four Levels*. Berrett-Koehler, 1994.

competente no mundo. Colocá-lo em uma sala onde um instrutor deposita informação de cima para baixo, onde o currículo foi definido por outra pessoa sem consultar suas necessidades, onde a progressão é determinada pela lógica da disciplina e não pela urgência da situação a resolver, é produzir precisamente a resistência que depois se atribui ao “caráter” do participante.

Havia também uma dimensão de humilhação que Knowles nomeou com clareza. O médico experiente que participa de um curso de atualização e é avaliado por prova de múltipla escolha. O engenheiro sênior que precisa cumprir horas de treinamento sobre conceitos que domina há quinze anos. O gerente que assiste a uma apresentação sobre “comunicação eficaz” construída com dinâmicas de grupo que ele associa ao ensino infantil. Nenhum deles “resiste à mudança”: todos resistem à infantilização. A diferença é epistemicamente fundamental.

O impasse teórico que Knowles identificou era mais profundo do que parecia. A pedagogia havia se desenvolvido como ciência da instrução de crianças porque crianças precisam ser conduzidas: dependem do adulto para definir o que aprender, não têm experiência acumulada como recurso, aprendem por acumulação de conteúdo que só fará sentido num futuro que não conseguem antecipar. Todas essas premissas são falsas quando aplicadas ao adulto. Como não havia teoria alternativa, o padrão vigente era aplicar à educação de adultos o mesmo modelo, com pequenas concessões cosméticas. Knowles foi o primeiro a articular sistematicamente que as premissas precisavam ser outras.

3.3. IDEIA CENTRAL

3.3.1. Os Seis Pressupostos da Andragogia

A teoria de Knowles se organiza em torno de seis pressupostos sobre o aprendiz adulto, desenvolvidos ao longo de suas obras principais [1, 4] e refinados na versão definitiva de *The Adult Learner*⁴. Cada um desses pressupostos descreve uma diferença estrutural entre a aprendizagem adulta e a infantil, e cada um tem implicações diretas para o design de qualquer produto educacional destinado a adultos.

O **autoconceito** (em inglês, *self-concept*, a imagem que o indivíduo tem de si mesmo como agente) é o pressuposto que ancora todos os outros. À medida que o indivíduo amadurece, esse autoconceito migra de dependência para autodirecionamento. A criança aceita que o adulto define o que ela deve aprender e em que ritmo, porque não tem ainda a identidade consolidada para questionar essa autoridade. O adulto, por sua vez, construiu ao longo da vida uma imagem de si mesmo como agente capaz de tomar decisões e dirigir sua própria trajetória. Quando colocado em situação de aprendizagem que o trata como dependente, ele experimenta uma dissonância entre esse autoconceito e o papel que lhe é atribuído. Essa dissonância produz resistência. Não resistência ao conteúdo: resistência ao papel. O produto educacional que ignora esse pressuposto e coloca o adulto na posição de receptor passivo está criando atrito antes de qualquer palavra do currículo ser dita.

A **experiência acumulada** (o repertório de vivências como recurso central de aprendizagem) é o pressuposto que reorganiza a hierarquia pedagógica. Esse repertório não é “contexto” ou “motivação”: é o substrato cognitivo sobre o qual qualquer conhecimento novo precisa ancorar para ser retido. Na pedagogia, a experiência do aprendiz tem valor marginal; a criança ainda não tem muito; o que importa é a experiência

⁴ KNOWLES, Malcolm S.; HOLTON, Elwood F.; SWANSON, Richard A. *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. 8. ed. Routledge, 2015. (1. ed. sob o título *The Adult Learner: A Neglected Species*. Gulf Publishing, 1973.)

do professor. Na andragogia, essa hierarquia se inverte: a experiência do aprendiz é o recurso central, e o papel do educador é ajudá-lo a explorá-la, questioná-la, expandir e corrigir o que for necessário.

A consequência prática é que métodos que ignoram a experiência do adulto (aulas expositivas longas, leitura passiva, exercícios genéricos desconectados de contextos reais) não são apenas ineficazes: são ativamente destrutivos, porque criam a sensação de que o que o adulto já sabe não importa. É aqui que a andragogia toca o tema da humilhação que Knowles nomeou com mais força: não é só que o adulto aprende menos; é que ele aprende que seu acervo de vida não tem valor para o processo ao qual foi submetido.

O terceiro pressuposto reorganiza o eixo temporal da aprendizagem. Knowles o chamou de **prontidão para aprender** (a disposição situacional e temporal do adulto para engajar-se com um conteúdo específico). Adultos se tornam dispostos a aprender as coisas que precisam saber para lidar com situações reais de suas vidas. Essa prontidão é contextual e temporal: não é uma disposição genérica para o aprendizado, mas uma abertura específica que surge quando o indivíduo se confronta com um problema que excede sua capacidade atual. Isso tem uma consequência crucial para o timing do produto educacional: o mesmo material pode ser irrelevante para um adulto hoje e absolutamente crítico para ele daqui a seis meses, dependendo da situação que enfrenta. Produtos educacionais que ensinam “para o futuro” ou que acumulam conteúdo “por precaução” violam esse pressuposto. O adulto não aprende o que pode precisar um dia; aprende o que precisa agora.

Da prontidão decorre naturalmente o quarto pressuposto, aquele que tem a consequência mais imediata para o design de qualquer produto educacional: a **orientação centrada no problema** (aprender em função de situações reais, não de disciplinas abstratas). Crianças têm ori-

entação centrada na matéria: aprendem matemática, história, ciências, línguas como disciplinas que serão úteis em algum momento não especificado. Adultos têm orientação centrada na situação: aprendem o que precisam para resolver uma questão específica que já existe. A consequência para o design instrucional é que a organização do conteúdo por disciplinas ou módulos temáticos é inadequada para adultos; o material deve ser organizado em torno de situações reais, de impasses concretos, de casos que o aprendiz reconhece como análogos ao que enfrenta. Um curso que ensina “gestão de conflitos” como disciplina abstrata terá desempenho muito inferior ao que ensina como conduzir a conversa difícil que o participante está evitando.

Se os pressupostos anteriores descrevem condições contextuais e cognitivas, o quinto atinge o nível dos estados internos: a **motivação interna** (em oposição a recompensas ou pressões externas). A motivação do adulto para aprender é predominantemente interna: satisfação profissional, autodesenvolvimento, sentido de realização, relevância para objetivos pessoais. A motivação externa (notas, diplomas, pressão hierárquica) pode mobilizar o corpo do adulto para dentro da sala de aula, mas não mobiliza o sistema cognitivo para a aprendizagem efetiva. Knowles não nega que incentivos externos existam ou que tenham algum efeito; afirma que, comparados com a motivação interna, são instrumentos fracos. O adulto que aprende porque quer, porque vê conexão direta entre o que está aprendendo e algo que importa para ele, aprende de modo qualitativamente distinto do adulto que aprende porque foi mandado.

O sexto pressuposto é o mais operacionalmente imediato de todos, e também o mais frequentemente ignorado por produtos educacionais que já chegam ao aprendiz adulto presumindo que ele entende por que está ali. Knowles o introduziu em edições posteriores porque as pesquisas de campo revelavam uma barreira de entrada que os cinco pressupos-

tos anteriores não cobriam: a **necessidade de saber** (o adulto precisa entender o porquê antes do como). Antes de se comprometer com qualquer processo de aprendizagem, o adulto precisa entender por que ele precisa aprender aquilo. Não uma justificativa retórica ou motivacional: uma explicação funcional de como esse conhecimento ou habilidade se conecta a uma necessidade real que ele reconhece como sua. A ausência dessa explicação não é um detalhe cosmético; é uma barreira de entrada que, se não for removida, compromete toda a sequência subsequente.

3.3.2. *A Estrutura Sistêmica dos Pressupostos*

Os seis pressupostos não são princípios independentes: formam um sistema cujo funcionamento exige que todos estejam presentes. A necessidade de saber cria a prontidão; a prontidão ativa a motivação interna; a motivação interna só sustenta o engajamento se o conteúdo se orienta ao problema real do aprendiz; a orientação ao problema faz sentido apenas quando a experiência do aprendiz é mobilizada como recurso; e tudo isso só funciona se o autoconceito do adulto como agente autodirecionado for respeitado ao longo de todo o processo. Remover qualquer um desses elementos não enfraquece o sistema: o colapsa.

Knowles formalizou esses pressupostos em quatro etapas de processo que o educador deve percorrer para criar as condições para a aprendizagem adulta: (1) criar um clima de aprendizagem que promova autodirecionamento; (2) diagnosticar necessidades de aprendizagem em colaboração com o aprendiz; (3) envolver o aprendiz no planejamento de sua própria experiência de aprendizagem; e (4) avaliar resultados em termos de aplicação real, não de desempenho em testes [1, 4].

Há uma tensão interna que Knowles reconheceu, especialmente em suas obras mais tardias: nem todo adulto chega a um processo de aprendizagem no estado de autodirecionamento que o modelo pressupõe. Adul-

tos que foram sistematicamente submetidos ao modelo pedagógico por anos de escolarização formal podem ter internalizado a dependência e resistir ao autodirecionamento, não porque não sejam capazes, mas porque nunca foram convidados a exercê-lo. Knowles chamou isso de “dependência condicionada” e recomendava que o educador de adultos fosse sensível ao nível de maturidade do aprendiz para evitar que a exigência de autodirecionamento prematuro produzisse ansiedade em vez de engajamento. O reconhecimento dessa limitação é também o ponto onde a andragogia convida ao questionamento: se a dependência pode ser condicionada, o que mais pode estar condicionado? Essa pergunta abriria, nas décadas seguintes, uma agenda de pesquisa que Knowles não antecipou completamente.

3.4. EVOLUÇÃO POSTERIOR

3.4.1. *Mezirow e a Aprendizagem Transformadora*

A resposta mais radical a essa pergunta veio de Jack Mezirow, sociólogo e educador da Columbia University. Em *Transformative Dimensions of Adult Learning* (“Dimensões Transformadoras da Aprendizagem de Adultos”), publicado em 1991⁵, Mezirow argumentou que a andragogia de Knowles, ao focar nos processos de aprendizagem, não havia interrogado suficientemente o conteúdo do que se aprende quando um adulto realmente muda: não apenas adquire novas informações ou habilidades, mas revisa as premissas fundamentais que organizam sua percepção do mundo.

A teoria da aprendizagem transformadora propõe que adultos operam a partir de “quadros de referência” (estruturas interpretativas construí-

⁵ MEZIROU, Jack. *Transformative Dimensions of Adult Learning*. Jossey-Bass, 1991.

das pela experiência, cultura e socialização ao longo da vida) que filtram a realidade, determinando o que o adulto percebe como relevante, o que descarta como irrelevante e como interpreta ambiguidades. A aprendizagem significativa, para Mezirow, é aquela que produz uma revisão crítica desses quadros, não apenas acumula informação dentro deles. Essa revisão raramente é confortável: envolve o que ele chamou de “dilemas desorientadores” (situações em que a realidade contradiz abertamente as premissas do aprendiz e o força a reconstruir seu sistema de interpretação).

Mezirow, ao mesmo tempo, produz uma extensão e uma crítica de Knowles. Extensão porque radicaliza a autonomia do aprendiz adulto: não apenas autodirecionamento no processo, mas agência crítica sobre os próprios pressupostos. Crítica porque sugere que Knowles havia descrito bem o “como” do aprendizado adulto sem questionar suficientemente o “por que mais profundo”: para que serve aprender, se não para transformar a forma como o adulto compreende e age sobre o mundo?

3.4.2. Freire e Knowles: Convergência Norte-Sul

A aproximação entre Paulo Freire e Malcolm Knowles é simultaneamente óbvia e paradoxal. Óbvia porque ambos rejeitam o que Freire chamou de “educação bancária” (o ensino como depósito de conteúdo em receptores passivos ⁶), ambos centralizam a experiência do aprendiz e ambos insistem na autonomia como condição da aprendizagem genuína. Paradoxal porque chegam a essas conclusões por trajetórias radicalmente diferentes: Freire a partir da educação popular com adultos analfabetos no Nordeste brasileiro, profundamente politizada e orientada para a consciência crítica; Knowles a partir da educação de adultos em contexto corporativo norte-americano, empiricamente orientada e deliberadamente

⁶ FREIRE, Paulo. *Pedagogia do Oprimido*. Paz e Terra, 1968. 58. ed., 2019.

não-ideológica.

A principal diferença entre os dois é que Freire trata a educação como ato político inerentemente situado em relações de poder, enquanto Knowles trata a educação de adultos como questão técnica de eficácia metodológica. Para Freire, não basta que o adulto aprenda de forma autônoma; é necessário que aprenda a questionar as condições que produzem sua realidade¹⁸. Para Knowles, o que importa é que o processo seja eficaz; as implicações políticas da educação ficam fora do escopo de sua teoria.

O peso dessa distinção é prático, não filosófico. A andragogia de Knowles pode ser, e frequentemente é, instrumentalizada em contextos corporativos para tornar mais eficiente a transmissão de valores e práticas que não são questionados pelo aprendiz, mas simplesmente aprendidos de modo mais palatável. Freire diria que isso não é educação libertadora: é pedagogia disfarçada de andragogia. O debate permanece aberto e legítimo, mas não invalida a utilidade técnica dos princípios de Knowles para o design instrucional.

3.4.3. A Andragogia no Século XXI

A chegada das tecnologias digitais e dos ambientes virtuais de aprendizagem recolocou os princípios andragógicos em pauta, com resultados mistos. De um lado, plataformas de aprendizagem online oferecem, ao menos em teoria, as condições que Knowles prescreveu: autonomia sobre o ritmo, possibilidade de escolha de conteúdo, aprendizagem contextualizada em problemas reais. De outro, a prática concreta dos MOOCs (Massive Open Online Courses, cursos abertos via internet a qualquer número de participantes, oferecidos por universidades como MIT, Stanford e Coursera) revelou um paradoxo: apesar de disponibilizarem conteúdo de alta qualidade de instituições de prestígio, suas taxas de con-

clusão ficam consistentemente entre 5% e 15%⁷, muito abaixo do que se observaria se os princípios andragógicos fossem aplicados.

O diagnóstico dessa falha é andragógico: MOOCs aplicam, na prática, um modelo pedagógico disfarçado de autonomia. O adulto tem liberdade de acessar o material quando quiser, mas o currículo continua organizado por disciplinas, avançando por progressão predefinida, desconectado das situações reais do aprendiz, sem diagnóstico de necessidades e sem comunidade de prática que ative a experiência acumulada. Flexibilidade de horário é uma concessão logística, não uma concessão andragógica.

A aprendizagem baseada em competências (do inglês *competency-based education*, modelo que organiza o currículo em torno de habilidades demonstráveis e aplicáveis, em vez de cargas horárias de disciplinas) é outra expressão institucional do pressuposto da orientação centrada no problema que Knowles prescrevia. A aprendizagem adaptativa (sistemas algorítmicos que ajustam o percurso ao desempenho do aprendiz) traduz tecnologicamente o pressuposto da prontidão: em vez de uma progressão linear igual para todos, o sistema detecta o que o aprendiz já sabe e direciona a atenção para o que ainda não domina.

3.4.4. Os Críticos

A andragogia não é sem críticos. Hartree⁸ argumentou que Knowles não havia demonstrado empiricamente que adultos e crianças diferem fundamentalmente em seus processos de aprendizagem; a distinção poderia ser de grau, não de natureza. Davenport e Davenport⁹ questionaram se a teoria seria descritiva (descreve como adultos aprendem) ou

⁷ JORDAN, Katy. Massive Open Online Course Completion Rates Revisited: Assessment, Length and Attrition. *The International Review of Research in Open and Distributed Learning*, v. 16, n. 3, 2015.

⁸ HARTREE, Anne. Malcolm Knowles's Theory of Andragogy: A Critique. *International Journal of Lifelong Education*, v. 3, n. 3, p. 203–210, 1984.

⁹ DAVENPORT, Joseph; DAVENPORT, Judith A. A Chronology and Analysis of the Andragogy Debate. *Adult Education Quarterly*, v. 35, n. 3, p. 152–159, 1985.

prescritiva (descreve como deveriam aprender), uma distinção que Knowles nunca resolveu completamente. Tennant¹⁰ apontou que o conceito de autodirecionamento pressupõe um sujeito liberal autônomo que pode não corresponder à experiência de adultos em contextos culturais coletivistas, onde a aprendizagem é fundamentalmente social e relacional.

Essas críticas são bem-fundadas. Knowles, em obras posteriores, reconheceu parcialmente cada uma dessas objeções, moderando sua posição original e afirmando que a andragogia não era um conjunto de verdades absolutas, mas um sistema de pressupostos que deveriam ser testados contextualmente. Essa posição mais modesta é também mais epistemicamente defensável: os princípios andragógicos funcionam melhor como perguntas diagnósticas (“Esse design respeita o autoconceito do aprendiz?”) do que como afirmações universais sobre como todo adulto aprende. O que as críticas não puseram em questão, porque a evidência empírica o tornava difícil, foi o custo prático de ignorar esses pressupostos, custo que o impacto real da andragogia no mundo continuou a documentar.

3.5. IMPACTO REAL NO MUNDO

3.5.1. *A Transformação do Treinamento Corporativo*

O impacto mais direto e economicamente mensurável da andragogia foi na transformação do treinamento corporativo. A partir dos anos 1980, sob influência crescente das ideias de Knowles, grandes corporações norte-americanas e europeias reformularam seus programas de desenvolvimento de pessoal, migrando de aulas expositivas para metodologias experienciais. Estudos de caso, simulações e projetos práticos

¹⁰ TENNANT, Mark. *Psychology and Adult Learning*. 3. ed. Routledge, 2006.

passaram a substituir a instrução direta, e o método de aprendizagem em ação (do inglês *action learning*, metodologia desenvolvida pelo pesquisador britânico Reginald Revans nos anos 1940¹¹, fundamentada em pressupostos semelhantes aos de Knowles: adultos aprendem mais efetivamente quando trabalham sobre problemas reais em grupos, com reflexão estruturada sobre a experiência) ganhou adoção ampla em programas corporativos de liderança.

A ATD (Association for Talent Development, maior organização norte-americana de profissionais de treinamento e desenvolvimento, anteriormente denominada ASTD), incorporou os princípios andragógicos como referencial central de suas certificações e metodologias, disseminando-os por empresas de múltiplos setores. Dados compilados pela ATD em seu relatório de 2022 indicam que empresas nos Estados Unidos investem mais de 90 bilhões de dólares anuais em treinamento e desenvolvimento¹². A questão da efetividade continua sendo o nó central do campo: estudos conduzidos por Bersin by Deloitte¹³ apontam que programas que incorporam princípios andragógicos (aprendizagem no fluxo do trabalho, conteúdo contextualizado, autodirecionamento) apresentam taxas de transferência para o ambiente de trabalho de 40% a 60%, contra menos de 20% de programas estruturados por modelo pedagógico.

3.5.2. Os Programas de MBA e a Educação Executiva

O redesenho dos programas de MBA e de educação executiva nas décadas de 1980 e 1990 foi, em grande medida, uma aplicação institucional dos princípios andragógicos, ainda que nem sempre nominalmente referenciados a Knowles. A introdução do método de casos como metodologia central em Harvard Business School reflete exatamente o pressuposto

¹¹ REVANS, Reginald. *The Origins and Growth of Action Learning*. Chartwell-Bratt, 1982.

¹² ASSOCIATION FOR TALENT DEVELOPMENT (ATD). *State of the Industry Report 2022*. ATD Press, 2022.

¹³ BERSIN, Josh. *The Corporate Learning Factbook: Benchmarks, Trends, and Analysis of the U.S. Training Market*. Bersin by Deloitte, 2014.

da orientação centrada no problema: o adulto aprende melhor quando o conteúdo é apresentado como situação a resolver do que como teoria a assimilar.

Programas de formação executiva em instituições como Wharton, INSEAD e IMD construíram propostas pedagógicas que explicitamente incorporam os pressupostos de Knowles: diagnóstico de necessidades individual antes do início do programa, composição de turmas por pares com experiências comparáveis, metodologias baseadas em projetos reais dos participantes, avaliação por impacto no retorno ao trabalho e não por desempenho em exames. Esses programas cobram preços que podem superar 50.000 dólares por participante precisamente porque oferecem o que o modelo andragógico promete: não transmissão de informação, mas transformação da capacidade de agir.

3.5.3. O Fracasso dos MOOCs como Evidência Andragógica

As taxas de conclusão dos MOOCs constituem um dos experimentos naturais mais instrutivos sobre as consequências de ignorar Knowles em escala. Quando plataformas como Coursera, edX e Khan Academy surgiram entre 2011 e 2013, havia expectativas de que a combinação de acesso gratuito, conteúdo de qualidade e flexibilidade de horário democratizaria a educação de alto nível globalmente. Os números decepcionaram: pesquisa de Jordan ⁶, sintetizando dados de 29 MOOCs com 1,7 milhão de matrículas, encontrou taxa mediana de conclusão de 12,6%. Cursos da MIT OpenCourseWare chegam a registrar taxas abaixo de 5%.

A análise andragógica é direta: MOOCs resolveram o problema de acesso sem resolver o problema andragógico. Nenhum dos seis pressupostos de Knowles é endereçado estruturalmente por um MOOC típico: não há diagnóstico de prontidão, não há mobilização da experiência do aprendiz, não há orientação ao problema real do participante, não há

comunidade que ative o autoconceito de agente, não há explicação personalizada da necessidade de saber, e a motivação é inteiramente dependente de fatores externos e voláteis. O resultado, para qualquer leitor de Knowles, era previsível.

3.5.4. Na Formação de Profissionais de Saúde

Na área da saúde, a influência de Knowles se materializa no redesenho da formação médica e de enfermagem a partir dos anos 1980. O movimento pela educação baseada em competências (do inglês *competency-based medical education*, modelo que organiza o currículo de residências médicas em torno de habilidades demonstráveis em contextos reais, em vez de cargas horárias por disciplinas) incorpora os pressupostos andragógicos de orientação ao problema e avaliação por desempenho real. Programas de residência médica reformulados com base nesses princípios em instituições como Mayo Clinic e Johns Hopkins reportaram melhorias mensuráveis tanto na satisfação dos residentes quanto nos índices de competência clínica avaliados por supervisores ¹⁴.

Em todos esses contextos (corporativo, executivo, universitário, clínico) o que a andragogia explica é a mesma coisa: por que métodos que respeitam o adulto como sujeito ativo produzem resultados que métodos que o tratam como receptor não produzem. Mas a teoria de Knowles nunca opera sozinha; ela define quem aprende, sem dizer o que deve ser ensinado nem como a estrutura do ensino deve ser construída. Essas perguntas pertencem a outros autores.

3.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

¹⁴ TEN CATE, Olle. Competency-Based Postgraduate Medical Education: Past, Present and Future. *GMS Journal for Medical Education*, v. 34, n. 5, 2017.

3.6.1. Knowles + Bloom: o *Quem e o Quê da Aprendizagem*

Esses outros autores respondem a perguntas que a andragogia deixa em aberto por construção, e cada resposta pressupõe Knowles sem o substituir. A taxonomia de objetivos educacionais de Benjamin Bloom¹⁵ é o caso mais imediato: responde a uma pergunta diferente da de Knowles, mas as respostas se pressupõem mutuamente. Knowles define o *quem*: um adulto autodirecionado, motivado internamente, orientado ao problema, com experiência que precisa ser mobilizada. Bloom define o *quê*: os diferentes níveis cognitivos que um objetivo de aprendizagem pode endereçar, do simples reconhecimento de informação à criação original de soluções.

A conexão prática é que a andragogia sem Bloom tende a produzir objetivos de aprendizagem vagos (“o aluno vai compreender gestão de projetos”) que não especificam o nível de competência esperado nem a forma de avaliá-lo. Bloom sem Knowles tende a produzir currículos bem estruturados cognitivamente, mas que ignoram a experiência do aprendiz, sua prontidão e sua motivação, resultando em conteúdo preciso entregue da forma errada para a pessoa errada no momento errado. A integração das duas teorias produz objetivos que são ao mesmo tempo andragogicamente relevantes (conectados a problemas reais do aprendiz) e cognitivamente precisos (especificam o nível de domínio esperado e como será verificado).

¹⁵ BLOOM, Benjamin S. (Org.). *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. David McKay, 1956. Revisado em: ANDERSON, Lorin W.; KRATHWOHL, David R. (Orgs.). *A Taxonomy for Learning, Teaching, and Assessing*. Longman, 2001.

3.6.2. Knowles + Wiggins e McTighe: Planejamento Reverso como Andragogia Operacionalizada

O planejamento reverso (*Understanding by Design*, ou Compreensão por Design) de Grant Wiggins e Jay McTighe¹⁶ pode ser lido como a operacionalização do pressuposto andragógico mais fundamental: começar pelo resultado desejado, não pelo conteúdo disponível. O processo de design reverso prescreve que o educador defina primeiro qual transformação o aprendiz deve ter completado ao final; só depois determine as evidências que confirmariam essa transformação; só então projete as experiências de aprendizagem que a produziriam.

O rigor andragógico dessa sequência está em partir da necessidade real do aprendiz (o que ele precisa ser capaz de fazer), não da lógica interna da disciplina. A segunda etapa (definir evidências de aprendizagem antes de definir atividades) impede a armadilha que Knowles identificou em currículos pedagógicos: atividades realizadas por sua própria lógica interna, sem conexão com transformações verificáveis no mundo real do aprendiz. A terceira etapa (projetar as experiências) é onde a experiência prévia do aprendiz deve ser mobilizada como recurso, não como contexto decorativo.

Uma limitação precisa ser nomeada: Wiggins e McTighe desenvolveram o planejamento reverso para contextos escolares, onde o conteúdo a ser ensinado é predefinido por currículos institucionais. Na andragogia, o pressuposto é que o diagnóstico de necessidades do aprendiz adulto participa da definição do que deve ser ensinado. O planejamento reverso adapta-se a esse contexto, mas requer uma etapa anterior (o diagnóstico andragógico de Knowles) para que os objetivos de compreensão reflitam necessidades reais e não apenas a lógica da disciplina.

¹⁶ WIGGINS, Grant; MCTIGUE, Jay. *Understanding by Design*. 2. ed. ASCD, 2005.

3.6.3. Knowles + Murphy: o Adulto que Abandona quando Não Percebe Valor

Lincoln Murphy, ao desenvolver o conceito de Customer Success como garantia estrutural de resultado ¹⁷, identificou um padrão que é andragogicamente previsível: o cliente, e por extensão o aluno adulto, que não percebe valor tangível nos primeiros trinta dias de engajamento tem probabilidade muito alta de abandonar o produto, independentemente de sua qualidade intrínseca. Murphy chama esse fenômeno de *churn* (cancelamento ou abandono do produto pelo cliente), e o situa como o principal risco de qualquer produto baseado em assinatura ou comprometimento contínuo.

O mecanismo andragógico subjacente é o da prontidão: o adulto que se matricula num produto educacional faz isso porque percebeu uma necessidade. Mas a prontidão não é permanente; ela surge em resposta a um problema específico e se dissipa se o produto não entrega progresso visível nesse problema em tempo razoável. O adulto que, nos primeiros trinta dias, não consegue ver como o que está aprendendo se conecta ao problema que o motivou a se matricular não cancela o autoconceito: cancela o produto. Esse abandono não é falha de marketing, falha de retenção ou falha de qualidade de conteúdo; é falha de design andragógico. O produto não demonstrou relevância enquanto a janela de prontidão ainda estava aberta.

A integração Knowles com Murphy produz uma regra de design: todo produto educacional para adultos deve entregar, nos primeiros momentos de engajamento, uma demonstração inequívoca de relevância para o problema real do aprendiz. Não uma promessa de relevância futura; uma experiência imediata de utilidade. É o que Knowles chamava de “necessidade de saber” aplicada ao design de produto. A maioria dos

¹⁷ MURPHY, Lincoln. *Customer Success: How Innovative Companies Are Reducing Churn and Growing Recurring Revenue*. Wiley, 2016.

produtos educacionais para adultos, porém, é projetada sem esse diagnóstico. Usam personas de marketing onde deveriam usar diagnóstico andragógico, e pagam esse preço em churn, em NPS baixo e em alunos que concluem módulos sem transformar coisa nenhuma.

PARTE II

Evolucao da Ciencia da Aprendizagem

4

COMEÇAR PELO FIM: WIGGINS, McTIGHE E O DESIGN REVERSO

“To begin with the end in mind means to start with a clear understanding of your destination.” _“Começar com o fim em mente significa partir com uma compreensão clara do seu destino.”. Grant Wiggins e Jay McTighe, *Understanding by Design*, 1998¹

4.1. CONTEXTO HISTÓRICO

Há um tipo específico de crise educacional que não se anuncia com colapso visível. Anuncia-se com uma anomalia discreta e persistente: o aluno que passa em todas as provas e não aprendeu nada. Não é ignorância. É algo mais estruturalmente perturbador: a dissociação entre o funcionamento formal da escola e o desenvolvimento de competência

¹ WIGGINS, Grant; McTIGHE, Jay. *Understanding by Design*. Association for Supervision and Curriculum Development (ASCD), 1998. 2. ed., 2005.

real, uma lacuna que pode existir por décadas sem que o sistema a reconheça como problema seu.

Nos Estados Unidos dos anos 1980 e início dos 1990, essa anomalia havia deixado de ser caso isolado.

A publicação do relatório *A Nation at Risk* (Uma Nação em Risco), em 1983, pela National Commission on Excellence in Education ², tornou-se o documento diagnóstico de uma geração. A linguagem era explicitamente alarmista: a educação americana estava em “maré de mediocridade” que, caso imposta por uma potência estrangeira, seria considerada ato de guerra. O dado central era simples e brutal. Estudantes americanos performavam consistentemente abaixo de seus pares em países industrializados; a alfabetização funcional de adultos era insuficiente para as demandas de um mercado em transformação; empresas reportavam que precisavam treinar do zero habilidades básicas em jovens recém-saídos do ensino médio. O relatório desencadeou uma reação política que definiria a agenda educacional norte-americana por ao menos duas décadas.

O que se seguiu foi o movimento de *standards* educacionais, uma tentativa de responder à crise definindo o que os alunos deveriam ser capazes de fazer ao final de cada ciclo escolar. A iniciativa *Goals 2000: Educate America Act*, aprovada pelo Congresso em 1994 ³, formalizou metas nacionais de aprendizagem e criou mecanismos de prestação de contas para estados e distritos. O pressuposto implícito: se havia crise, era porque os objetivos não estavam claros o suficiente. Bastaria defini-los com precisão e o sistema se alinharia naturalmente.

O pressuposto estava errado. Ou, ao menos, era incompleto de uma forma que levaria tempo para ser percebida.

² NATIONAL COMMISSION ON EXCELLENCE IN EDUCATION. *A Nation at Risk: The Imperative for Educational Reform*. U.S. Department of Education, 1983.

³ UNITED STATES CONGRESS. *Goals 2000: Educate America Act* (Pub. L. 103-227). Washington, D.C.: U.S. Government Printing Office, 1994.

Definir o que os alunos devem saber não resolve o problema de como se organiza o ensino para que eles efetivamente saibam. A forma dominante de organizar o ensino, aquela que sobrevivia às reformas porque estava cristalizada nos livros didáticos, nas grades curriculares e na formação dos professores, era o que Wiggins e McTighe chamariam de *forward design* (design prospectivo): o currículo organizado a partir do conteúdo disponível, não a partir dos resultados esperados. Começa-se pelo material; espera-se que o aluno, ao absorvê-lo, desenvolva competência. Não havia garantia nenhuma de que isso ocorresse. Na escala do sistema americano, havia evidência crescente de que não ocorria.

Grant Wiggins e Jay McTighe chegaram a esse problema por trajetórias convergentes, ambas enraizadas na prática de avaliação educacional. Wiggins havia trabalhado como professor antes de se tornar consultor de avaliação e reformador curricular; fundou o *Center on Learning, Assessment, and School Structure* (CLASS) e dedicou anos ao estudo do que torna uma avaliação autêntica, capaz de verificar competência real em vez de reprodução mecânica de conteúdo⁴. McTighe havia desenvolvido trabalho semelhante no sistema educacional de Maryland, coordenando iniciativas de reforma curricular e avaliação por competências. Os dois tinham, por isso, uma perspectiva que a maioria dos reformadores da era dos *standards* não tinha: começaram pela pergunta sobre como saber se um aluno aprendeu, não pela pergunta sobre o que ensinar.

Essa diferença de ponto de partida é o que distingue *Understanding by Design* (Compreensão por Design) das demais reformas dos anos 1990. Não é um manifesto sobre o que ensinar. É uma metodologia sobre como projetar o ensino a partir do que se quer que o aluno seja capaz de fazer. O primeiro passo dessa metodologia, como se verá, é parar de começar pelo conteúdo.

⁴ WIGGINS, Grant. *Assessing Student Performance: Exploring the Purpose and Limits of Testing*. Jossey-Bass, 1993.

4.2. O PROBLEMA QUE WIGGINS E McTIGHE ESTAVAM RESOLVENDO

Understanding by Design parte de um diagnóstico que, enunciado com clareza, soa quase trivial: o design educacional dominante opera na direção errada. O que Wiggins e McTighe fizeram foi levar esse diagnóstico a sério o suficiente para construir uma metodologia inteira a partir dele.

A mais ubíqua das falhas que identificaram é o que chamaram de *coverage* (cobertura). O professor organiza o curso em torno do conteúdo que precisa ser “coberto”: cada semana, um conjunto de tópicos; cada tópico, um conjunto de páginas ou aulas correspondentes; ao final do semestre, o material foi todo apresentado. A medida de sucesso é a completude do programa, não a competência dos alunos. O modelo opera uma confusão fundamental entre *apresentar* e *ensinar*. O conteúdo foi apresentado, mas o aluno, confrontado com uma situação real que exige mobilizar aquele conhecimento, se revela incapaz de fazê-lo. O professor ensinou; o aluno não aprendeu. Conclusão habitual: o problema é o aluno. Conclusão correta: o design estava errado desde o início.

A segunda falha é o que chamaram de *activity trap* (armadilha da atividade). O professor projeta aulas engajantes, dinâmicas, bem executadas. Os alunos participam, discutem, constroem algo, saem satisfeitos. As atividades foram projetadas a partir de sua lógica interna, são interessantes, são interativas, são animadas pelo entusiasmo do professor, sem conexão explícita com transformações verificáveis na competência dos alunos. A satisfação da atividade oculta a ausência de aprendizagem substantiva. As avaliações chegam como choque: as atividades foram excelentes, mas os alunos não desenvolveram o que deveriam desenvolver.

Há uma terceira, menos óbvia e mais reveladora sobre a psicologia docente. Wiggins e McTighe a descrevem como o paradoxo do professor apaixonado pelo assunto e estruturalmente indiferente ao resultado. O professor de história medieval apaixona-se pela riqueza do período, pela densidade das fontes primárias, pela complexidade das relações feudais. Esse apaixonamento produz aulas de genuína qualidade intelectual. Mas o professor não se perguntou, antes de projetar o curso, o que quer que os alunos sejam capazes de fazer com esse conhecimento ao final. O resultado é um curso que serve ao entusiasmo do professor, não ao desenvolvimento do aluno. A paixão pelo conteúdo substituiu a clareza sobre o propósito.

O diagnóstico é preciso e incômodo porque ataca algo que a cultura escolar tende a proteger: a ideia de que um professor comprometido com seu conteúdo é, por definição, um bom professor. Wiggins e McTighe separam essas duas coisas sem acusação, mas também sem ambiguidade. Comprometimento com o conteúdo é necessário. Insuficiente.

O que falta é comprometimento com o resultado.

As três falhas têm em comum um pressuposto que os autores tornaram explícito: o design educacional dominante parte do conteúdo disponível, seleciona atividades coerentes com esse conteúdo, avalia ao final o quanto o aluno absorveu. Essa direção não garante coerência entre o que se ensina e o que se quer que o aluno seja capaz de fazer. A única forma de garantir essa coerência é inverter a ordem: começar pelo resultado desejado e trabalhar retroativamente até chegar ao conteúdo e às atividades. É essa inversão que o próximo item descreve em sua estrutura técnica precisa.

4.3. IDEIA CENTRAL

4.3.1. Os Três Estágios do *Backward Design*

O coração de *Understanding by Design* é um modelo de três estágios para o design curricular, que os autores chamam de *backward design* para enfatizar que a ordem de planejamento é o inverso da ordem de apresentação ¹. O professor não projeta o curso na sequência em que vai executá-lo; projeta na sequência inversa, começando pelo que deveria estar no final.

O primeiro estágio é a definição dos *resultados desejados*: o que, ao final da unidade ou do curso, o aluno deve ser capaz de entender, fazer ou transferir para situações novas? Wiggins e McTighe distinguem aqui entre três níveis de resultado, organizados em torno do conceito de *big ideas* (ideias essenciais de alta transferência). Nem todo conteúdo é equivalente. Alguns conceitos e habilidades permitem ao aluno não apenas reproduzir o que aprendeu, mas aplicá-lo em contextos que o professor não antecipou. São esses que precisam estar no centro do design. Identificar quais são esses conceitos e torná-los o centro gravitacional do planejamento é o trabalho desse estágio.

A distinção é operacionalmente importante porque currículos tendem a tratar todo conteúdo como igualmente relevante. O modelo propõe uma hierarquia explícita: há conteúdos que o aluno deve apenas conhecer (informações periféricas que contextualizam sem ser centrais); há conteúdos que o aluno deve compreender (conceitos que informam o raciocínio dentro da disciplina); e há conteúdos com os quais o aluno deve desenvolver compreensão profunda e transferível, aqueles que produzem a capacidade de ver o mundo de forma diferente ¹. Projetar o curso sem essa hierarquia é projetar um mapa onde todas as estradas têm a mesma largura.

O segundo estágio é a definição das *evidências de aprendizagem*: antes de projetar qualquer atividade de ensino, o designer educacional precisa especificar como saberá, ao final, se o aluno atingiu os resultados do estágio anterior. Que desempenhos, produções ou demonstrações seriam evidência de compreensão genuína? Que erros ou equívocos típicos revelariam que a aprendizagem superficial mascarou a ausência de compreensão real?

Esse estágio é o mais perturbador para a prática docente convencional, porque inverte a relação entre ensino e avaliação. No modelo dominante, a avaliação vem ao final, como verificação do que foi ensinado. Em *Understanding by Design*, a avaliação é projetada antes do ensino, porque ela é a especificação do que o ensino deve produzir. A pergunta “como vou saber que aprendeu?” precede e condiciona a pergunta “o que vou ensinar?”. Com isso, as avaliações deixam de ser medidas de cobertura de conteúdo e se tornam medidas de desenvolvimento de competência. Wiggins havia trabalhado extensivamente nesse ponto antes do livro, desenvolvendo o conceito de *avaliação autêntica* (tarefas de desempenho que replicam a complexidade das situações reais nas quais a competência deve ser mobilizada) como alternativa à avaliação por testes de reprodução ⁴.

O terceiro estágio, e apenas o terceiro, é quando se projetam as *experiências de aprendizagem*: as atividades, as sequências de instrução, os materiais, as discussões, os exercícios, tudo que constitui o “conteúdo” visível do curso. No modelo convencional, esse estágio é o primeiro e frequentemente o único. Em *Understanding by Design*, é o último, subordinado aos dois anteriores. A pergunta muda de “que atividades posso fazer sobre esse tema?” para “que atividades produzirão os desempenhos que especifiquei como evidência de aprendizagem para os resultados que defini?”

A inversão é estrutural, não cosmética. Não é uma questão de “começar pensando nos objetivos antes de fazer as aulas”, como tantas reformas educacionais superficialmente pregaram. É uma questão de mudar a direção causal do design: o OUTPUT define o PROCESSO, que define o INPUT. Não o inverso.

4.3.2. A Teoria da Compreensão que Fundamenta o Modelo

Subjacente aos três estágios há uma teoria da compreensão que Wiggins e McTighe desenvolvem com cuidado, porque sem ela o modelo se torna apenas uma reorganização procedimental sem fundamento conceitual. O que distingue o segundo estágio de uma avaliação qualquer é a noção de compreensão genuína, e essa noção precisa ser precisa o suficiente para ser operacionalizável.

A compreensão, para eles, não é um estado binário nem uma progressão linear de simples a complexo. É uma capacidade de realizar seis operações específicas com o conhecimento, que denominam *seis facetas da compreensão*¹: (1) *explicar*, articular princípios, teorias e processos que dão conta de um fenômeno; (2) *interpretar*, conferir significado a dados, textos, experiências; (3) *aplicar*, usar o conhecimento em situações novas e diversas; (4) *ter perspectiva*, ver pontos de vista e fundamentos críticos; (5) *demonstrar empatia*, perceber o que outros veem e sentem; e (6) *ter autoconhecimento*, reconhecer os próprios padrões de raciocínio, preconceitos e limites.

O modelo das seis facetas serve a um propósito específico: especificar que tipo de evidência é indicativa de compreensão genuína. Um aluno que consegue reproduzir uma explicação demonstra algo, mas não a capacidade de interpretar ou aplicar. Um aluno que aplica um princípio a casos novos demonstra algo mais. A compreensão que os autores associam ao desenvolvimento de expertise envolve mobilizar as seis facetas

de forma integrada: não apenas saber, mas saber por quê, saber quando, saber para quem e saber o que não se sabe.

É essa distinção entre reprodução e compreensão genuína que torna o modelo metodologicamente rigoroso. Sem uma teoria do que conta como evidência de compreensão, o segundo estágio se tornaria arbitrário, qualquer avaliação serviria. Com ela, o design tem um critério preciso: as evidências devem ser avaliações autênticas, capazes de revelar compreensão real, não apenas cobertura de conteúdo.

4.4. EVOLUÇÃO POSTERIOR

4.4.1. Do Currículo ao Corporate Learning

Understanding by Design foi publicado originalmente em 1998 pela ASCD (Association for Supervision and Curriculum Development)¹ e revisado em segunda edição em 2005⁵. A adoção inicial foi lenta e geograficamente concentrada, como é típico em reformas curriculares. Os primeiros sistemas a implementar o UbD em escala eram distritos com lideranças reformistas, dispostos a investir em formação docente e a enfrentar a resistência que qualquer inversão de prática estabelecida inevitavelmente produz.

O que aconteceu nas décadas seguintes foi uma disseminação que os autores não tinham previsto com clareza: a transposição do modelo para o campo do treinamento e desenvolvimento corporativo, onde encontrou terreno fértil por razões que são, em si mesmas, instrutivas.

O treinamento corporativo havia chegado, no início dos anos 2000, a uma crise análoga à que a educação escolar vivia nos anos 1980. Estudos do campo de Aprendizagem e Desenvolvimento (L&D) revelavam consis-

⁵ WIGGINS, Grant; McTIGHE, Jay. *Understanding by Design*. 2. ed. revisada e expandida. ASCD, 2005.

tentemente que grandes percentuais dos investimentos em treinamento não produziam mudança mensurável de comportamento no trabalho ⁶. O diagnóstico era estruturalmente o mesmo: programas projetados a partir do conteúdo que o facilitador sabia ensinar, não a partir do comportamento que a organização precisava que os funcionários desenvolvessem. A armadilha da cobertura era onipresente; a avaliação de reação (o participante gostou do treinamento?) era confundida com avaliação de impacto (o participante mudou seu comportamento de trabalho?).

O modelo de Kirkpatrick ⁷, que havia estabelecido quatro níveis de avaliação de treinamento (reação, aprendizado, comportamento e resultados), já pressupunha a necessidade de definir o resultado esperado antes de projetar o treinamento, mas nunca havia sido operacionalizado como metodologia de design. A articulação entre os níveis de Kirkpatrick e os três estágios do UbD produziu, em práticas diferentes e com terminologias distintas, o que passou a ser chamado de *outcome-based training design*, cuja distinção fundamental era a mesma: o que o participante aprende durante o treinamento não é o mesmo que o que é capaz de fazer no trabalho depois.

A convergência com o movimento de OKRs (*Objectives and Key Results*) foi, nesse contexto, natural. A lógica dos OKRs, desenvolvida por Andy Grove na Intel nos anos 1970 e popularizada por John Doerr ⁸, é estruturalmente análoga ao modelo: defina o resultado a ser alcançado (o Objective) com antecedência; especifique as evidências mensuráveis de que o resultado foi alcançado (os Key Results); apenas então organize as iniciativas. Quando essa lógica chegou ao campo de L&D, a sobreposição produziu o que alguns praticantes chamaram de *learning OKRs*: a aplicação da lógica de resultados-chave ao design de programas de

⁶ BRINKERHOFF, Robert O. *The Success Case Method: Find Out Quickly What's Working and What's Not*. Berrett-Koehler, 2003.

⁷ KIRKPATRICK, Donald L.; KIRKPATRICK, James D. *Evaluating Training Programs: The Four Levels*. 3. ed. Berrett-Koehler, 2006.

⁸ DOERR, John. *Measure What Matters: OKRs; The Simple Idea that Drives 10x Growth*. Portfolio/Penguin, 2018.

aprendizagem, garantindo que cada programa tivesse um resultado de comportamento definido e uma métrica de evidência antes de qualquer conteúdo ser desenvolvido.

4.4.2. Bloom e o Design Thinking como Desenvolvimentos Paralelos

O modelo de três estágios pressupõe a capacidade de especificar resultados de aprendizagem com precisão. Mas o que significa especificar um resultado com precisão? A Taxonomia de Bloom ⁹, e mais especificamente sua revisão por Anderson e Krathwohl publicada em 2001 ¹⁰, funcionou como complemento operacional indispensável para responder a essa pergunta.

A taxonomia original organizava objetivos educacionais em seis categorias hierárquicas: conhecimento, compreensão, aplicação, análise, síntese e avaliação. A versão revisada substituiu os substantivos por verbos, tornando a taxonomia utilizável como ferramenta de especificação de objetivos: lembrar, compreender, aplicar, analisar, avaliar, criar. A diferença não é apenas terminológica. Os verbos tornam a taxonomia operacionalizável de uma forma que os substantivos não permitiam: um objetivo especificado como “o aluno vai *analisar* a estrutura argumentativa de textos históricos primários” é mensurável de uma forma que “o aluno vai compreender textos históricos” não é. Usado para especificar os resultados do primeiro estágio, Bloom transforma o backward design de um princípio filosófico numa ferramenta de design.

Uma linha de desenvolvimento distinta, mas conceitualmente legítima, conecta o modelo ao *design thinking* aplicado à educação, que ganhou tração nos anos 2010. O design thinking, codificado pela d.school de Stanford e pela firma IDEO, parte de um princípio análogo: defina o

⁹ BLOOM, Benjamin S. (Org.). *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. David McKay, 1956.

¹⁰ ANDERSON, Lorin W.; KRATHWOHL, David R. (Orgs.). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman, 2001.

problema do usuário antes de projetar a solução. Ambas as abordagens emergem de uma mesma crítica ao design centrado no produtor, o que o professor quer ensinar, o que o designer quer construir, em favor do design centrado no resultado do usuário. O que o design thinking adicionou foi a ênfase na iteração rápida e na prototipagem, que o UbD, como metodologia voltada a currículos de longa duração, não havia desenvolvido com o mesmo rigor.

4.5. IMPACTO REAL NO MUNDO

4.5.1. Escala e Alcance Institucional

Understanding by Design tornou-se, nas duas décadas seguintes à sua publicação, um dos textos de referência mais influentes na formação de professores e no desenvolvimento curricular nos Estados Unidos. A ASCD reporta que o livro vendeu mais de um milhão de cópias ¹¹, número incomum para obras de teoria educacional. O impacto, porém, não se mede por vendas. Mede-se pela penetração institucional.

Numerosos estados americanos incorporaram os princípios do UbD em seus frameworks de desenvolvimento curricular e em seus programas de certificação docente. O International Baccalaureate (IB), programa de educação internacional presente em mais de 5.000 escolas em 140 países, adotou estruturas de design curricular coerentes com o backward design em seus programas ¹². O sistema de educação de Singapura, consistentemente apontado como referência em avaliações como o PISA, opera com uma filosofia de design curricular que converge com os princípios de Wiggins e McTighe: resultados claramente definidos, avaliações

¹¹ ASCD. *Understanding by Design Professional Development Workbook*. ASCD, 2004. (Dados de vendas referenciados em comunicações institucionais da ASCD, 2018.)

¹² INTERNATIONAL BACCALAUREATE ORGANIZATION. *Principles into Practice: A Guide to the Approaches to Teaching and Learning*. IBO, 2017.

autênticas e sequências de instrução subordinadas aos dois.

No campo das certificações profissionais, o impacto é visível nos sistemas de credenciamento que migraram de modelos baseados em currículo para modelos baseados em competência. O movimento de *competency-based education* nos Estados Unidos, formalizado por iniciativas do Departamento de Educação nos anos 2010, opera na mesma lógica: define o que o profissional deve ser capaz de fazer, estabelece as evidências de que é capaz, e permite que os candidatos demonstrem competência por caminhos variados, incluindo experiência prévia validada. Western Governors University e College for America construíram programas inteiros sobre essa base.

4.5.2. O Fracasso como Evidência Negativa

Um indicador indireto do impacto do UbD é o fracasso sistemático das reformas que ignoraram seus princípios. A *No Child Left Behind Act* de 2001¹³ introduziu testes padronizados obrigatórios em linguagem e matemática para todos os alunos de 3ª a 8ª série, com consequências severas para escolas que não atingissem as metas. O resultado, documentado em múltiplos estudos¹⁴, foi o fenômeno que a pesquisa educacional passou a chamar de *teaching to the test*: professores reorganizaram seus currículos para maximizar o desempenho nas provas padronizadas, abandonando conteúdos não testados e reduzindo o tempo dedicado a competências de ordem superior que não aparecem em testes de múltipla escolha.

O problema não era a ênfase em resultados; era que os resultados foram especificados sem os critérios do segundo estágio. Definir um resultado como “nota satisfatória no teste X” não é o mesmo que definir o que o aluno deve ser capaz de fazer. A confusão entre métrica e competência

¹³ UNITED STATES CONGRESS. *No Child Left Behind Act of 2001* (Pub. L. 107-110). Washington, D.C.: U.S. Government Printing Office, 2002.

¹⁴ RAVITCH, Diane. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. Basic Books, 2010.

produziu o que qualquer leitura cuidadosa de *Understanding by Design* previu: alunos que passam no teste sem desenvolver a competência que o teste, supostamente, mede. A *No Child Left Behind* queria resultados; não queria backward design. São coisas diferentes.

4.5.3. *Corporate Learning em Escala*

No campo corporativo, a ATD (*Association for Talent Development*) estimou que empresas americanas gastaram mais de 180 bilhões de dólares em treinamento e desenvolvimento em 2021¹⁵. A proporção desse investimento que produz mudança verificável de comportamento no trabalho permanece perturbadoramente baixa: estimativas variam, mas há consenso de que menos de 20% do treinamento transfere para o ambiente de trabalho de forma mensurável⁶. A adoção de princípios de backward design em L&D produz resultados consistentemente superiores em estudos comparativos¹⁶, mas a adoção permanece parcial. A cultura de treinamento centrado em conteúdo é resistente, porque é mais fácil de produzir e mais imediatamente visível para compradores que não sabem o que perguntar.

Isso tem uma implicação direta para o design de produtos educacionais: o mercado cobra, eventualmente, o preço de prometer conteúdo em vez de competência.

4.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

¹⁵ ASSOCIATION FOR TALENT DEVELOPMENT (ATD). *State of the Industry Report 2022*. ATD Press, 2022.

¹⁶ SITZMANN, Traci; WEINHARDT, Justin M. Training Engagement Theory: A Multilevel Perspective on the Effectiveness of Work-Related Training. *Journal of Management*, v. 44, n. 2, p. 732–756, 2018.

4.6.1. Bloom: o Vocabulário de Precisão que o UbD Exige

A relação entre *Understanding by Design* e a Taxonomia de Bloom é de hierarquia funcional, mas a direção da dependência importa. Bloom fornece o vocabulário de especificação que o primeiro estágio exige, mas a taxonomia, por si só, não resolve o problema que o UbD ataca. É possível escrever objetivos de aprendizagem taxonomicamente precisos, usando os verbos corretos do nível adequado, e ainda assim operar no modelo convencional, definindo objetivos depois do conteúdo em vez de antes. Bloom especifica *o quê* com precisão; UbD determina *para quê* e *em que ordem*.

O que os dois sistemas compartilham é a separação entre conhecimento de reprodução e conhecimento de transferência. Bloom a articula como diferença de nível cognitivo: lembrar e compreender são reprodução; aplicar, analisar, avaliar e criar são transferência. Wiggins e McTighe articulam a mesma distinção como diferença entre cobertura e compreensão genuína. Abordagens distintas do mesmo problema. A integração produz design instrucional em que a hierarquia de resultados (UbD) é especificada com a precisão taxonômica que torna cada resultado mensurável (Bloom).

4.6.2. Gagné: a Estrutura Interna de Cada Sessão

Se Bloom serve ao primeiro estágio, o modelo de Robert Gagné serve ao terceiro, fornecendo uma teoria da estrutura interna das experiências de aprendizagem que o UbD não desenvolve.

Os Nove Eventos de Instrução de Gagné¹⁷, formulados originalmente em *The Conditions of Learning* (1965), descrevem a sequência interna de um episódio de aprendizagem eficaz: ganhar a atenção do aprendiz; in-

¹⁷ GAGNÉ, Robert M. *The Conditions of Learning*. Holt, Rinehart and Winston, 1965. 4. ed.: *The Conditions of Learning and Theory of Instruction*. Holt, Rinehart and Winston, 1985.

formar o objetivo; estimular a recordação de conhecimento anterior relevante; apresentar o conteúdo; fornecer orientação; elicitare a performance (prática); fornecer retorno corretivo; avaliar a performance; e aumentar a retenção e a transferência. Cada evento tem fundamento nas condições cognitivas necessárias para que a aprendizagem ocorra: atenção como condição de processamento, conhecimento prévio como andaime para o novo, prática como condição de consolidação, retorno como condição de correção.

O UbD não especifica a estrutura interna das aulas, apenas a lógica do design curricular. Gagné opera no nível que o UbD deixa em aberto: dado que você definiu o resultado, as evidências e o conteúdo, como estruturar cada sessão para garantir que a aprendizagem ocorra? A articulação dos dois sistemas produz design instrucional em dois níveis: macro (o curso como um todo, via UbD) e micro (cada sessão de aprendizagem, via Gagné). Uma limitação vale nomear: os Nove Eventos foram formulados em um contexto de instrução direta e controlada, e sua aplicação a ambientes mais abertos ou baseados em projetos exige adaptação. O princípio subjacente, porém, permanece robusto.

4.6.3. *Schwartz: o Ponto de Partida que o UbD Não Especifica*

A conexão com Eugene Schwartz¹⁸ é menos óbvia, mas operacionalmente relevante para quem projeta produtos educacionais com dimensão comercial. Schwartz desenvolveu o conceito de *níveis de consciência do comprador*: o comprador pode estar inteiramente inconsciente do problema, consciente do problema mas não da solução, consciente da solução mas não do produto específico, ou plenamente informado sobre o mercado. A comunicação eficaz de marketing precisa estar calibrada para o nível de consciência do comprador, porque uma mensagem que

¹⁸ SCHWARTZ, Eugene. *Breakthrough Advertising*. Bottom Line Books, 1966. Reimpresso por Boardroom Inc., 2004.

pressupõe que o receptor está mais informado do que é resulta em incompreensão; pressupor que está menos informado resulta em irrelevância.

A transposição para o design instrucional é direta: o resultado desejado do primeiro estágio precisa estar calibrado para o nível de consciência e competência do aprendiz que chega ao produto. Um produto que define como resultado “o aluno vai criar estratégias de posicionamento competitivo” para um público que ainda não compreende o que é posicionamento está pressupondo um nível de competência anterior que o aprendiz não tem. O gap entre resultado desejado e nível de entrada não é obstáculo pedagógico; é falha de design. O UbD especifica a direção; Schwartz especifica o ponto de partida calibrado. Os dois são necessários para que o produto educacional seja, ao mesmo tempo, pedagogicamente eficaz e comercialmente honesto.

5

A PIRÂMIDE QUE MUDA COMO SE ENSINA: WILLIAM GLASSER E A RETENÇÃO REAL

“Nós aprendemos 10% do que vemos, 20% do que ouvimos, 30% do que vemos, 50% do que vemos e ouvimos, 70% do que discutimos, 80% do que experienciamos pessoalmente, e 95% do que ensinamos a outros.” William Glasser, The Quality School, 1990¹

5.1. CONTEXTO HISTÓRICO

A psiquiatria americana dos anos 1950 e 1960 vivia uma tensão que raramente aparece nas narrativas lineares da história intelectual: de um lado, o modelo psicanalítico freudiano, dominante nas faculdades de medicina e nos consultórios de elite, comprometido com a exploração do passado inconsciente como chave para o sofrimento presente; de outro,

¹ GLASSER, William. *The Quality School: Managing Students Without Coercion*. Harper & Row, 1990. 2. ed. revisada: HarperPerennial, 1998.

uma insatisfação crescente de clínicos que trabalhavam com populações marginalizadas, jovens delinquentes, internos de hospitais psiquiátricos estaduais, pacientes sem recursos para anos de diván. O conflito não era apenas teórico. Era uma disputa sobre o que a psiquiatria podia oferecer a quem, em quanto tempo, e com que evidência de resultado.

William Glasser nasceu em 1925 em Cleveland, Ohio, formou-se em engenharia química antes de mudar para medicina e concluiu sua especialização em psiquiatria na UCLA em 1957. Não chegou ao problema da educação por interesse primário. Chegou por necessidade clínica. Seu trabalho na Ventura School for Girls, instituição correcional para jovens delinquentes no sul da Califórnia, tornou insustentável a aplicação do modelo terapêutico convencional. As jovens não tinham tempo, não tinham recursos para o processo analítico prolongado, e o ambiente institucional era estruturalmente incompatível com a produção de insight retrospectivo como mecanismo de mudança. Glasser precisava de outra coisa.

O que desenvolveu, descrito inicialmente em *Reality Therapy* (1965)², era uma inversão deliberada da premissa freudiana. Em vez de explorar o passado para explicar o presente, propunha focar no comportamento atual e nas escolhas disponíveis agora. O fundamento teórico viria mais tarde, mas a intuição clínica era clara: pessoas mudam quando percebem que têm controle sobre suas escolhas presentes, não quando compreendem as origens de seus padrões passados. A denominação mudaria com o tempo, de Control Theory para Choice Theory, mas o núcleo permaneceria estável: comportamento é escolha, escolha tem propósito, propósito é a satisfação de necessidades básicas que Glasser identificaria como sobrevivência, pertencimento, poder, liberdade e diversão.

A educação entrou nessa teoria por consequência lógica. Se o comportamento é escolha orientada por necessidades, e se a escola é o am-

² GLASSER, William. *Reality Therapy: A New Approach to Psychiatry*. Harper & Row, 1965.

biente onde crianças e adolescentes passam a maior parte de seus anos formativos, então a pergunta sobre por que tantos alunos escolhem não aprender, não se engajar, não completar o que começaram, não era uma pergunta sobre déficit cognitivo ou desordem emocional. Era uma pergunta sobre design: o sistema escolar estava projetado de uma forma que não satisfazia as necessidades que motivam o comportamento humano. *Schools Without Failure* (1969)³ foi o primeiro diagnóstico sistemático desse argumento, e também o mais incômodo: a escola, como instituída, não estava falhando apesar de seu design, mas por causa dele.

Nos anos 1970 e 1980, Glasser aprofundava sua teoria em *Control Theory in the Classroom* (1986)⁴, enquanto pesquisadores como Robert Gagné e Benjamin Bloom construíam arquiteturas formais de objetivos educacionais e condições de aprendizagem. Os dois projetos intelectuais tinham pouca comunicação direta, o que, retrospectivamente, é uma das ironias mais produtivas da história educacional: Glasser perguntava por que o aluno se engaja; Bloom e Gagné perguntavam como estruturar o que o aluno aprende. A síntese viria mais tarde, por iniciativa de outros.

The Quality School (1990)¹ marca o momento em que a teoria clínica se integra a uma proposta pedagógica concreta. É nesse livro que a chamada Pirâmide de Retenção, ou Cone de Aprendizagem em sua versão mais rigorosa historicamente, aparece associada ao argumento central: a forma de aprender, não o conteúdo aprendido, determina o quanto permanece. O que esse argumento afirma exatamente, de onde vêm os dados que o sustentam, e por que ele resiste às críticas que recebeu são as questões que a próxima seção trata antes de qualquer outra.

³ GLASSER, William. *Schools Without Failure*. Harper & Row, 1969.

⁴ GLASSER, William. *Control Theory in the Classroom*. Harper & Row, 1986.

5.2. O PROBLEMA QUE GLASSER ESTAVA RESOLVENDO

Há um fenômeno que qualquer professor experiente reconhece, mas que raramente é tratado como o problema central que deveria ser: o aluno que esteve presente durante toda a aula, respondeu às perguntas, saiu satisfeito, e três dias depois não recorda praticamente nada do que foi ensinado. Não é falha de memória patológica. É o funcionamento normal de um sistema cognitivo que processa informação recebida passivamente como ruído de fundo, preservando apenas os fragmentos que encontraram algum ponto de ancoragem na estrutura de conhecimento existente.

Glasser chegou a esse diagnóstico por uma rota específica. Seu trabalho com jovens em contextos correcionais havia produzido uma observação recorrente: o ensino convencional não apenas falhava em produzir aprendizagem duradoura, falhava de uma forma que o aluno percebia como fracasso pessoal. A escola estava estruturada para produzir mais perdedores do que vencedores, como ele argumentou em *Schools Without Failure*³, e a lógica interna desse sistema era a da avaliação comparativa, da hierarquia de desempenho, da passividade como norma e da cobertura de conteúdo como métrica de sucesso. Exatamente os mecanismos que Wiggins e McTighe, duas décadas depois, identificariam como raízes estruturais do problema de design.

A ilusão da transmissão está no núcleo desse problema. O professor apresenta o conteúdo; o conteúdo é recebido; o aluno foi ensinado. A cadeia parece coerente até que se examina a diferença entre receber informação e processá-la de forma que ela se integre à memória de longo prazo como conhecimento acionável. Essas são operações cognitivas distintas. A primeira é passiva e praticamente automática; a segunda exige esforço, conexão, elaboração, alguma forma de ativação do aprendiz como agente do processo.

O modelo escolar convencional, na formulação de Glasser, não satisfaz a necessidade de poder (competência, realização), nem a de diversão (engajamento, descoberta), nem a de pertencimento (aprendizagem colaborativa). Satisfaz, quando muito, a necessidade de sobrevivência, na medida em que a aprovação é condição para a continuidade social. Motivação baseada em sobrevivência produz aprendizagem instrumental e de curto prazo, aquela que dura até a prova e desaparece depois. O aluno aprende para não ser punido, não porque perceba relevância no que está aprendendo.

A consequência é um ciclo descrito em *The Quality School*¹: o conteúdo não é percebido como relevante; o aprendiz não se engaja; a aprendizagem é superficial; a avaliação confirma o não aprendido; o sistema interpreta isso como falha individual; a identidade de não aprendiz se instala. O problema não é o aluno. É o projeto do sistema que produziu esse resultado como resultado esperado.

O diagnóstico tem uma implicação metodológica precisa: se o engajamento do aprendiz é condição, não variável opcional, da aprendizagem real, então o design de qualquer experiência educacional precisa começar pela pergunta sobre como o aprendiz será ativado, não pela pergunta sobre qual conteúdo será transmitido. A forma precede o conteúdo. Não no sentido estético, mas no sentido causal: a modalidade de aprendizagem determina o processamento cognitivo, que determina a fixação, que determina se alguma competência foi de fato desenvolvida.

Insuficiente transmitir. Necessário ativar.

5.3. IDEIA CENTRAL

5.3.1. A Pirâmide de Retenção: Estrutura, Atribuição e Robustez do Princípio

A chamada Pirâmide de Retenção circula na literatura educacional e nos materiais de treinamento corporativo com uma persistência que é, ela mesma, um fenômeno digno de análise. Taxas de consolidação por modalidade de aprendizagem: leitura, 5 a 10%; aula expositiva, 5%; material audiovisual, 20%; demonstração, 30%; discussão em grupo, 50%; prática ativa, 75 a 80%; ensinar outros, 90%. As variações numéricas entre versões são pequenas e consistentes no que importa: a progressão das modalidades passivas para as ativas produz retorno crescente.

A questão da autoria é genuinamente complexa. O antecessor direto da pirâmide é o *Cone of Experience* (Cone de Experiência) de Edgar Dale, publicado em 1946 em *Audio-Visual Methods in Teaching*⁵. Dale propunha uma hierarquia de formatos de aprendizagem, da mais abstrata (leitura, linguagem verbal) à mais concreta (experiência direta), sem atribuir percentuais específicos. Os números apareceram em versões posteriores do cone, e sua origem está associada ao National Training Laboratories (NTL) de Bethel, Maine, que desde os anos 1950 conduzia pesquisas sobre dinâmicas de grupo e aprendizagem experiencial⁶. As taxas foram amplamente circuladas nos materiais do NTL, e o nível de documentação primária para os valores específicos é limitado.

Glasser não inventou a hierarquia no sentido técnico. O que fez foi integrar esses dados de retenção, qualquer que seja sua origem empírica precisa, em um argumento teórico coerente sobre por que as modalidades ativas produzem mais fixação do que as passivas. Esse argumento tem fundamento independente nas ciências cognitivas: processamento

⁵ DALE, Edgar. *Audio-Visual Methods in Teaching*. Dryden Press, 1946. 3. ed.: Dryden Press/Holt, Rinehart and Winston, 1969.

⁶ NATIONAL TRAINING LABORATORIES INSTITUTE. *The Learning Pyramid*. Bethel, ME: NTL Institute for Applied Behavioral Science, c. 1960s. (Dados de retenção circulados em materiais institucionais; base metodológica primária de acesso limitado. Ver discussão crítica em Thalheimer, 2006.)

elaborativo, teoria da codificação dupla, o efeito de geração, a hipótese do processamento desejável de dificuldade de Robert Bjork⁷. A pirâmide, como representação, é um artefato pedagógico imperfeito. O princípio que ela ilustra é empiricamente sólido.

5.3.2. *Por que as Modalidades Ativas Retêm Mais*

O mecanismo cognitivo subjacente à progressão da hierarquia é, em grande medida, compreendido. A memória de longo prazo opera por consolidação e recuperação, e ambos os processos são facilitados pelo processamento elaborativo: a atividade mental pela qual o aprendiz conecta informação nova à estrutura de conhecimento existente, produz exemplos, resolve problemas, articula explicações, antecipa objeções⁸. Quanto mais elaboração, maior a consolidação. Quanto maior a consolidação, mais robusto o traço de memória e mais flexível a recuperação posterior.

A leitura e a aula expositiva produzem pouca elaboração porque o aprendiz está predominantemente recebendo. O esforço cognitivo é de decodificação e compreensão imediata, não de integração ativa. A discussão produz mais elaboração porque exige que o aprendiz articule sua compreensão, responda a objeções, ajuste sua posição. A prática ativa exige a aplicação do conceito em contexto, exercício de recuperação com retroalimentação imediata. Ensinar é a modalidade de maior elaboração porque exige que o aprendiz organize o conhecimento de forma transmissível, identifique o que não compreende bem o suficiente para explicar, e responda às perguntas de outro aprendiz.

Há uma simetria notável entre essa hierarquia e o que Glasser descreve como satisfação das necessidades básicas. Aulas expositivas não

⁷ BJORK, Robert A. Memory and Metamemory Considerations in the Training of Human Beings. In: METCALFE, Janet; SHIMAMURA, Arthur P. (Orgs.). *Metacognition: Knowing About Knowing*. MIT Press, 1994. p. 185–205.

⁸ ROEDIGER, Henry L.; KARPICKE, Jeffrey D. The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, v. 1, n. 3, p. 181–210, 2006.

acionam a necessidade de poder (competência percebida) nem de diversão. A discussão mobiliza pertencimento e, em algum grau, realização. A prática ativa atende poder e liberdade. Ensinar satisfaz as três simultaneamente: poder, pertencimento, diversão. A progressão de fixação acompanha a progressão de satisfação de necessidades, o que sugere que o engajamento motivacional e o processamento cognitivo não são variáveis independentes: são dimensões do mesmo processo.

5.3.3. Escolha de Qualidade e Relevância Percebida

O conceito de escolha de qualidade (*quality work*), desenvolvido em *The Quality School*¹ e aprofundado em *Choice Theory* (1998)⁹, constitui o segundo eixo da contribuição de Glasser. A premissa é que o aprendiz não pode produzir aprendizagem de qualidade sem perceber que o que está aprendendo tem relevância real para seus objetivos. Não relevância futura e abstrata (“isso vai ser útil na sua vida adulta”), mas conexão imediata e específica entre o conteúdo e alguma necessidade ou objetivo que o aprendiz reconhece como seu agora.

Isso não é um princípio motivacional no sentido superficial. É uma afirmação sobre a estrutura do processamento cognitivo. Informação percebida como irrelevante é descartada pelo sistema de memória de trabalho com eficiência considerável: ela não chega a ser elaborada, porque o sistema cognitivo não aloca recursos de atenção para o que não apresenta valor de sobrevivência ou realização. A relevância não é um bônus que facilita a aprendizagem. É uma condição de acesso ao processamento que produz permanência.

Glasser articula isso em termos de seu modelo de cinco necessidades: o aprendiz engaja com o que percebe como caminho para satisfazer poder, pertencimento, liberdade ou diversão. O conteúdo que não se co-

⁹ GLASSER, William. *Choice Theory: A New Psychology of Personal Freedom*. HarperCollins, 1998.

necta a nenhuma dessas necessidades é percebido como obstáculo, ou, na melhor das hipóteses, como neutro. Neutro é insuficiente para produzir o processamento elaborativo que a consolidação exige.

A implicação para o design instrucional é precisa: antes de determinar qualquer sequência de conteúdo, quem projeta a experiência educacional deve responder à pergunta sobre como o aprendiz vai perceber a relevância imediata daquilo que está sendo ensinado. Não como argumento retórico de abertura de aula. Como elemento estrutural do design, embutido na forma como os conceitos são apresentados, nas atividades que os acompanham, nos problemas que o aluno é convidado a resolver. Relevância que precisa ser anunciada explicitamente já falhou em ser construída.

5.3.4. A Crítica Metodológica e a Robustez do Argumento

Qualquer tratamento honesto do legado de Glasser precisa incluir a crítica metodológica que a pirâmide de retenção recebeu, especialmente a partir dos anos 2000. Will Thalheimer, pesquisador de aprendizagem e desempenho, conduziu análise sistemática em 2006 buscando as fontes primárias dos percentuais atribuídos ao NTL e a Glasser¹⁰. A conclusão foi que os dados específicos não tinham suporte em estudos controlados publicados, e que os percentuais haviam sido propagados como fato empírico sem a base metodológica que um fato empírico exige.

Essa crítica é metodologicamente correta e pedagogicamente inconsequente. Inconsequente não porque os percentuais específicos não importem, mas porque o princípio central que a hierarquia ilustra, que modalidades que exigem processamento elaborativo produzem melhor fixação do que modalidades de recepção passiva, tem suporte independente robusto na psicologia cognitiva e nas ciências da aprendizagem⁸.

¹⁰ THALHEIMER, Will. *People Remember 10%, 20%.. Oh Really?* Work-Learning Research, 2006. Disponível em: https://www.worklearning.com/2006/05/01/people_remember/. Acesso em: jan. 2025.

Os números específicos são, provavelmente, aproximações razoáveis de ordens de magnitude verdadeiras, não medições precisas de constantes psicológicas.

O que a crítica de Thalheimer ilumina é um padrão recorrente na circulação de ideias educacionais: princípios corretos ganham representações quantitativas precisas que excedem a evidência disponível, porque a precisão aparente facilita a comunicação e a persuasão. A versão honesta do argumento é menos memorável do que a versão quantificada. Essa tensão o campo não resolveu.

5.4. EVOLUÇÃO POSTERIOR DA IDEIA

O princípio estabelecido por Glasser permaneceu, por algum tempo, mais influente como diagnóstico do que como agenda de reforma. A transposição para prática pedagógica organizada requereu um campo intermediário, e esse campo tomou forma nos anos 1980 e 1990 sob o rótulo de aprendizagem ativa.

O trabalho de Charles Bonwell e James Eison, *Active Learning: Creating Excitement in the Classroom* (1991)¹¹, sistematizou o campo e forneceu uma revisão de literatura que demonstrava consistência entre estudos: estratégias que envolvem os estudantes ativamente em tarefas de pensar produzem ganhos mensuráveis de compreensão quando comparadas a métodos predominantemente expositivos.

O que Bonwell e Eison não resolveram, e que permaneceria como tensão persistente no campo, é a especificidade das condições em que a aprendizagem ativa produz ganhos. O problema não é que o princípio funcione melhor em média. É que “aprendizagem ativa” é uma categoria

¹¹ BONWELL, Charles C.; EISON, James A. *Active Learning: Creating Excitement in the Classroom*. ASHE-ERIC Higher Education Report n. 1. George Washington University, 1991.

suficientemente ampla para incluir práticas muito distintas com mecanismos muito distintos, e a evidência granular sobre qual abordagem funciona melhor para qual tipo de conteúdo, com qual perfil de aprendiz, em qual estágio de desenvolvimento de competência, é consideravelmente menos robusta do que a narrativa popular do campo sugere.

Isso não invalida o princípio. Invalida a simplificação.

Richard Hake, em um estudo de 1998 que se tornou referência no campo, comparou 6.542 estudantes de física em 62 cursos e encontrou que cursos com métodos interativos produziam ganhos de compreensão conceitual que eram, em média, o dobro dos produzidos por cursos tradicionais de aula expositiva¹². O tamanho do efeito era suficientemente robusto para sustentar a conclusão, mesmo com variância considerável entre contextos. O estudo de Hake forneceu, para o movimento de aprendizagem ativa, a evidência em larga escala que a pirâmide de Glasser não tinha: não apenas uma hierarquia plausível, mas um dado controlado de impacto real.

O campo continuou a se ramificar. No ensino superior, as pedagogias de sala de aula invertida (*flipped classroom*) redistribuíram o tempo de exposição passiva para fora da sala, reservando o tempo presencial para práticas de aplicação e discussão. Na formação profissional, modelos como o *problem-based learning* substituíram a aula expositiva pelo problema como unidade organizadora do currículo. Todas essas variantes partem do mesmo axioma: a elaboração ativa do aprendiz não é um complemento à instrução; é a condição da instrução. O que os contextos variaram foi a forma de produzir essa elaboração, não o compromisso com ela. E é precisamente esse compromisso que o impacto real do princípio permite medir.

¹² HAKE, Richard R. Interactive-Engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses. *American Journal of Physics*, v. 66, n. 1, p. 64–74, 1998.

5.5. IMPACTO REAL NO MUNDO

5.5.1. Educação Médica e o Redesign do Internato

Um dos setores onde o princípio de aprendizagem ativa foi mais consequentemente aplicado é a educação médica. A residência médica, como modelo de formação, é estruturalmente um experimento em aprendizagem experiencial: o residente aprende fazendo, sob supervisão crescentemente reduzida, com pacientes reais, em condições de pressão real.

O problema que emergiu nos anos 1980 e 1990 foi a combinação de aprendizagem experiencial com carga de trabalho extrema: residentes com turnos de 36 horas ou mais, fadiga comprometendo tanto o desempenho clínico quanto a consolidação do aprendido. O Institute of Medicine americano, em seu relatório *To Err is Human* (1999)¹³, identificou a fadiga dos profissionais de saúde como fator de risco em erros médicos. A resposta regulatória foi a limitação de horas de trabalho para residentes, implementada pelo Accreditation Council for Graduate Medical Education (ACGME) em 2003.

A limitação de horas criou um problema pedagógico: menos exposição a casos significava menos aprendizagem experiencial. A solução desenvolvida por programas de vanguarda foi o uso sistemático de simulação clínica de alta fidelidade, um ambiente onde o residente pratica procedimentos e tomada de decisão em condições que replicam a realidade sem expor pacientes a risco. A lógica é diretamente derivada do princípio de Glasser: a simulação é prática ativa com feedback imediato, o que produz fixação superior à observação ou à aula sobre procedimentos. Estudos comparativos confirmam a superioridade da simulação sobre métodos passivos em procedimentos técnicos e em habilidades de trabalho em equipe¹⁴.

¹³ INSTITUTE OF MEDICINE. *To Err Is Human: Building a Safer Health System*. National Academies Press, 1999.

¹⁴ COOK, David A. et al. Comparative Effectiveness of Instructional Design Features in Simulation-Based Education:

5.5.2. O Mercado Digital e a Crise de Retenção em Escala

No campo da educação digital, a tensão entre escalabilidade e eficácia pedagógica é estrutural, e ela manifesta a mesma contradição que Glasser identificou no modelo escolar convencional. O mercado global de e-learning foi estimado em 399 bilhões de dólares em 2022, com projeção de crescimento consistente¹⁵. O formato predominante nesse mercado é o vídeo assíncrono: alta escala, custo marginal próximo de zero, formato que reproduz fielmente a aula expositiva passiva.

O dado mais citado no debate sobre eficácia é a taxa de conclusão de plataformas de vídeo-aula: menos de 15% em Massive Open Online Courses (MOOCs), com algumas plataformas reportando taxas ainda menores em cursos não certificados¹⁶. A interpretação dominante é que o formato passivo de consumo de vídeo replica, em escala, exatamente o problema da aula expositiva convencional: o aprendiz consome sem elaborar, e sem elaboração não há consolidação. A taxa de conclusão é, nessa leitura, um indicador de engajamento comprometido desde o design.

Essa crítica é legítima, mas exige precisão. A taxa de conclusão de MOOCs é uma métrica de comportamento, não de aprendizagem. Estudantes que acessam materiais sem intenção de completar o curso, que usam o conteúdo como referência pontual, que abandonam após obter o que precisavam, são contabilizados como não conclusões. A questão sobre o que foi retido por quem completou o curso não é, em geral, o que os dados de plataforma medem. Isso não invalida a preocupação com permanência do aprendido, mas exige que o argumento seja construído com mais precisão do que a estatística bruta permite.

Systematic Review and Meta-analysis. *Medical Teacher*, v. 35, n. 1, p. e844–e875, 2013.

¹⁵ GLOBAL MARKET INSIGHTS. *E-Learning Market Size, By Technology, By Application, By End-use, Growth Potential, Competitive Market Share & Forecast, 2022–2032*. Global Market Insights Inc., 2022.

¹⁶ JORDAN, Katy. MOOC Completion Rates: The Data. 2015. Disponível em: <http://www.katyjordan.com/MOOCproject.html>. Acesso em: jan. 2025.

A resposta das plataformas mais sofisticadas foi a incorporação de elementos de ativação dentro de formatos predominantemente passivos: quizzes de recuperação espaçados ao longo do vídeo, exercícios de aplicação imediata entre módulos, fóruns de discussão estruturada, projetos práticos com feedback de instrutores ou pares. O boom dos cohort-based courses (CBCs) nos anos 2010, com plataformas como Maven, On Deck e Reforge construindo modelos de negócio em torno de grupos de aprendizes que avançam juntos com discussão estruturada e aplicação imediata, foi a versão mais radical dessa resposta: em vez de corrigir o vídeo passivo com elementos ativos, substituir a lógica do produto inteiro. A evidência sobre quizzes de recuperação espaçada é robusta ⁸; sobre fóruns de discussão, mista; sobre aprendizagem em coorte, promissora mas dependente da qualidade da facilitação. O design instrucional ótimo para plataformas digitais de grande escala continua sendo uma questão em aberto.

5.5.3. *Treinamento Corporativo e o Modelo 70-20-10*

No campo do desenvolvimento de liderança, o modelo 70-20-10, desenvolvido por Morgan McCall, Michael Lombardo e Robert Eichinger no Center for Creative Leadership nos anos 1980 ¹⁷, propõe que 70% do desenvolvimento de líderes ocorre por experiências no trabalho, 20% por interações com outros e feedback, e apenas 10% por treinamentos formais. O modelo não é uma refutação de Glasser; é uma aplicação setorial do mesmo princípio.

A implicação prática é que organizações que investem a maior parte de seu orçamento de L&D em treinamentos formais presenciais ou online estão alocando recursos na porção de menor retorno. O treinamento formal é necessário para instalar o quadro conceitual; a transferência real

¹⁷ MCCALL, Morgan W.; LOMBARDO, Michael M.; MORRISON, Ann M. *The Lessons of Experience: How Successful Executives Develop on the Job*. Free Press, 1988.

ocorre na prática com feedback. Essa distinção entre instalação de conhecimento e desenvolvimento de competência é exatamente o que Glasser articulou, em linguagem distinta, como a diferença entre a escola que transmite e a escola que transforma. Que a pesquisa corporativa tenha chegado ao mesmo ponto por outra rota, três décadas depois, confirma a extensão e a força do diagnóstico original.

5.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

A utilidade analítica de Glasser para a APE não está em suas proposições de forma isolada: está na articulação com os outros autores do livro que respondem às perguntas que ele deixa em aberto. Wiggins e McTighe definem o que precisa ser praticado: o OUTPUT. Gagné especifica como estruturar cada episódio de prática internamente. Knowles explica o que torna o adulto capaz de se engajar com relevância. E Ericsson responde à pergunta sobre a qualidade que a prática precisa atingir para produzir competência real, e não apenas familiaridade com o conteúdo.

5.6.1. *Glasser e Ericsson: o que a Prática Precisa Ser*

K. Anders Ericsson dedicou décadas ao estudo do desenvolvimento de expertise, e a conclusão central de seu programa de pesquisa, detalhado em *Peak: Secrets from the New Science of Expertise* (2016)¹⁸, é que o que diferencia experts de praticantes experientes não é a quantidade de prática, mas sua qualidade. Prática deliberada é prática com objetivos específicos, feedback imediato sobre o desempenho em relação a esses objetivos, e operação consistente na zona de desenvolvimento proximal, onde a dificuldade é suficientemente alta para exigir esforço sem ser tão alta que

¹⁸ ERICSSON, K. Anders; POOL, Robert. *Peak: Secrets from the New Science of Expertise*. Houghton Mifflin Harcourt, 2016.

produza paralisia.

A relação com Glasser é de complementaridade precisa. Glasser estabelece a hierarquia de modalidades: mais elaboração, mais fixação. Ericsson especifica o que a elaboração precisa ser para produzir desenvolvimento de competência real, em vez de mera repetição de padrões já consolidados. A diferença entre prática e prática deliberada é a diferença entre o músico que toca o mesmo repertório por décadas sem melhorar e o que trabalha especificamente as passagens em que trava. Ambos praticam. Apenas um desenvolve.

Para o design de produtos educacionais, a articulação dos dois é operacionalmente crítica. Não basta incluir prática; ela precisa ser projetada com os elementos que Ericsson identifica como condições do desenvolvimento: objetivos específicos, feedback qualificado e imediato, progressão de dificuldade. Prática sem feedback é repetição. Repetição sem progressão é estagnação.

5.6.2. Glasser e Knowles: Relevância como Condição Compartilhada

A convergência entre Glasser e Malcolm Knowles é uma das mais diretas e consequentes do livro. O segundo princípio da andragogia de Knowles, a orientação para aprendizagem, afirma que adultos aprendem mais eficazmente quando o conteúdo é apresentado em função de problemas reais que precisam resolver, não em função de disciplinas ou tópicos abstratos¹⁹. O adulto não aprende porque o conteúdo é importante em abstrato; aprende porque percebe conexão entre o conteúdo e uma situação que enfrenta agora.

Glasser chega ao mesmo ponto por outra rota. Sua teoria das necessidades prevê que o engajamento real, o tipo que produz processamento elaborativo e retenção de longo prazo, ocorre quando o aprendiz percebe

¹⁹ KNOWLES, Malcolm S.; HOLTON, Elwood F.; SWANSON, Richard A. *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. 8. ed. Routledge, 2015.

o conteúdo como caminho para a satisfação de uma necessidade presente. A necessidade de poder (competência, realização), em particular, é o motor principal do aprendizado autodirigido no adulto. Quando o conteúdo não se conecta a uma necessidade reconhecida como urgente, o adulto desengaja, não por falta de disciplina, mas porque o sistema cognitivo está funcionando de forma adaptativa: não alocar recursos para o que não apresenta valor de retorno.

O diagnóstico compartilhado tem uma implicação que alcança o design comercial do produto, não apenas sua pedagogia: o ponto de partida, sua promessa e sua estrutura de entrada, precisa estabelecer a relevância antes que qualquer conteúdo seja entregue. O aprendiz precisa saber, com especificidade, o que será capaz de fazer de diferente ao final que não conseguia ao início. Sem isso, o processamento do conteúdo subsequente fica comprometido desde a abertura.

5.6.3. *Glasser e Gagné: Retenção como Produto de Sequência*

O modelo de Gagné dos Nove Eventos de Instrução²⁰ é, entre outras coisas, uma teoria sobre como maximizar a probabilidade de consolidação dentro de um episódio de aprendizagem controlado. Cada evento tem fundamento nas condições cognitivas necessárias: atenção como condição de processamento inicial, evocação de conhecimento prévio como andaime para o novo, prática como condição de consolidação, feedback como condição de correção e refinamento.

O que a hierarquia de Glasser acrescenta a Gagné é uma perspectiva macro que os Nove Eventos não oferecem: não apenas como estruturar cada episódio de aprendizagem internamente, mas qual modalidade predominante o episódio deve adotar para maximizar retenção. Os dois frameworks operam em níveis diferentes, o que torna a integração pro-

²⁰ GAGNÉ, Robert M. *The Conditions of Learning and Theory of Instruction*. 4. ed. Holt, Rinehart and Winston, 1985.

dutiva sem que um invalide o outro.

Um produto educacional bem projetado usa o backward design de Wiggins e McTighe para definir o OUTPUT, usa a hierarquia de modalidades de Glasser para escolher as atividades que produzirão maior fixação para aquele OUTPUT específico, e usa os Nove Eventos de Gagné para estruturar cada atividade internamente. A cadeia é lógica e não redundante: cada framework resolve um nível de problema que os outros deixam em aberto. É sobre como operar essa cadeia de forma intencional que a próxima seção trata.

6

O QUE SIGNIFICA SABER: BENJAMIN BLOOM E A ARQUITETURA DA COMPETÊNCIA

—*“O achado mais impressionante é que, nas melhores condições de aprendizagem que conseguimos conceber, a tutoria individual, o estudante médio situa-se 2 desvios padrão acima do estudante médio instruído pelos métodos convencionais de grupo.”* Benjamin Bloom, Educational Researcher, 1984¹

6.1. CONTEXTO HISTÓRICO

O período que vai do final da Segunda Guerra ao início dos anos 1960 produziu, nos Estados Unidos, uma contradição que a história da educação ainda não absorveu completamente. De um lado, o sistema educacional americano expandia em escala sem precedente: a Lei de Reajustamento

¹ BLOOM, Benjamin S. The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, v. 13, n. 6, p. 4–16, 1984.

dos Militares de 1944, conhecida como G.I. Bill, havia enviado mais de dois milhões de veteranos para faculdades e universidades que não estavam, em nenhum sentido relevante, preparadas para recebê-los. De outro, os instrumentos disponíveis para avaliar o que esses estudantes aprendiam eram, na prática, instrumentos para medir o que eles conseguiam lembrar. A avaliação educacional americana dos anos 1940 e 1950 era, predominantemente, uma tecnologia de mensuração da memória de curto prazo, disfarçada de mensuração de competência.

A University of Chicago era um ambiente particular. Ralph Tyler, que ali dirigia o Departamento de Educação, havia articulado nos anos 1930 e 1940 o argumento de que o currículo deveria ser organizado a partir de objetivos comportamentais observáveis, não a partir de tradições disciplinares ou conveniências institucionais². O chamado Estudo dos Oito Anos, que Tyler coordenou entre 1933 e 1941, havia demonstrado que estudantes formados em escolas experimentais, organizadas em torno de objetivos precisos, performavam ao menos tão bem quanto, e frequentemente melhor do que, estudantes de escolas tradicionais nos exames de admissão universitária. A evidência era robusta. O campo a ignorava.

Benjamin Samuel Bloom chegou à Universidade de Chicago em 1940 para cursar o doutorado sob orientação de Tyler. Nascido em 1913 em Lansford, Pennsylvania, filho de imigrantes judeus poloneses, havia concluído a graduação e o mestrado na Pennsylvania State University com notável rapidez. O que o distinguiu, já no início, não era apenas inteligência analítica: era uma incapacidade constitutiva de aceitar ambiguidade normativa onde precisão era possível. Quando os examinadores da universidade produziam questões de prova, Bloom queria saber qual operação cognitiva exatamente a questão estava medindo. A pergunta parecia óbvia. Ninguém havia tentado respondê-la com rigor.

Em 1948, durante a convenção da American Psychological Associa-

² TYLER, Ralph W. *Basic Principles of Curriculum and Instruction*. University of Chicago Press, 1949.

tion em Boston, um grupo de examinadores universitários concordou que valia a pena tentar. A ideia era construir uma taxonomia de objetivos educacionais na área cognitiva que pudesse servir como linguagem comum para examinadores, professores e designers de currículo em todo o país. O grupo incluía Edward Furst, Walker Hill, David Krathwohl e Max Engelhart, entre outros. Bloom coordenaria o trabalho. Levaria oito anos.

Esse intervalo não é trivial. O lançamento do Sputnik em outubro de 1957 transformou a educação americana em questão de segurança nacional. O National Defense Education Act de 1958 despejou recursos federais em matemática, ciências e línguas. A demanda por objetivos educacionais precisos e mensuráveis, que em 1948 era interesse de especialistas, havia se tornado imperativo político. *Taxonomy of Educational Objectives*, Handbook I: Cognitive Domain, publicado em 1956³, chegou ao campo exatamente quando o campo mais precisava de uma gramática para articular o que as escolas deveriam produzir e como verificar que produziram.

A tensão intelectual do período era real. O behaviorismo, dominante na psicologia americana, descrevia aprendizagem em termos de respostas condicionadas a estímulos, uma linguagem que tornava difícil distinguir entre memorização e compreensão genuína. A revolução cognitiva, que ganharia tração com Jerome Bruner e outros a partir de 1960, ainda não havia articulado um vocabulário alternativo com a precisão que o trabalho de design educacional exigia. Bloom construía a ponte entre as duas: retinha do behaviorismo o compromisso com a observabilidade e a mensurabilidade, mas introduzia uma hierarquia de operações mentais que a escola de Skinner não conseguia acomodar.

O que viria depois, em 1984, era diferente em escala e em urgência.

³ BLOOM, Benjamin S. et al. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. David McKay Company, 1956.

O estudo publicado na revista *Educational Researcher* ² não era produto de um projeto colaborativo de longo prazo. Era a síntese de dois estudos de doutoramento, conduzidos pelos orientandos Joanne Anania e Joseph Burke, que haviam comparado os resultados de estudantes sob três condições distintas de instrução. O resultado era tão discrepante em relação ao que o campo considerava possível que o próprio Bloom o descreveu como surpreendente. Dois desvios padrão de diferença entre tutoria individual e instrução convencional de grupo. O dado mais citado na história da pesquisa educacional americana. E, como frequentemente acontece com achados perturbadores, o mais citado e o menos compreendido na prática.

6.2. O PROBLEMA QUE BLOOM ESTAVA RESOLVENDO

O dado mais citado na história da pesquisa educacional americana foi também o menos operacionalizado com rigor. Para entender por que, é necessário separar os dois problemas distintos que Bloom estava resolvendo simultaneamente, e que a recepção do seu trabalho tendeu a fundir.

O primeiro era um problema de mensuração. Os examinadores da University of Chicago constatavam, com regularidade, que as questões de prova produzidas pelos professores mediam, em sua esmagadora maioria, um único tipo de operação cognitiva: a recuperação de informação específica da memória. Definições, datas, nomes, fórmulas isoladas de qualquer contexto de aplicação. O resultado era que um estudante poderia saber, com precisão técnica, que a fotossíntese é o processo pelo qual plantas convertem luz solar em energia química, e ser completamente incapaz de explicar por que folhas mortas não a realizam. Saber a defini-

ção não era equivalente a entender o processo. Entender o processo não era equivalente a ser capaz de aplicá-lo em situações novas. Esses eram saltos cognitivos reais, não apenas diferenças de grau.

A taxonomia foi construída para dar nome a esses saltos.

O segundo problema era de distribuição. No contexto da expansão pós-guerra do ensino superior americano, tornava-se visível um padrão que os professores conheciam anedoticamente mas que ninguém havia quantificado com rigor: estudantes que dominavam conteúdos em aula, que conseguiam reproduzir com precisão o que haviam sido ensinados, falhavam sistematicamente quando confrontados com situações que exigiam transferência para contextos novos. A qualificação, o diploma, a nota, nenhum desses sinais previa com segurança razoável a capacidade de operar no campo real. Havia uma lacuna estrutural entre o conhecimento escolar e o conhecimento operacional.

John Carroll, em um artigo de 1963 que seria central para o trabalho posterior de Bloom, havia formulado uma hipótese direta: o grau em que um estudante aprende determinada tarefa é função da proporção entre o tempo que ele efetivamente usa para aprender e o tempo que ele necessita para aprender⁴. Se Carroll estava certo, a maioria das diferenças de desempenho entre estudantes não era produto de capacidade inata, mas de tempo e condições de instrução. Estudantes mais lentos em determinado conteúdo não eram estudantes incapazes: eram estudantes que precisavam de mais tempo e melhores condições. A questão pedagógica era, portanto, não “quem é capaz?” mas “como organizar a instrução para que mais estudantes alcancem domínio?”

Bloom desenvolveu o conceito de *mastery learning* a partir dessa premissa⁵. Se a maioria dos estudantes podia aprender a maioria dos conteúdos, desde que as condições fossem corretas, então o problema da

⁴ CARROLL, John B. A Model of School Learning. *Teachers College Record*, v. 64, n. 8, p. 723–733, 1963.

⁵ BLOOM, Benjamin S. *Mastery Learning*. In: BLOOM, James H. (Org.). *Mastery Learning: Theory and Practice*. Holt, Rinehart and Winston, 1971. p. 47–63.

educação de massa era um problema de design instrucional, não de distribuição de talento. Uma afirmação com implicações políticas evidentes num país que havia institucionalizado, por décadas, a seleção educacional como substituto da instrução.

O problema do 2 Sigma veio depois, e era mais agudo. As pesquisas de Anania e Burke mostravam que estudantes submetidos a tutoria individual chegavam a resultados que a instrução coletiva, mesmo bem organizada, não conseguia replicar. Não era uma questão de margem. Dois desvios padrão correspondem a passar do 50º para o 98º percentil de uma distribuição normal. Nenhuma variável isolada em educação produz esse efeito com consistência. O problema, que Bloom nomeou com exatidão, era encontrar métodos de instrução coletiva capazes de produzir o que a tutoria individual produzia. O problema não foi resolvido. Continua em aberto.

Por que a tutoria produz esse efeito, quais mecanismos cognitivos específicos ela ativa e por que esses mecanismos são estruturalmente incompatíveis com a instrução de grupo nas condições convencionais: a taxonomia de Bloom responde a essa questão com precisão que o dado estatístico, por si, não oferece. O ponto de partida não é o 2 Sigma, mas a arquitetura cognitiva que o explica.

6.3. IDEIA CENTRAL

6.3.1. *Os Seis Níveis e a Natureza Hierárquica da Cognição*

A taxonomia de Bloom não é uma escala de dificuldade. Esse é o equívoco mais persistente em sua circulação popular. Dificuldade é relativa ao aprendiz, ao conteúdo e ao contexto. O que a taxonomia descreve é uma hierarquia de operações cognitivas qualitativamente distintas: cada

nível pressupõe o anterior e o utiliza como instrumento para uma operação de ordem mais elevada ¹.

O nível 1, Lembrar, envolve recuperar informação específica da memória de longo prazo. Definir, listar, reconhecer, identificar. A operação é de recuperação, não de processamento. Um estudante que sabe que a taxa de desconto é o custo de oportunidade do capital está operando no nível 1.

O nível 2, Entender, exige interpretar, explicar, classificar, comparar, inferir. A operação central é a construção de significado, traduzir o que foi memorizado em uma representação compreendida. O estudante que consegue explicar por que a taxa de desconto afeta o valor presente de um fluxo de caixa opera no nível 2.

O nível 3, Aplicar, requer executar ou implementar procedimentos em situações específicas. A distinção em relação ao nível 2 é operacional: não basta compreender, é necessário usar. O estudante que calcula o valor presente de um projeto com uma taxa de desconto fornecida está no nível 3. Essa é a fronteira mais importante da taxonomia para produtos educacionais. É onde a maioria dos cursos termina de fato, embora raramente o anuncie dessa forma.

O nível 4, Analisar, envolve decompor o material em suas partes constituintes e identificar como elas se relacionam entre si e com a estrutura maior. Diferenciar, organizar, atribuir causalidade. O estudante que examina por que dois projetos com o mesmo valor presente líquido podem ter perfis de risco radicalmente distintos está analisando, não apenas calculando.

O nível 5, Avaliar, corresponde a fazer julgamentos baseados em critérios e padrões explícitos. Verificar, criticar, comparar alternativas segundo parâmetros definidos. Não é opinião: é avaliação com critérios formulados. Há uma diferença técnica real entre “este projeto parece

melhor” e “este projeto é superior segundo os critérios de retorno ajustado ao risco porque, dadas as variâncias observadas e a correlação com o portfólio existente, a relação risco-retorno marginal é favorável nas condições especificadas.”

O nível 6, Criar, é o único que não apenas usa conhecimento existente, mas produz algo novo a partir dele. Planejar, projetar, construir, gerar. É o nível onde o conhecimento se converte em agência produtiva.

A revisão de Anderson e Krathwohl (2001) ⁶ operou sobre a estrutura original em dois movimentos principais. Substituiu os substantivos por verbos, tornando explícita a natureza processual de cada patamar: Lembrar, Entender, Aplicar, Analisar, Avaliar, Criar. E inverteu os dois níveis superiores, colocando Criar acima de Avaliar, com o argumento de que a criação genuína exige a coordenação ativa de todas as operações anteriores, o que a torna cognitivamente mais exigente do que a avaliação de objetos existentes.

A revisão introduziu também uma segunda dimensão: o Tipo de Conhecimento, cruzando os seis processos cognitivos com quatro categorias, conhecimento factual, conceitual, procedimental e metacognitivo. O resultado é uma matriz 6×4 que permite especificar, com razoável precisão, o tipo exato de competência que um objetivo educacional visa desenvolver ⁷. “O estudante será capaz de explicar o conceito de risco” e “o estudante será capaz de aplicar procedimentos de análise de risco em um portfólio real” ocupam células completamente distintas na matriz. São objetivos pedagógicos separados, com designs instrucionais próprios, formas de avaliação distintas e critérios de sucesso incompatíveis entre si.

⁶ ANDERSON, Lorin W.; KRATHWOHL, David R. (Orgs.). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman, 2001.

⁷ KRATHWOHL, David R. A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, v. 41, n. 4, p. 212-218, 2002.

6.3.2. O Efeito 2 Sigma: Estrutura do Dado e do Problema

O artigo de 1984² reportava os resultados de três condições experimentais aplicadas a conteúdos do ensino médio de ciências e matemática. A condição de controle era instrução convencional de grupo: professor ensinando para uma turma, distribuição de resultados assumindo a forma aproximadamente normal esperada, com média no 50º percentil por definição. A segunda condição era *mastery learning*: instrução coletiva com avaliações formativas frequentes, feedback diagnóstico e atividades corretivas para estudantes que não atingissem o critério de domínio antes de avançar. Os estudantes nessa condição performaram, em média, no 84º percentil relativo ao grupo controle. Uma diferença de aproximadamente um desvio padrão.

A terceira condição era tutoria individual. Um tutor por estudante, capaz de adaptar explicações, detectar lacunas específicas, fornecer feedback imediato e calibrado ao estado cognitivo atual. Os estudantes sob tutoria performaram, em média, no 98º percentil em relação ao grupo controle. Dois desvios padrão.

Uma magnitude de efeito de $d = 2,0$ é, em qualquer ciência social, resultado que exige escrutínio antes de generalização. Bloom era suficientemente rigoroso para saber disso. Os dois estudos de Anania e Burke haviam chegado à mesma conclusão com populações distintas. A literatura mais ampla sobre tutoria convergia no mesmo sentido: Hattie, em síntese de mais de 800 meta-análises publicada em 2009⁸, identificou a tutoria individual como uma das intervenções de maior efeito documentado na literatura educacional, com tamanho de efeito médio de 0,40 acima de intervenções já eficazes. O dado de 2 sigma não era artefato de amostra.

O que produzia a diferença? Bloom identificou três mecanismos: fe-

⁸ HATTIE, John. *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Routledge, 2009.

edback imediato e específico ao desempenho, adaptação da instrução ao ritmo e ao perfil cognitivo do aprendiz individual, e participação ativa contínua que a dinâmica de um para um estruturalmente impõe. Cada um desses mecanismos é tecnicamente reproduzível. O problema é a combinação dos três em escala. Feedback individualizado exige alguém capaz de avaliar cada resposta de cada estudante com a frequência necessária para ser útil. Adaptação instrucional requer que o tutor conheça o estado cognitivo atual do aprendiz, não a média esperada para sua faixa etária. Participação ativa contínua é estruturalmente impossível quando um professor divide atenção com trinta estudantes ao mesmo tempo.

Bloom terminou o artigo com a pergunta que define o problema: pode alguma combinação de métodos de instrução coletiva reproduzir os resultados da tutoria individual? ². Não era retórica. Era um programa de pesquisa com prazo em aberto.

6.4. EVOLUÇÃO POSTERIOR DA IDEIA

A resposta do campo ao trabalho de Bloom tomou forma em duas linhas que raramente se articularam: a revisão da taxonomia, que chegaria em 2001, e o desenvolvimento de métodos de instrução voltados ao domínio do conteúdo, testados a partir das décadas de 1970 e 1980.

A revisão de Anderson e Krathwohl (2001) ³ foi o desenvolvimento mais consequente da taxonomia original. Krathwohl havia coautorado não apenas o Handbook I, mas o Handbook II, sobre o Domínio Afetivo (1964) ⁹, o que lhe dava uma perspectiva longitudinal sobre as limitações práticas da versão original que poucos pesquisadores possuíam. A influência da psicologia cognitiva dos anos 1980 e 1990, com sua ênfase em

⁹ KRATHWOHL, David R.; BLOOM, Benjamin S.; MASIA, Bertram B. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook II: Affective Domain*. David McKay Company, 1964.

processos mentais em vez de estados estáticos, fundamentou as mudanças de linguagem realizadas em 2001 e tornou possível uma discriminação mais precisa entre objetivos que, escritos em linguagem natural, parecem equivalentes mas impõem condições de instrução completamente distintas.

O movimento de *mastery learning* teve trajetória diferente. Bloom havia formulado os princípios centrais em 1971⁸, derivando-os diretamente do modelo de Carroll. A estrutura era simples: definir com precisão o critério de domínio de cada unidade, fornecer instrução de qualidade, avaliar formativamente com frequência, oferecer feedback diagnóstico e atividades corretivas para quem não atingisse o critério, e só então avançar. O tempo de instrução tornava-se variável; o critério de domínio, constante. Na instrução convencional, a proporção é inversa.

Os resultados foram expressivos o suficiente para atrair atenção, e inconsistentes o suficiente para gerar controvérsia. Robert Slavin conduziu em 1987 uma síntese de evidências que concluía que os efeitos do *mastery learning* eram menores do que os reportados por Bloom quando se aplicavam parâmetros metodológicos mais rigorosos na seleção dos estudos¹⁰. O debate refletia, em parte, a dificuldade real de implementar *mastery learning* com fidelidade em contextos escolares normais, onde pressão de currículo, heterogeneidade de turmas e limitações de tempo tornam a variabilidade de instrução difícil de sustentar.

Guskey (2007) revisou o campo três décadas depois e argumentou que a evidência acumulada, quando analisada com exigência metodológica adequada, sustenta efeitos de tamanho moderado a grande para implementações bem conduzidas, compatíveis com os valores que Bloom havia reportado¹¹. A questão permanece em disputa, o que é sinal de que o campo chegou a uma complexidade empírica que as formulações

¹⁰ SLAVIN, Robert E. *Mastery Learning Reconsidered. Review of Educational Research*, v. 57, n. 2, p. 175–213, 1987.

¹¹ GUSKEY, Thomas R. Closing Achievement Gaps: Revisiting Benjamin S. Bloom's "Learning for Mastery". *Journal of Advanced Academics*, v. 19, n. 1, p. 8–31, 2007.

iniciais não capturavam inteiramente.

A influência da taxonomia sobre o design de certificações profissionais foi mais silenciosa e mais duradoura do que sua influência sobre a pesquisa acadêmica. O CFA Institute adotou os níveis de Bloom explicitamente na definição de seus objetivos de aprendizagem, distribuindo o currículo pelos três primeiros níveis com predominância crescente de Aplicar nas etapas avançadas. A Association of American Medical Colleges reorganizou o MCAT usando a taxonomia revisada como referência. O Conselho de Reguladores de Enfermagem americano estruturou o NCLEX em torno de hierarquias cognitivas derivadas do trabalho de Bloom. O que a taxonomia forneceu, em cada um desses contextos, foi uma linguagem compartilhada para especificar o que significa, concretamente, ter aprendido algo a ponto de ser competente para exercer.

Os verbos de Bloom migraram para o mundo corporativo com rapidez. Os objetivos de aprendizagem de programas de treinamento organizacional, escritos no formato “ao final deste programa, o participante será capaz de..”, são uma tradução direta da linguagem de Tyler e Bloom. Que essa tradução frequentemente resultasse em promessas de Nível 4 entregues com design de Nível 2 não invalida a influência: revela um problema de apropriação parcial que é, em si, uma das questões centrais para qualquer arquitetura de produto educacional responsável.

6.5. IMPACTO REAL NO MUNDO

6.5.1. *Certificações Profissionais como Laboratório da Taxonomia*

O contraste entre essa apropriação superficial e o uso rigoroso da taxonomia em contextos de alta consequência tem uma explicação simples: onde o custo de uma discrepância entre promessa e entrega é mensurá-

vel e imputável, a estrutura de Bloom foi aplicada com fidelidade que o mercado de treinamento corporativo raramente alcançou. O caso mais documentado é o CFA, credencial global para analistas financeiros.

O currículo dos três níveis do exame distribui objetivos de aprendizagem com cuidado explícito pelos seis patamares da taxonomia, com a distribuição se deslocando de Lembrar e Entender no Nível I para ênfase crescente em Analisar e Avaliar no Nível III. A justificativa não é apenas pedagógica: é epistêmica. Um analista financeiro que apenas lembra e entende conceitos de valuation não é funcionalmente diferente de um estudante que acabou de terminar o semestre. O que a credencial atesta é capacidade de operar sob condições de ambiguidade real, com dados incompletos, em situações que nenhum manual previu exatamente dessa forma ¹².

A medicina é ainda mais consequente. O United States Medical Licensing Examination, que define a trajetória de formação médica americana, reorganizou seus objetivos ao longo dos anos 2000 para privilegiar explicitamente os patamares superiores da taxonomia. A questão típica do Step 2 não pergunta “qual é o mecanismo da hipertensão arterial?” (Nível 2). Apresenta um caso clínico com dados que o candidato precisa integrar, analisar e usar para tomar uma decisão diagnóstica ou terapêutica específica (Níveis 3 a 5). Exames baseados em casos dessa natureza não estão testando memória. Estão testando se o candidato opera nos níveis da taxonomia que importam para um médico em exercício real.

O Quadro Europeu de Qualificações, que harmoniza credenciais educacionais nos países da União Europeia, utiliza descritores derivados da taxonomia de Bloom para especificar o que cada patamar de qualificação deve demonstrar ⁶. O mesmo princípio subjaz a frameworks de competência em recursos humanos corporativos, que distinguem entre “conhe-

¹² EUROPEAN COMMISSION. *European Qualifications Framework for Lifelong Learning*. Publications Office of the European Union, 2008.

cer” e “poder demonstrar” como critérios de qualificação para posições de responsabilidade crescente. Em ambos os casos, o que a taxonomia viabiliza é uma linguagem compartilhada para uma distinção que antes era implícita e disputada.

6.5.2. O Problema do 2 Sigma e a Inteligência Artificial

A pergunta que Bloom formulou em 1984 recebeu, nas décadas seguintes, respostas tecnológicas parciais que ele não poderia ter antecipado completamente. Kurt VanLehn, em revisão sistemática publicada em 2011¹³, comparou o tamanho do efeito de três categorias de instrução: tutores humanos individuais, sistemas de tutoria inteligente baseados em computador, e instrução convencional. Os dados eram reveladores. Tutores humanos produziam efeitos médios de 0,79 desvio padrão acima da instrução convencional. Sistemas de tutoria inteligente, 0,76 desvio padrão. A diferença entre tutor humano e sistema computacional era estatisticamente não significativa na maioria dos contextos.

Esses valores não alcançam os 2 sigma de Bloom. Mas são substancialmente superiores à instrução convencional, e equivalentes aos que tutores humanos entregam em tempo real. O que os sistemas de tutoria inteligente fazem, e por que funcionam, é uma questão técnica com implicação direta para o design de produtos educacionais. Eles fornecem feedback imediato, calibrado ao desempenho atual do aprendiz. Eles adaptam a sequência de problemas com base no histórico de respostas. Eles identificam erros sistemáticos e oferecem intervenções específicas, não genéricas. Nenhuma dessas operações é isoladamente impossível em instrução coletiva bem conduzida. O que a tecnologia torna possível é escalá-las para número arbitrário de aprendizes simultaneamente.

O limite da geração de sistemas estudada por VanLehn era o patamar

¹³ VanLEHN, Kurt. The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist*, v. 46, n. 4, p. 197–221, 2011.

taxonômico em que conseguiam operar com eficácia. Feedback sobre resposta de múltipla escolha ou sobre cálculo matemático é tecnicamente tratável. Feedback sobre argumento analítico, sobre design de produto, sobre julgamento avaliativo: esses eram problemas em aberto. A geração de sistemas de IA baseados em modelos de linguagem de grande escala que emergiu a partir de 2020 modificou esse panorama de forma substantiva, tornando possível, pela primeira vez em escala, feedback automatizado sobre texto argumentativo, análise de casos e alguns tipos de produção. Se isso resolve o problema do 2 Sigma ou apenas o aproxima parcialmente, os estudos controlados com a rigorosidade que Bloom exigiria estão apenas começando.

Há um risco simétrico e oposto ao otimismo tecnológico: o risco de que plataformas de IA educacional operem nos Níveis 1 e 2 da taxonomia com eficiência impressionante, reforçando exatamente o viés que Bloom havia diagnosticado como o problema central da educação de massa. Recuperação de informação automatizada, compreensão verificada por questionários gerados por IA, tudo isso é tecnicamente tratável e pedagogicamente insuficiente. A questão não é se a tecnologia pode fazer isso: é se o design dos produtos a utilizará para subir a taxonomia ou para escalar o que já era fácil de escalar.

6.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

Responder à pergunta que a seção anterior deixa em aberto, sobre como usar a tecnologia para subir a taxonomia em vez de escalar o que já era simples de escalar, exige entender como a estrutura de Bloom articula com os demais frameworks deste livro. Não porque todos os outros autores derivem de Bloom, mas porque a taxonomia fornece a linguagem

de especificação que os outros frameworks precisam para ser operacionalizados com precisão. É a articulação que define o que uma arquitetura de produto pode, honestamente, prometer.

6.6.1. Bloom e Wiggins & McTighe: o Mapa do Território

O backward design de Wiggins e McTighe começa com uma pergunta: o que o estudante deve ser capaz de fazer, entender e conhecer ao final da experiência de aprendizagem? A resposta a essa pergunta, quando bem formulada, é inevitavelmente uma afirmação sobre um nível específico da taxonomia de Bloom¹⁴. “Entender o conceito de força” é uma declaração de intenção que pode significar coisas muito diferentes dependendo do nível que se tem em mente: compreender a definição (Nível 2), aplicar a segunda lei de Newton em problemas específicos (Nível 3), analisar sistemas mecânicos compostos com variáveis múltiplas (Nível 4). O UbD sem a taxonomia produz declarações de intenção que parecem específicas mas não são. A taxonomia sem o UbD produz mapeamentos cognitivos sem direção de design.

A integração mais produtiva está no conceito de “understandings” de Wiggins e McTighe. Quando os autores definem que compreensão genuína exige que o estudante seja capaz de explicar, interpretar, aplicar, mudar perspectiva, demonstrar empatia e exercer autoconhecimento⁹, eles estão descrevendo, em termos funcionais, operações que cobrem os Níveis 2 a 6 da taxonomia. A taxonomia fornece a estrutura hierárquica. O UbD fornece a orientação de propósito.

¹⁴ WIGGINS, Grant; MCTIGHE, Jay. *Understanding by Design*. 2. ed. ASCD, 2005.

6.6.2. Bloom e Glasser: Subir a Hierarquia Exige Mudar as Modalidades

O capítulo anterior estabeleceu que diferentes formas de aprendizagem produzem diferentes níveis de fixação, com as ativas consistentemente superiores às passivas¹⁵. A relação com Bloom é de complementaridade estrutural: não se pode subir na taxonomia usando modalidades inadequadas ao patamar visado.

Exposição passiva é suficiente, em condições razoáveis, para desenvolver Lembrar (Nível 1) e parte do Entender (Nível 2). Não é suficiente para Aplicar. Um estudante pode assistir a quarenta horas de videoaulas sobre análise de balanços e não conseguir analisar o balanço de uma empresa real. Não porque não prestou atenção: porque Aplicar requer prática em contexto de aplicação, com feedback sobre o desempenho na tarefa. Glasser determina a profundidade de processamento cognitivo que cada modalidade permite atingir; Bloom especifica o grau de competência que essa profundidade pode desenvolver. Os dois argumentos são necessários em conjunto. Separadamente, cada um responde apenas metade da pergunta.

6.6.3. Bloom e Gagné: Cada Nível Requer Eventos Específicos

O modelo de Gagné dos Nove Eventos de Instrução¹⁶ especifica as condições instrucionais que maximizam a probabilidade de aprendizagem em cada episódio. A integração com Bloom revela que o peso relativo de cada evento varia substancialmente dependendo do nível taxonômico visado.

O Evento 7, provocar o desempenho (*eliciting performance*), tem natureza diferente conforme o patamar taxonômico. Para Lembrar, elicitar

¹⁵ GLASSER, William. *The Quality School: Managing Students Without Coercion*. Harper & Row, 1990.

¹⁶ GAGNÉ, Robert M. *The Conditions of Learning and Theory of Instruction*. 4. ed. Holt, Rinehart and Winston, 1985.

desempenho pode ser um quiz de recuperação. Para Aplicar, exige uma tarefa genuína com dados novos em contexto que o estudante não viu antes. Para Avaliar ou Criar, o produto elicitado tem que ser avaliável contra critérios explícitos. Gagné fornece o esqueleto do episódio instrucional. Bloom especifica o tipo de operação cognitiva que cada evento precisa mobilizar.

O Evento 8, fornecer feedback (*providing feedback*), muda de natureza de forma ainda mais pronunciada. Feedback sobre recuperação de memória pode ser binário: correto ou incorreto. Feedback sobre análise ou avaliação precisa ser qualitativo, precisar onde o raciocínio divergiu do que a tarefa exigia e por que esse desvio importa. Esse é precisamente o tipo de feedback que a tutoria individual fornece e que a instrução coletiva convencional não consegue escalar. Gagné o nomeia como evento necessário. Bloom quantificou o que sua ausência custa em termos de aprendizagem: um desvio padrão, no mínimo, entre o que o estudante poderia atingir e o que efetivamente atinge.

7

A ENGENHARIA DA EXCELÊNCIA: ANDERS ERICSSON E A PRÁTICA DELIBERADA

“O que distingue as pessoas que alcançam desempenho de nível mundial não é simplesmente a quantidade de prática que realizaram, mas a qualidade específica dessa prática.” K. Anders Ericsson, Peak, 2016¹

7.1. CONTEXTO HISTÓRICO

Em algum momento dos anos 1960 e 1970, a psicologia experimental tinha um problema que raramente nomeava com clareza. O campo estudava aprendizagem com sofisticação crescente, produzia modelos sobre memória, atenção e processamento de informação, mas permanecia notavelmente silencioso sobre uma questão que qualquer observador atento ao desempenho humano real poderia formular: por que algumas pessoas

¹ ERICSSON, K. Anders; POOL, Robert. *Peak: Secrets from the New Science of Expertise*. Houghton Mifflin Harcourt, 2016.

continuam melhorando indefinidamente enquanto outras param? A psicologia comportamental havia se concentrado nas condições de reforço. A psicologia cognitiva nascente se debruçava sobre os mecanismos de representação mental. Nenhuma das duas havia tratado o desempenho como o problema central que era: não como estado, mas como trajetória.

A exceção parcial vinha da pesquisa em xadrez. Adriaan de Groot, psicólogo holandês, havia publicado em 1946 um estudo sobre como mestres de xadrez percebem e analisam posições no tabuleiro, obra que só alcançaria audiência anglófona na edição de 1965 ². O que de Groot demonstrou era contra-intuitivo: grandes mestres não calculavam mais movimentos do que jogadores de nível médio. Percebiam mais. Em frações de segundo, reconheciam padrões estruturais que jogadores menos experientes levavam minutos para identificar, quando conseguiam. A expertise, nesse caso, não era uma questão de velocidade de processamento bruto: era uma questão de organização do conhecimento.

Herbert Simon e William Chase formalizaram esse achado em 1973, com o conceito de “chunks”, unidades de informação organizadas em estruturas reconhecíveis que peritos armazenam na memória de longo prazo ³. Um grande mestre de xadrez com experiência de décadas havia acumulado, segundo as estimativas de Simon e Chase, entre 50.000 e 100.000 padrões de tabuleiro que podia reconhecer diretamente, sem análise sequencial. A memória de curto prazo, limitada a cerca de sete unidades de informação simultâneas para qualquer ser humano, deixava de ser o gargalo quando cada “unidade” poderia ser um padrão complexo de quinze peças com implicações estratégicas codificadas. A expertise, nessa leitura, era fundamentalmente um problema de estrutura de representação mental.

Karl Anders Ericsson chegou à Suécia em 1947 e estudou psicologia

² DE GROOT, Adriaan D. *Thought and Choice in Chess*. Mouton, 1965.

³ SIMON, Herbert A.; CHASE, William G. Skill in Chess. *American Scientist*, v. 61, n. 4, p. 394–403, 1973.

em Estocolmo nos anos 1970, num período em que essa tradição de pesquisa sobre expertise estava consolidando seu vocabulário conceitual mas ainda carecia de uma teoria unificada sobre como tais representações se formavam. O que o atraía não era a memória de peritos como fenômeno estático: era o processo pelo qual ela se tornava o que era. Como se passa de novato a especialista? Que atividades específicas produzem a acumulação de representações que de Groot e Simon haviam descrito? A pergunta parecia óbvia, retrospectivamente. Ninguém havia tentado respondê-la com rigor empírico.

A trajetória de Ericsson levou-o para os Estados Unidos, onde faria sua carreira, primeiro em Boulder, no Colorado, depois por décadas em Tallahassee, na Florida State University. O contato direto com Simon foi formativo. Mas a divergência também era real: Simon havia documentado o produto da expertise. Ericsson queria estudar sua produção. A distinção parece tênue. Suas implicações não são.

O estudo que alterou a trajetória do campo foi publicado em 1993 no *Psychological Review*, com Ralf Krampe e Clemens Tesch-Römer como coautores⁴. O objeto era a Academia de Música Hanns Eisler em Berlim, mais especificamente seus alunos de violino, divididos em três grupos por nível projetado de carreira: os que tinham potencial para carreiras solo internacionais, os destinados a orquestras de primeira linha, e os que seguiriam para o ensino. Todos haviam começado a estudar violino por volta dos cinco anos. Todos tinham passado pelos mesmos professores nas fases iniciais. A pergunta era direta: o que diferenciava, em termos de comportamento e prática acumulada, os que estavam na trajetória de elite?

A resposta não era talento inato. Era prática. Um tipo específico de prática. Aos vinte anos, os violinistas do grupo superior haviam acumu-

⁴ ERICSSON, K. Anders; KRAMPE, Ralf T.; TESCH-RÖMER, Clemens. The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review*, v. 100, n. 3, p. 363–406, 1993.

lado em média dez mil horas do que Ericsson chamaria de prática deliberada. Os do grupo intermediário, sete mil e quinhentas. Os do grupo de ensino, cinco mil. A lacuna não era de aptidão: era de atividade estruturada e acumulada ao longo de anos, iniciada nos mesmos instrumentos pelas mesmas famílias e guiada, nas fases decisivas, por professores capazes de exigir mais do que a zona de conforto do estudante permitia.

Esse dado, tirado de seu contexto, alimentaria uma das maiores deformações conceituais da literatura popular sobre desempenho. Em seu contexto, era a base empírica de uma teoria sobre como a excelência se produz que o próprio Ericsson passaria as décadas seguintes refinando, defendendo e, em aspectos críticos, corrigindo.

7.2. O PROBLEMA QUE ERICSSON ESTAVA RESOLVENDO

O mito do talento inato é mais resistente do que parece porque resolve um problema real: a assimetria de desempenho entre pessoas submetidas às mesmas condições aparentes de instrução. Se dois estudantes passam pelo mesmo professor pelo mesmo número de anos e chegam a resultados radicalmente distintos, a explicação mais econômica, no sentido cognitivo do termo, é que eram diferentes desde o início. Talentos distintos. Aptidões desiguais. Inteligências de outro calibre. A explicação é conveniente porque exime todos os atores da análise mais desconfortável: a instrução foi diferente, a prática foi diferente, as condições de feedback divergiam, as horas de estudo efetivo não eram as mesmas.

Há evidência robusta de que predisposições genéticas específicas facilitam determinadas aquisições. Ninguém sério nega isso. O problema é a escala em que o talento inato é invocado como explicação, muitas vezes para domínios onde as evidências indicam exatamente o oposto. A al-

tura em basquete é um preditor legítimo de certos tipos de performance. O senso tonal absoluto tem componente genético identificável. Mas a proporção de desempenho excepcional que pode ser atribuída a predisposições inatas, separada da prática acumulada nas condições certas, é substancialmente menor do que a narrativa cultural dominante sobre o gênio espontâneo sugere ².

O segundo problema que Ericsson identificou é mais sofisticado e, para a teoria de produtos educacionais, mais relevante. Chama-se, na sua formulação, a “ilusão da experiência”. A suposição de que mais anos de prática implicam maior competência não é uma lei: é uma generalização falsa. Médicos com vinte anos de carreira clínica cometem, em certas tarefas diagnósticas específicas, erros que residentes de segundo ano cometem com menor frequência. Não porque sejam menos inteligentes: porque a prática não estruturada, repetida indefinidamente, consolida os procedimentos que já existem sem necessariamente melhorar os que são deficientes.

Paul Fitts e Michael Posner haviam descrito esse fenômeno estruturalmente em 1967, com o modelo de três estágios de aquisição de habilidades: cognitivo, associativo e autônomo ⁵. No estágio autônomo, a habilidade está tão automatizada que pode ser executada sem atenção consciente. O problema é que a automatização também bloqueia o melhoramento. Uma vez que uma habilidade atingiu esse patamar de automatização, ela é executada com menos variância e menos melhoria ao longo do tempo. O perito que deixa de impor demandas de melhoria a si mesmo não está praticando: está repetindo. A repetição não é prática deliberada.

A pergunta central que Ericsson formulou pode ser expressa com precisão: o que, na atividade de prática, determina se uma habilidade continuará a melhorar ou entrará em platô? A resposta não está nas horas

⁵ FITTS, Paul M.; POSNER, Michael I. *Human Performance*. Brooks/Cole, 1967.

totais. Está na estrutura da atividade dentro dessas horas. A distinção importa porque elimina a conclusão falsa de que experientes são necessariamente competentes, e de que novatos com menos horas são necessariamente menos capazes em tarefas específicas onde a prática do novato foi mais deliberada do que a do veterano. Nomear esse mecanismo com precisão era, por isso, a tarefa que restava.

7.3. IDEIA CENTRAL

7.3.1. *Prática Deliberada: Estrutura e Componentes*

A prática deliberada, no sentido técnico que Ericsson conferiu ao termo, não é qualquer atividade de prática. É uma forma específica de atividade que satisfaz condições estruturais precisas, e que é qualitativamente diferente tanto da prática ingênua quanto da prática intencional comum¹.

A prática ingênua é o que a maioria das pessoas faz: repetir a atividade com a esperança de que a competência venha com o tempo. Tocar as mesmas músicas no violino que já se sabe tocar. Realizar as cirurgias já conhecidas. Conduzir reuniões com o formato habitual. A atividade produz conforto e velocidade de execução, mas raramente melhora além de um platô que se atinge relativamente cedo.

A prática intencional é um grau acima: há esforço consciente para melhorar, mas sem estrutura clara de como esse esforço deve ser aplicado. O profissional que decide “praticar mais” sem especificar o quê, como e com qual critério de avaliação está engajado em prática intencional, não em prática deliberada.

A prática deliberada satisfaz, segundo o modelo de Ericsson², quatro condições interdependentes. A primeira é que ela opera fora da zona de

conforto: exige esforço cognitivo real e desconforto, porque se uma tarefa pode ser executada facilmente, ela já está automatizada e não produz adaptação adicional. A segunda é que ela é projetada com objetivo específico e bem definido, não como “melhorar em geral” mas como “melhorar o controle de arco nos compassos 34 a 41 da sonata”. A especificidade é funcional, não arbitrária: permite que o feedback seja informativo.

A terceira condição é o feedback imediato e específico, seja fornecido por um professor, por um instrumento de medição, ou pela própria estrutura da tarefa. Sem feedback, não há como saber se a adaptação ocorreu. Sem feedback específico, não há como saber qual aspecto da performance mudou. A quarta condição é que a atividade requer atenção plena. Não é possível engajar na prática deliberada “no piloto automático”. A atenção consciente mantida sobre a execução é o mecanismo pelo qual o sistema cognitivo registra discrepâncias entre o desempenho atual e o desempenho alvo.

Essas quatro condições em conjunto produzem algo que a prática ingênua não produz: representações mentais progressivamente mais sofisticadas do domínio.

7.3.2. Representações Mentais: O Mecanismo Central

O conceito de representação mental é, na teoria de Ericsson, mais do que terminologia técnica: é o mecanismo pelo qual a prática deliberada produz seu efeito². Peritos e novatos diferem não principalmente na velocidade de processamento ou na capacidade de memória de curto prazo: diferem na qualidade e na riqueza das estruturas mentais através das quais percebem, organizam e avaliam informação no domínio de expertise.

Um médico especialista que examina um raio-x de tórax não “vê mais” no sentido de processar mais pixels. Percebe de forma qualitativamente

diferente porque sua representação mental do campo inclui padrões, anomalias e relações que o residente recém-formado ainda não codificou como unidades reconhecíveis ⁶. O mesmo ocorre com o grande mestre de xadrez diante do tabuleiro, com o designer de produto ao examinar uma interface, com o analista financeiro ao ler um balanço. A expertise não é velocidade: é estrutura de percepção.

A prática deliberada produz representações mentais cada vez mais detalhadas, mais hierarquicamente organizadas e mais funcionalmente conectadas ao desempenho. O feedback específico durante a prática é o mecanismo pelo qual o praticante detecta a discrepância entre a representação atual (o que ele “vê” na situação) e a representação mais sofisticada do perito (o que deveria ser visto). Cada ciclo bem estruturado produz uma atualização incremental da representação.

Essa concepção do que a prática deliberada produz é radicalmente diferente da noção popular de que praticar é “fazer mais vezes até que fique fácil”. Quando uma tarefa fica fácil, a prática deliberada precisa ser redirecionada para o que ainda é difícil. O conforto é sinal de automatização. A automatização é inimiga da melhoria adicional. A teoria de Ericsson é, nesse sentido, uma teoria sobre a gestão contínua do desconforto produtivo. Esse caráter estruturalmente exigente seria tanto a fonte de sua força explicativa quanto a causa das distorções que sofreu ao entrar na literatura popular.

7.4. EVOLUÇÃO POSTERIOR DA IDEIA

⁶ ERICSSON, K. Anders. Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains. *Academic Medicine*, v. 79, n. 10, p. S70–S81, 2004.

7.4.1. O Problema das 10.000 Horas

O estudo de 1993 forneceu um número que Malcolm Gladwell imortalizaria em *Outliers*, publicado em 2008⁷, com consequências que Ericsson passou anos subsequentes tentando corrigir. A “regra das 10.000 horas”, como Gladwell a formulou, sugeria que qualquer pessoa podia alcançar maestria em qualquer campo investindo esse tempo mínimo de prática. A afirmação era generosa, otimista e substancialmente incorreta.

Ericsson havia encontrado que os violinistas mais promissores da Academia de Berlim haviam acumulado cerca de dez mil horas de prática deliberada até os vinte anos. Esse número não era uma regra universal: era uma média de um grupo específico em um domínio específico. Em outros domínios, o tempo necessário varia enormemente. Em xadrez, grandes mestres tipicamente precisam de onze mil horas antes de atingir o nível de grande mestre; jogadores de tênis de alto nível, de menos. A variação reflete diferenças na estrutura do domínio, na disponibilidade de professores qualificados, e na intensidade possível de prática deliberada sustentada⁸.

O problema mais sério com a leitura de Gladwell não era o número. Era o adjetivo que havia sido omitido. As horas que Ericsson havia estudado não eram horas de atividade qualquer no domínio: eram horas de atividade estruturada, fora da zona de conforto, com feedback específico e atenção plena. Horas de apresentação em palcos não contavam. Horas de improviso informal não contavam. A distinção não é pedantismo: é a diferença entre uma teoria sobre como a excelência se produz e uma afirmação factualmente vaga sobre quanto tempo “praticar” requer.

A meta-análise de MacNamara, Hambrick e Oswald, publicada em 2014 em *Psychological Science*⁹, encontrou que a prática deliberada ex-

⁷ GLADWELL, Malcolm. *Outliers: The Story of Success*. Little, Brown and Company, 2008.

⁸ ERICSSON, K. Anders; LEHMANN, Andreas C. Expert and Exceptional Performance: Evidence of Maximal Adaptation to Task Constraints. *Annual Review of Psychology*, v. 47, p. 273–305, 1996.

⁹ MACNAMARA, Brooke N.; HAMBRICK, David Z.; OSWALD, Frederick L. Deliberate Practice and Performance in

plica em média 26% da variância em desempenho em jogos, 21% em música, 18% em esportes e 4% em educação. Ericsson contestou parte dos resultados metodológicos, argumentando que muitos estudos incluídos na meta-análise não mediam prática deliberada no sentido técnico do termo, mas sim horas de atividade no domínio em geral ². O debate reflete uma tensão real: a dificuldade de operacionalizar o conceito de forma suficientemente precisa para estudos em larga escala.

O que permanece não contestado é o resultado do estudo original. Violinistas de elite diferem de violinistas de nível médio não em horas totais de contato com o instrumento, mas em horas de prática estruturada especificamente para melhorar aspectos identificados como deficientes, com feedback de professores qualificados que podiam nomear com precisão o que precisava mudar.

7.4.2. Expansões: Neurociência, Perseverança e Aprendizado de Máquina

A controvérsia sobre a popularização não impediu que pesquisadores de outras tradições encontrassem, independentemente, convergências com o núcleo do argumento de Ericsson. Daniel Coyle publicou *The Talent Code* em 2009 ¹⁰ com o argumento de que a neurociência havia encontrado o substrato biológico da prática deliberada. A mielinização, processo pelo qual células de suporte neuronal envolvem axônios em bainhas de mielina que aceleram a transmissão de sinais, havia sido documentada como processo sensível à ativação repetida de circuitos específicos. R. Douglas Fields, neurocientista do Instituto Nacional de Saúde americano, havia mostrado que células oligodendrócitos respondem à atividade elétrica nos axônios produzindo mielina adicional, com efeitos

Music, Games, Sports, Education, and Professions: A Meta-Analysis. *Psychological Science*, v. 25, n. 8, p. 1608–1618, 2014.

¹⁰ COYLE, Daniel. *The Talent Code: Greatness Isn't Born. It's Grown. Here's How*. Bantam Books, 2009.

mensuráveis na velocidade de condução neural ¹¹.

A leitura de Coyle era que a prática deliberada é o mecanismo pelo qual se ativa a mielinização nos circuitos relevantes para a habilidade sendo desenvolvida. A hipótese era sedutora: oferecia uma explicação biológica para por que a prática deliberada produz efeitos que a prática ingênua não produz. A relação entre mielinização e comportamento de alto desempenho é mais complexa do que a narrativa popularizada sugeria, e o processo não se presta facilmente à instrumentação pedagógica direta. Mas a convergência entre a perspectiva comportamental de Ericsson e a perspectiva neurológica de Fields reforçou a legitimidade científica do conceito como fenômeno real com correlatos biológicos identificáveis.

Angela Duckworth tomou trajetória diferente. Seu conceito de “grit”, a combinação de paixão persistente e perseverança de longo prazo, foi apresentado inicialmente em artigos com Ericsson como coautor ¹² antes de se tornar o argumento central de Grit, publicado em 2016 ¹³. A relação entre grit e prática deliberada é de pré-condição: a prática deliberada é, por definição, desconfortável e exige esforço sustentado ao longo de períodos longos. Sem perseverança, não se sustenta o tempo necessário para produzir representações mentais de nível perito. Duckworth identificava o componente motivacional que Ericsson havia deixado em segundo plano.

O estudo publicado em *Social Psychological and Personality Science* em 2011 ¹⁹, conduzido por Duckworth e Ericsson com outros pesquisadores, examinou participantes do National Spelling Bee americano. O resultado era preciso: a quantidade de prática deliberada acumulada pre-

¹¹ FIELDS, R. Douglas. White Matter in Learning, Cognition and Psychiatric Disorders. *Trends in Neurosciences*, v. 31, n. 7, p. 361–370, 2008.

¹² DUCKWORTH, Angela L.; KIRBY, Teri A.; TSUKAYAMA, Eli; BERSTEIN, Heather; ERICSSON, K. Anders. Deliberate Practice Spells Success: Why Grittier Competitors Triumph at the National Spelling Bee. *Social Psychological and Personality Science*, v. 2, n. 2, p. 174–181, 2011.

¹³ DUCKWORTH, Angela. *Grit: The Power of Passion and Perseverance*. Scribner, 2016.

viamente previa o desempenho melhor do que o grau de grit. Mas o grau de grit previa qual dos participantes acumularia mais prática deliberada. Os dois construtos não são intercambiáveis: são sequenciais. O grit determina a quantidade de prática deliberada que o estudante realizará. A prática deliberada determina o nível de desempenho alcançado.

Mais recentemente, a estrutura de ciclos de feedback com objetivo específico e ajuste incremental encontrou paralelo inesperado em sistemas de aprendizado de máquina supervisionado. Tais modelos são expostos a dados com rótulos corretos (equivalente ao feedback), ajustam parâmetros internos (equivalente à adaptação da representação mental) e são testados em dados novos para verificar generalização. A analogia tem limites óbvios: sistemas de aprendizado de máquina não têm “desconforto” e não precisam de atenção consciente. Mas o paralelo funcional reforça que a lógica de ciclo deliberado, feedback específico e adaptação incremental não é peculiaridade do desempenho humano: é uma estrutura de aprendizagem com alcance mais amplo.

7.5. IMPACTO REAL NO MUNDO

A estrutura identificada por Ericsson não esperou pela publicação do artigo de 1993 para aparecer em domínios de alto risco. O simulador de voo como dispositivo de treinamento é, retrospectivamente, um dos exemplos mais bem-sucedidos de aplicação implícita dos princípios de prática deliberada: permite que pilotos pratiquem procedimentos de emergência com feedback imediato sobre cada ação, em condições que replicam a experiência real com custo e risco eliminados ². O debriefing pós-simulação, quando conduzido com especificidade suficiente para identificar exatamente o que o piloto fez diferente do procedimento correto e

quando, satisfaz as condições de feedback que Ericsson havia descrito como necessárias. O que a aviação militar havia descoberto empiricamente, a psicologia da expertise descreveria conceitualmente duas décadas depois.

A medicina revisou seus modelos de residência parcialmente sob a influência direta dessa literatura. Ericsson publicou em *Academic Medicine* em 2004 uma análise das implicações da prática deliberada para o treinamento médico ¹¹, argumentando que a exposição clínica bruta, o modelo tradicional de “ver e fazer”, não satisfaz as condições que produzem competência demonstrável. Um residente que realiza cento e cinquenta procedimentos em que o feedback se limita ao resultado (paciente evoluiu bem ou mal) acumula exposição, não prática deliberada.

O feedback necessário seria: qual aspecto específico da técnica na cirurgia de número noventa e três diferiu do que um especialista de referência teria feito, e por quê esse desvio importa? A pergunta parece simples. Na prática clínica tradicional, raramente é formulada em termos operacionais suficientemente precisos para gerar adaptação.

No esporte de elite, a influência é mais direta. O modelo de periodização do treinamento físico, que divide o ciclo anual em fases com objetivos distintos e feedback específico sobre adaptações fisiológicas mensuráveis, converge com os princípios de Ericsson, embora derivado de tradições independentes. O que Ericsson acrescentou foi a extensão do raciocínio para além das habilidades físicas: a representação mental do atleta, sua capacidade de perceber e antecipar padrões de jogo, é treinável com a mesma lógica que se aplica às habilidades motoras.

O fracasso das “escolas de talento” que prometem identificar e desenvolver talentos excepcionais sem estruturar a prática com rigor deliberado é documentável e segue um mecanismo preciso. O modelo de muitas academias esportivas jovens seleciona crianças com caracterís-

ticas físicas favoráveis e as coloca em treinamento de volume elevado sem feedback específico sobre aspectos técnicos identificados como deficientes. O resultado são atletas que atingem platôs precocemente. A seleção ativa uma profecia autocumprida: crianças com características iniciais favoráveis são identificadas como talentos, recebem mais atenção de treinadores, acumulam prática de maior qualidade, e confirmam a previsão inicial. Não porque eram mais talentosas desde o início, mas porque receberam mais prática deliberada desde que foram selecionadas¹⁷.

O mesmo padrão reproduz-se em contextos corporativos. Programas de “desenvolvimento de lideranças” que consistem em exposição a conteúdo e workshops de reflexão, sem ciclos estruturados de aplicação com feedback sobre o desempenho real nas situações-alvo, produzem o equivalente gerencial do atleta em platô: profissionais que reconhecem os conceitos, articulam os princípios, e retornam ao mesmo repertório comportamental de antes. A estrutura importa mais do que o conteúdo. Ericsson havia documentado isso em domínios artísticos e esportivos. O princípio se generaliza. E é precisamente essa generalização que conecta o estudo de violinistas berlinenses às questões de design que este livro examina.

7.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

7.6.1. *Ericsson e Polanyi: o Mecanismo de Formação do Tácito*

O capítulo sobre Michael Polanyi estabeleceu que o conhecimento tácito, aquilo que sabemos sem conseguir dizer como sabemos, é a camada mais profunda e mais valiosa da expertise operacional. O violinista que sabe “como o arco deve soar” antes de conseguir verbalizar os critérios

técnicos que distinguem o som certo do errado é um caso entre muitos. O médico que reconhece a anormalidade antes de completar o raciocínio diagnóstico, o designer que identifica o problema de interface antes de articular os princípios de usabilidade que o justificam: todos operam com conhecimento tácito no sentido de Polanyi¹⁴.

O que Ericsson acrescenta a Polanyi é o mecanismo. O filósofo descreveu a estrutura do conhecimento tácito e argumentou que ele é irreduzível à verbalização completa. Ericsson descreveu como esse conhecimento se forma. A prática deliberada é o processo pelo qual as representações mentais do perito, inicialmente explícitas e conscientes, se tornam progressivamente automáticas e sub-simbólicas. O que começa como regra verbal (“manter o cotovelo a noventa graus durante o movimento”) se transforma, após milhares de repetições com feedback específico, em ajuste proprioceptivo que ocorre sem instrução consciente. O conhecimento deixou de ser proposicional. Tornou-se tácito no sentido preciso de Polanyi.

Dois corolários seguem para o design de produtos. O primeiro: o conhecimento tácito de um perito não pode ser transferido diretamente por exposição, porque não existe em forma que possa ser exposta. Pode apenas ser reconstituído no aprendiz através de ciclos de prática deliberada suficientemente bem estruturados para acionar o mesmo processo de consolidação que ocorreu no especialista. O segundo: o que o especialista enuncia durante o ensino não é o conhecimento relevante, é uma aproximação proposicional de algo que resiste à proposicionalização completa. O aprendiz precisa dessa enunciação como orientação, mas a aprendizagem real ocorre na prática que segue, não na exposição à verbalização.

¹⁴ POLANYI, Michael. *The Tacit Dimension*. Doubleday, 1966.

7.6.2. *Ericsson e Bloom: a Prática Deliberada como Motor Taxonômico*

O capítulo sobre Bloom estabeleceu que a taxonomia descreve níveis qualitativamente distintos de operação cognitiva, e que o Nível 3 (Aplicar) constitui o patamar mínimo de competência real em qualquer domínio¹⁵. A relação com Ericsson é de especificação mecanística: a prática deliberada é o processo pelo qual o aprendiz sobe os níveis da taxonomia.

É impossível subir de Entender para Aplicar sem prática deliberada. Compreender o conceito de taxa de desconto (Nível 2) não produz, por si, a capacidade de aplicar o cálculo de valor presente em situações de ambiguidade real (Nível 3). A transição requer ciclos de aplicação em contextos variados, com feedback sobre cada tentativa, suficientemente numerosos para construir representações do que é “aplicação correta” em condições distintas. Isso é precisamente o que a prática deliberada produz.

A subida de Aplicar para Analisar (Nível 4) requer ainda mais: o aprendiz precisa não apenas executar procedimentos corretamente, mas perceber a estrutura relacional dos problemas, identificar o que varia e o que permanece constante entre situações superficialmente similares. Essa forma de percepção não se forma por exposição adicional ao conteúdo. Forma-se pela prática de análise, com feedback sobre a qualidade do raciocínio analítico, não apenas sobre o resultado final.

Ericsson e Bloom, lidos juntos, produzem uma proposição com consequências diretas para design: cada transição entre níveis taxonômicos requer um conjunto de ciclos de prática deliberada projetados especificamente para as representações que esse nível exige. Não há atalho proposicional. Não há número de videoaulas que substitua.

¹⁵ BLOOM, Benjamin S. et al. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. David McKay Company, 1956.

7.6.3. *Ericsson e Gagné: Prática Deliberada nos Eventos de Instrução*

Robert Gagné descreveu os Nove Eventos de Instrução como condições que, quando presentes em um episódio instrucional bem estruturado, maximizam a probabilidade de aprendizagem¹⁶. Dois deles são particularmente relevantes aqui: o Evento 7, provocar o desempenho, e o Evento 8, fornecer feedback.

No modelo de Gagné, provocar o desempenho é um evento instrucional entre outros. No modelo de Ericsson, o desempenho deliberado estruturado é a atividade central, em torno da qual todo o restante do design instrucional existe como preparação ou suporte. A diferença de ênfase reflete diferenças de contexto: Gagné descrevia episódios de instrução formal; Ericsson estudava trajetórias de anos de desenvolvimento de expertise. Mas a convergência é real: os dois concordam que provocar o desempenho e fornecer feedback sobre ele são as operações que produzem aprendizagem real.

A contribuição de Ericsson ao modelo de Gagné é a especificação de que o desempenho precisa ser deliberado, não apenas elicitado. Não basta “elicitado desempenho” como tarefa qualquer: a tarefa deve estar fora da zona de conforto atual do aprendiz, focada em um aspecto específico que ainda não está automatizado, e estruturada para que o feedback possa ser informativo sobre exatamente o que deve mudar. Quando o Evento 7 de Gagné é preenchido com prática deliberada no sentido de Ericsson, o resultado instrucional é qualitativamente diferente do que quando é preenchido com repetição confortável de habilidades já dominadas.

¹⁶ GAGNÉ, Robert M. *The Conditions of Learning and Theory of Instruction*. 4. ed. Holt, Rinehart and Winston, 1985.

8

A ARQUITETURA DO ENSINAR: ROBERT GAGNÉ E O DESIGN INSTRUCIONAL SISTEMÁTICO

“A aprendizagem não ocorre de uma só vez. Ocorre em etapas, e cada etapa cria as condições para a seguinte.” Robert M. Gagné, The Conditions of Learning, 1965¹

8.1. CONTEXTO HISTÓRICO

Há algo paradoxal no fato de que a disciplina hoje chamada de design instrucional nasceu não de uma universidade, mas de uma guerra. A Segunda Guerra Mundial colocou os Estados Unidos diante de um problema de engenharia humana de escala sem precedente: era necessário treinar, em meses, milhões de recrutas sem formação militar anterior

¹ GAGNÉ, Robert M. *The Conditions of Learning and Theory of Instruction*. 4. ed. Holt, Rinehart and Winston, 1985.

para operar equipamentos complexos, tomar decisões sob pressão e executar procedimentos que, mal aprendidos, resultariam em baixas. A pedagogia disponível em 1941 não tinha resposta para esse problema. A educação formal pressupunha anos, professores individualizados, progresso gradual. O conflito não permitia nenhum dos três.

O que emergiu desse impasse foi menos uma solução pedagógica do que uma resposta de engenharia. Psicólogos, especialistas em comportamento e analistas de sistemas foram recrutados para desenvolver materiais de treinamento que pudessem funcionar independentemente de quem os aplicasse. O foco deslocou-se do professor para o material. Da habilidade individual do instrutor para a estrutura da instrução em si. Essa inversão, que pareceu puramente pragmática na época, seria o fundamento epistemológico sobre o qual Robert Gagné construiria sua obra nas décadas seguintes.

Robert Mills Gagné nasceu em 1916 e formou-se em psicologia em Yale em 1940, um ano antes da entrada americana na guerra. Trabalhou durante o conflito analisando o desempenho de pilotos e desenvolvendo procedimentos de treinamento para a Air Force americana ¹. A experiência foi decisiva, não porque confirmou o que ele já sabia, mas porque o forçou a perguntar que a psicologia acadêmica da época não havia feito: por que certos procedimentos de treinamento funcionavam independentemente do instrutor, enquanto outros dependiam inteiramente da competência pessoal do professor? O que havia na estrutura do procedimento bem-sucedido que o tornava replicável?

A psicologia behaviourista dominante naquele momento tinha uma resposta, mas ela era insatisfatória: estímulo, resposta, reforço. O modelo de Skinner descrevia como comportamentos eram adquiridos, mas não como habilidades complexas eram desenvolvidas em sequência, cada uma dependendo das anteriores. A instrução de pilotagem não podia ser

reduzida a cadeias de reforço. Havia uma arquitetura subjacente que o behaviorismo não capturava.

A pesquisa que Gagné iniciou na Air Force Personnel and Training Research Center no início dos anos 1950 iria explorar exatamente essa arquitetura. O trabalho culminaria na primeira edição de *The Conditions of Learning*, publicado em 1965, que propunha uma teoria unificada de como diferentes tipos de aprendizagem ocorrem e quais condições precisam estar presentes para que cada um deles se realize ¹. O livro chegaria à quarta edição em 1985, revisado e expandido, mas com a estrutura central intacta.

O contexto intelectual dos anos 1960 tornava o projeto de Gagné simultaneamente necessário e controverso. A revolução cognitiva estava deslocando o behaviorismo como paradigma dominante, e pesquisadores como Jerome Bruner e David Ausubel desenvolviam modelos de aprendizagem que levavam a sério os processos internos de representação e organização do conhecimento. Gagné não era alheio a essa tradição: sua taxonomia de resultados de aprendizagem incorpora distinções que pressupõem processamento interno sofisticado. Mas ele manteve um compromisso com a operacionalização que muitos cognitivistas consideravam redutor. Para Gagné, uma teoria de aprendizagem que não se traduzia em prescrições concretas de design instrucional era de utilidade limitada ².

A Air Force continuou financiando pesquisa aplicada em treinamento ao longo dos anos 1950 e 1960. O que estava em jogo não era apenas eficiência pedagógica: era a capacidade do sistema de defesa americano de manter superioridade técnica num contexto de corrida armamentista. Pilotos precisavam ser treinados mais rapidamente, com maior consistência, para operar equipamentos progressivamente mais complexos. O

² GAGNÉ, Robert M. *Instructional Technology: Foundations. Review of Educational Research*, v. 58, n. 3, p. 265–291, 1988.

design instrucional nasceu, portanto, sob pressão de alto custo de falha. Não havia tolerância para variação de qualidade dependente do instrutor. A instrução precisava funcionar como sistema, não como arte individual.

8.2. O PROBLEMA QUE GAGNÉ ESTAVA RESOLVENDO

O problema que Gagné nomeou com precisão tinha dois componentes, e a confusão entre eles é, até hoje, a fonte de boa parte do fracasso instrucional sistemático. O primeiro componente era a dependência do talento: a constatação de que bons resultados de aprendizagem, quando ocorriam, eram produto da habilidade pessoal do instrutor e não da estrutura da instrução. Isso tornava os resultados educacionais essencialmente não-replicáveis. Um bom professor produzia boa aprendizagem. Na ausência desse professor, os resultados desabavam. O sistema não era um sistema: era a soma de performers individuais.

O segundo componente era mais profundo e menos frequentemente articulado. Mesmo quando instrutores talentosos conseguiam ensinar bem, eles raramente conseguiam explicar o que faziam. O conhecimento sobre como ensinar era, ele mesmo, tácito no sentido de Polanyi: presente na execução, ausente na descrição. Um instrutor experiente de pilotagem sabia intuitivamente em que momento introduzir um conceito novo, quando era preciso deixar o aluno errar antes de corrigir, quanto feedback era informativo sem ser paralisante. Mas quando pedido para articular esses julgamentos em regras transmissíveis, o instrutor encontrava dificuldade genuína. A expertise instrucional resistia à verbalização da mesma forma que qualquer outra expertise complexa.

O projeto de Gagné era converter esse conhecimento tácito em conhecimento proposicional: extrair da prática dos melhores instrutores

os princípios que tornavam sua instrução eficaz, e traduzi-los em sequências prescritivas que qualquer instrutor treinado pudesse seguir com resultados consistentes. Isso não era democratização superficial do talento. Era engenharia de sistemas aplicada à instrução.

Há um terceiro problema implícito que o projeto abordava, e que se tornaria cada vez mais central nas décadas seguintes: a heterogeneidade dos resultados de aprendizagem. O fracasso de boa parte da instrução formal reside não apenas na ausência de bons instrutores, mas na suposição de que “aprender” é uma categoria homogênea que responde a um único tipo de intervenção. Aprender um fato histórico é uma coisa. Dominar um procedimento cirúrgico é outra. Desenvolver a capacidade de resolver problemas de otimização em contextos ambíguos é outra ainda. Cada um desses resultados envolve processos cognitivos distintos, e exige, portanto, requisitos instrucionais distintos ¹.

O enunciado é simples. Na prática, é sistematicamente ignorado. Programas de treinamento corporativos aplicam sequências de exposição a conteúdo indistintamente a todos os tipos de objetivo, como se a transmissão de informação e o desenvolvimento de habilidade fossem o mesmo problema. Gagné havia nomeado essa confusão nos anos 1960. Ela persiste intacta.

8.3. IDEIA CENTRAL

8.3.1. *Os Cinco Tipos de Resultado de Aprendizagem*

A pedra angular do sistema de Gagné é a taxonomia dos resultados de aprendizagem. Ao contrário de Bloom, que organizou resultados em hierarquia de complexidade crescente dentro de um único domínio cognitivo, Gagné propôs cinco categorias qualitativamente distintas de resul-

tado, cada uma com condições de aprendizagem específicas e irreduzíveis entre si ¹.

A primeira categoria é a informação verbal: o aprendiz é capaz de enunciar, recordar ou reconhecer proposições. Saber que Paris é a capital da França, que a Lei de Ohm relaciona tensão, corrente e resistência, que a Constituição de 1988 define os direitos fundamentais. A informação verbal não é habilidade: o aprendiz pode enunciar a Lei de Ohm sem ser capaz de aplicá-la a um problema de circuito real. Sua presença é condição necessária, mas não suficiente, para o desenvolvimento de habilidades intelectuais no mesmo domínio.

A segunda categoria são as habilidades intelectuais, que Gagné subdivide em discriminações, conceitos concretos, conceitos definidos, regras e regras de ordem superior. Essa subcategoria é a mais extensamente desenvolvida no sistema, porque representa a maior parte do que se entende por “competência” em contextos profissionais. Uma discriminação é a capacidade de distinguir estímulos: saber que dois compostos químicos têm propriedades distintas. Um conceito concreto é a capacidade de identificar instâncias de uma categoria: reconhecer um melanoma numa imagem dermatológica. Um conceito definido envolve critérios verbalizáveis: definir o que constitui um conflito de interesse numa transação financeira.

Regras são procedimentos aplicáveis: calcular a taxa interna de retorno. Regras de ordem superior são combinações de regras aplicadas a situações novas, o que se aproxima do que Bloom chamaria de Análise e Síntese.

A terceira categoria são as estratégias cognitivas: os processos que o aprendiz usa para gerenciar sua própria aprendizagem e pensamento. Como o aprendiz organiza informação nova, como monitora sua própria compreensão, como seleciona estratégias de solução de problemas.

Gagné reconhecia que estratégias cognitivas são aprendidas, mas que seu desenvolvimento é lento e altamente dependente de experiência acumulada. Não se ensinam estratégias cognitivas da mesma forma que se ensinam regras procedimentais.

A quarta categoria são as atitudes, predisposições que influenciam escolhas de ação. Um profissional de saúde pode saber os procedimentos de segurança e ter a habilidade de executá-los, mas não os executar consistentemente porque não internalizou a atitude de cautela sistemática. Atitudes não são ensinadas por exposição a informação. São moldadas por modelos, por experiências com consequências reais, por estruturas de aprovação social.

A quinta categoria são as habilidades motoras: coordenação de movimentos físicos, desde a caligrafia até a cirurgia laparoscópica. Habilidades motoras têm estrutura de aprendizagem distinta das demais: exigem prática física repetida, feedback proprioceptivo, e deterioram com ausência de prática de forma que habilidades intelectuais geralmente não deterioram.

A importância dessa taxonomia não é apenas classificatória. É prescritiva: cada categoria exige condições de instrução radicalmente diferentes, e confundir as categorias produz instrução sistematicamente ineficaz.

8.3.2. Os Nove Eventos de Instrução

Os Nove Eventos de Instrução representam a contribuição mais operacional de Gagné, e a que mais diretamente se presta à aplicação em design de produtos educacionais¹. Os eventos não são uma lista de “boas práticas”: são condições que, quando presentes em sequência, ativam os processos cognitivos necessários à aprendizagem. A ausência de qualquer evento não é uma falha de elegância: é uma falha estrutural que

compromete a cadeia de processamento.

O Evento 1 é obter atenção. Antes que qualquer aprendizagem possa ocorrer, o sistema cognitivo do aprendiz precisa estar orientado para o estímulo relevante. Isso parece trivial, mas a atenção não é um estado binário: é um recurso que compete com outros estímulos, com pensamentos intrusivos, com preocupações não relacionadas ao conteúdo. Gagné não propunha truques de entretenimento: propunha que o início de qualquer episódio instrucional incluía deliberadamente um estímulo que oriente o processamento para o que virá.

O Evento 2 é informar o objetivo. O aprendiz precisa saber o que será capaz de fazer ao final do episódio. Isso não é formalidade burocrática: é um mecanismo cognitivo com função precisa. Quando o aprendiz sabe o resultado esperado, ele pode ativar esquemas de processamento relevantes, alocar atenção de forma seletiva, e usar o objetivo como critério de avaliação do próprio desempenho. Um episódio instrucional sem objetivo explícito obriga o aprendiz a adivinhar o que é relevante. A adivinhação produz processamento difuso, não aprendizagem estruturada.

O Evento 3 é ativar conhecimento prévio. Ausubel havia demonstrado que aprendizagem significativa ocorre quando novo conteúdo pode ser ancorado em estruturas cognitivas existentes³. Gagné incorpora esse princípio operacionalmente: antes de apresentar conteúdo novo, o design instrucional deve ativar deliberadamente o que o aprendiz já sabe e que é relevante para o que virá. Isso não é revisão por cortesia: é preparação cognitiva. O novo conhecimento não flutua livre no espaço mental: ele se conecta a estruturas preexistentes, e a qualidade dessa conexão determina a qualidade da retenção.

O Evento 4 é apresentar o conteúdo. É o único evento que a maioria dos produtos educacionais executa com alguma consistência, e mesmo aqui, a maioria falha por tratar apresentação como suficiente. Gagné era

³ AUSUBEL, David P. *Educational Psychology: A Cognitive View*. Holt, Rinehart and Winston, 1968.

preciso: apresentação não é aprendizagem. É o fornecimento do material sobre o qual o processamento subsequente vai trabalhar. A qualidade da apresentação importa, mas ela não substitui os eventos que a seguem.

O Evento 5 é guiar a aprendizagem. Aqui o design instrucional fornece ao aprendiz estruturas que facilitam o processamento: exemplos elaborados, analogias, esquemas organizadores, modelos parcialmente resolvidos que demonstram o raciocínio antes de exigir que o aprendiz o reproduza de forma independente. O ponto de Gagné era que aprendizagem não guiada é ineficiente: o aprendiz sem orientação tende a processar o conteúdo de forma superficial, armazenar apenas o que é mais saliente, e perder as conexões estruturais que tornam o conhecimento generalizável.

O Evento 6 é provocar o desempenho. O aprendiz precisa produzir uma resposta ativa que demonstre a aprendizagem pretendida. Não reconhecimento passivo: produção. Não verificação de compreensão: execução. A distinção é crítica. A capacidade de reconhecer a resposta correta quando apresentada não implica a capacidade de gerá-la sem apoio. O Evento 6 força o processamento de recuperação, que é estruturalmente diferente do processamento de reconhecimento e produz traços de memória substancialmente mais robustos ⁴.

O Evento 7 é fornecer feedback. Sem feedback, o Evento 6 é truncado: o aprendiz produz uma resposta sem saber se ela corresponde ao critério. O feedback que Gagné descreve não é validação emocional: é informação diagnóstica sobre a qualidade do desempenho em relação ao padrão estabelecido pelo objetivo. Feedback de qualidade insuficiente (correto/incorreto) é melhor do que ausência de feedback, mas substancialmente inferior ao feedback que identifica especificamente onde o desempenho divergiu do padrão e por que essa divergência importa.

⁴ ROEDIGER, Henry L.; KARPICKE, Jeffrey D. Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychological Science*, v. 17, n. 3, p. 249–255, 2006.

O Evento 8 é avaliar o desempenho. Distinto do feedback do Evento 7, a avaliação do Evento 8 verifica se a aprendizagem se consolidou a um nível que satisfaz o critério de domínio estabelecido pelo objetivo. É uma verificação de estado, não de processo. O design instrucional precisa especificar que nível de desempenho constitui evidência de aprendizagem, e a avaliação precisa ser adequadamente sensível para distinguir desempenho que satisfaz esse critério de desempenho que o aproxima sem atingi-lo.

O Evento 9 é promover transferência. A aprendizagem que não transfere para contextos novos é aprendizagem de utilidade limitada. O Evento 9 deliberadamente expõe o aprendiz a variações do problema aprendido, a contextos distintos dos usados na instrução, a situações que exigem adaptação do conhecimento adquirido. Sem esse evento, o conhecimento permanece inerte: disponível quando o contexto é idêntico ao da instrução, indisponível quando o contexto varia mesmo superficialmente.

A sequência não é arbitrária. Cada evento cria condições cognitivas que o seguinte pressupõe: atenção é condição para que o objetivo seja processado, o objetivo ativa estruturas que tornam o conhecimento prévio relevante, e o conhecimento prévio ativado cria os ganchos que tornam a apresentação do conteúdo significativa. A cadeia estende-se até o Evento 9. Saber que ela existe, e por que cada elo importa, é condição para projetar produtos que não quebrem onde a maioria quebra.

8.4. EVOLUÇÃO POSTERIOR DA IDEIA

O modelo ADDIE (Analysis, Design, Development, Implementation, Evaluation), que se tornaria o framework dominante de design instrucional a partir dos anos 1970, é em grande medida uma operacionalização do

raciocínio de Gagné em nível de processo organizacional ⁵. Enquanto os Nove Eventos descrevem a microestrutura de um episódio instrucional, o framework descreve o macroprocesso de desenvolvimento de um programa: análise das necessidades e do público, design do currículo e das sequências, desenvolvimento dos materiais, implementação e avaliação de resultados. Sua origem é atribuída a pesquisadores da Florida State University nos anos 1970, instituição onde Gagné estava lotado no mesmo período, e a convergência tem origem comum.

O trabalho de Dick e Carey, formalizado no livro *The Systematic Design of Instruction*, publicado em 1978 ⁶, expandiu o sistema de Gagné numa direção mais explicitamente processual. Os autores produziram um modelo que especificava cada etapa do desenvolvimento instrucional com critérios mensuráveis de qualidade: desde a análise de tarefas até a revisão com base em dados de avaliação. O modelo Dick-Carey tornou-se, nas décadas seguintes, o padrão de referência em programas de formação de designers instrucionais nas universidades americanas.

A chegada do e-learning nos anos 1990 submeteu o sistema de Gagné a uma adaptação que revelou tanto sua robustez quanto suas limitações. A robustez: os Nove Eventos são independentes do meio. Podem ser implementados em instrução presencial, por texto impresso, em plataforma digital, em simulação. A estrutura cognitiva que eles ativam não muda com o canal.

A limitação era de outra ordem: a implementação digital de eventos como “fornecer feedback” exigiu o desenvolvimento de sistemas de avaliação automatizada que, nas primeiras gerações de e-learning, eram grosseiros demais para satisfazer os critérios de qualidade especificados por Gagné. O resultado foi uma expansão massiva de produtos que implementavam os Eventos 1 a 5 com competência variável e simplesmente

⁵ BRANCH, Robert Maribe. *Instructional Design: The ADDIE Approach*. Springer, 2009.

⁶ DICK, Walter; CAREY, Lou; CAREY, James O. *The Systematic Design of Instruction*. 8. ed. Pearson, 2015.

ignoravam os Eventos 6 a 9.

No Brasil, Andrea Filatro produziu, a partir dos anos 2000, uma das sistematizações mais rigorosas do design instrucional contextualizado para o ambiente educacional brasileiro⁷. O trabalho não é simples tradução: é adaptação que leva em conta as especificidades do ensino superior brasileiro, a cultura de avaliação formativa, e as condições de implementação de e-learning em contextos com restrições de infraestrutura. A contribuição mais relevante para o presente argumento é a distinção entre design instrucional fixo e design instrucional aberto: o primeiro produz materiais fechados que prescrevem a sequência instrucional; o segundo cria estruturas que permitem adaptação em tempo real ao perfil e ao desempenho do aprendiz⁸.

A crítica mais consistente ao sistema de Gagné veio, paradoxalmente, de dentro da tradição de design instrucional. Michael Hannafin e Susan Land argumentaram, nos anos 1990, que os ambientes de aprendizagem abertos (Open Learning Environments) são epistemicamente superiores aos ambientes prescritivos de Gagné para determinados tipos de resultado, especialmente os que envolvem raciocínio em domínios mal-definidos⁸. A instrução altamente estruturada de Gagné seria otimizada para aquisição de habilidades procedimentais bem-definidas, mas subótima para o desenvolvimento de pensamento crítico em domínios onde os próprios critérios de qualidade são contestados.

A crítica é válida e aponta para um limite real do sistema. Gagné nunca afirmou que os Nove Eventos eram adequados a todo tipo de resultado: sua taxonomia existe precisamente para distinguir resultados que requerem instrução diferente. O problema prático é que o sistema foi aplicado indiscriminadamente, com a prescrição de Gagné para habilidades procedimentais sendo usada para tentar desenvolver estratégias

⁷ FILATRO, Andrea. *Design Instrucional Contextualizado: Educação e Tecnologia*. Senac São Paulo, 2004.

⁸ HANNAFIN, Michael J.; LAND, Susan M. The Foundations and Assumptions of Technology-Enhanced Student-Centered Learning Environments. *Instructional Science*, v. 25, n. 3, p. 167–202, 1997.

cognitivas e disposições que requerem ambientes estruturalmente distintos. O inverso também é verdadeiro: onde o sistema foi aplicado aos domínios para os quais foi construído, os resultados documentam uma eficácia que a crítica acadêmica raramente reconhece com a mesma clareza com que aponta os limites.

8.5. IMPACTO REAL NO MUNDO

O impacto de Gagné nos sistemas de formação de alto risco é tão profundo que se tornou invisível: a maioria dos profissionais que projeta programas em aviação, medicina ou operações militares aplica princípios derivados do seu sistema sem ter lido uma linha de sua obra.

O caso dos pilotos comerciais é um exemplo de implementação implícita dos Nove Eventos com fidelidade notável: os briefings pré-voos ativam conhecimento prévio e informam objetivos, as sessões de simulação provocam desempenho em condições progressivamente mais exigentes, os debriefings fornecem feedback diagnóstico detalhado sobre cada decisão tomada. A progressão de exercícios de simulação promove transferência por exposição sistemática a variações de cenário. Que isso seja chamado de “certificação por proficiência demonstrada” ou de “treinamento baseado em competências” é terminologia. A estrutura subjacente é de Gagné.

A indústria global de treinamento corporativo apresenta o contraste mais revelador. Ela movimenta, segundo estimativas da indústria, mais de 360 bilhões de dólares anuais em investimento⁹. Uma proporção substancial desse investimento produz resultados que não são mensurados de forma rigorosa. Onde há medição, os resultados são consistentemente

⁹ STATISTA. *Corporate Training Market; Global Spending*. Disponível em: [statista.com](https://www.statista.com). Acesso em 2024.

decepcionantes. O estudo de Saks, publicado em 2006 no *Journal of Vocational Behavior*¹⁰, documentou que a maioria do conhecimento adquirido em programas corporativos não é transferida para o trabalho real, com estimativas que variam entre 10% e 40% dependendo do domínio e do método. A análise de por que a transferência falha invariavelmente aponta para a ausência sistemática dos eventos 6, 7, 8 e 9.

Programas de certificação profissional de alto stakes, como o CFA, o USMLE (exame de licenciatura médica americano) e os exames de certificação do Project Management Institute, implementam os Nove Eventos com rigor que o mercado de treinamento corporativo raramente alcança. O CFA explicita os objetivos de aprendizagem por nível (informar o objetivo), organiza o currículo em progressão que parte de conceitos fundamentais para análise integrada (ativar conhecimento prévio, apresentar conteúdo, guiar aprendizagem), e exige prática extensiva com questões que replicam o formato do exame real (provocar desempenho). O sistema fornece feedback por área temática nos resultados de avaliação e administra exames que progressivamente exigem integração de conceitos de múltiplas áreas, o que corresponde ao Evento 9 de transferência.

A taxa de aprovação no nível 1 do CFA raramente supera 40%, o que reflete a seriedade com que o sistema trata o critério de domínio do Evento 8.

Na educação médica, a reforma do treinamento baseado em competências, que ganhou tração a partir dos anos 2000 com o trabalho do Accreditation Council for Graduate Medical Education nos Estados Unidos, é estruturalmente uma implementação dos princípios de Gagné. O modelo anterior, “ver um, fazer um, ensinar um”, foi substituído por progressão documentada de competências com avaliação formativa em cada etapa. O uso de simuladores para prática deliberada antes da exposição

¹⁰ SAKS, Alan M. Problem and Solutions to the Training Transfer Problem. In: *The Handbook of Human Resource Management*. Blackwell Publishing, 2006.

a pacientes reais, e os sistemas de feedback estruturado de supervisores: todos esses elementos correspondem a eventos específicos da sequência de Gagné.

O fracasso de cursos corporativos bem-intencionados segue um padrão documentável. O curso foi desenvolvido por especialistas com domínio técnico impecável do conteúdo. O conteúdo, organizado em sequência lógica. Os materiais, produzidos com qualidade. E no final, os participantes saíram com pouca ou nenhuma mudança de comportamento observável. O diagnóstico invariavelmente revela a mesma estrutura: os eventos 1 a 5 foram executados com cuidado variável, os eventos 6 e 7 foram reduzidos a questionários de múltipla escolha ao final de cada módulo, e os eventos 8 e 9 estavam ausentes por completo. O produto havia sido projetado para cobrir conteúdo, não para produzir aprendizagem.

Entender por que essa confusão persiste exige olhar para além do sistema de Gagné isolado. O que torna os eventos 6 e 7 tão sistematicamente negligenciados não é apenas ignorância do modelo: é a confluência de pressões que os outros autores deste livro ajudam a nomear.

8.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

8.6.1. *Gagné e Bloom: Tipos de Aprendizagem e Níveis da Taxonomia*

A relação entre Gagné e Bloom é de complementaridade com tensão. Bloom organizou resultados cognitivos em hierarquia de complexidade: Lembrar, Entender, Aplicar, Analisar, Avaliar, Criar. Gagné organizou resultados em categorias qualitativamente distintas: informação verbal, habilidade intelectual, estratégia cognitiva, atitude, habilidade motora. As duas taxonomias não são concorrentes: são ortogonais, e a interse-

ção entre elas é operacionalmente útil.

Uma habilidade intelectual pode existir em qualquer nível da taxonomia de Bloom. Discriminar estímulos é uma operação próxima do Nível 1 de Bloom. Aplicar regras procedimentais corresponde ao Nível 3. Desenvolver regras de ordem superior para resolver problemas novos corresponde aos Níveis 4 e 5. O que Gagné acrescenta a Bloom é a especificação das condições de aprendizagem: não apenas o nível de complexidade do resultado, mas a natureza do processo cognitivo que o produz e os eventos que precisam estar presentes para que esse processo ocorra.

A implicação para design instrucional é que identificar o nível de Bloom do objetivo instrucional é necessário mas não suficiente para determinar o design instrucional adequado. Um objetivo no Nível 3 de Bloom (Aplicar) que envolve habilidade intelectual procedimental requer condições distintas de um objetivo no mesmo nível que envolve atitude ou estratégia cognitiva. Bloom diz o quê do resultado. Gagné diz que tipo de resultado é e, portanto, quais condições são necessárias.

8.6.2. Gagné e Glasser: Eventos de Instrução como Engenharia da Retenção

O capítulo sobre William Glasser estabeleceu que diferentes formas de engajamento com o conteúdo produzem diferentes níveis de retenção, com as modalidades ativas consistentemente superiores às passivas¹¹. A integração com Gagné é direta: os Eventos 6 e 7 (provocar desempenho e fornecer feedback) são precisamente as condições que ativam as formas ativas de aprendizagem que Glasser identifica como produtoras de maior retenção.

A instrução que se limita aos Eventos 1 a 5 corresponde, na hierar-

¹¹ GLASSER, William. *Choice Theory in the Classroom*. HarperCollins, 1998.

quia de Glasser, às modalidades passivas: ouvir, ler, ver. A instrução que implementa os Eventos 6 a 9 força o aprendiz a praticar, a demonstrar, a ensinar, que são as formas de engajamento associadas a retenção substancialmente maior. Gagné fornece a sequência que torna possível chegar aos eventos mais eficazes; o modelo de Glasser fornece a evidência empírica de por que esses eventos específicos produzem retenção superior.

Há uma tensão na integração: Glasser descreve a estrutura neurológica da retenção de forma que poderia sugerir que os eventos anteriores são dispensáveis, desde que se chegue à prática ativa suficientemente cedo. Gagné argumentaria que os eventos anteriores são condições funcionais: a prática ativa sem ativação de conhecimento prévio, sem objetivo claro, sem conteúdo apresentado adequadamente, produz prática confusa, não aprendizagem estruturada. A pirâmide de Glasser não funciona sem a arquitetura de Gagné que a precede.

8.6.3. Gagné e Sweller: Eventos, Carga Cognitiva e Progressão de Complexidade

John Sweller desenvolveu a Teoria da Carga Cognitiva a partir dos anos 1980, argumentando que a memória de trabalho tem capacidade limitada e que o design instrucional deve gerenciar deliberadamente a carga imposta ao aprendiz para evitar sobrecarga que comprometa a aprendizagem¹². A integração com Gagné é de convergência com especificação adicional.

A sequência proposta por Gagné, quando implementada com fidelidade, opera por construção como gerenciadora de carga cognitiva. Ativar conhecimento prévio (Evento 3) reduz a demanda imposta pelo conteúdo novo porque fornece estruturas às quais o novo conhecimento

¹² SWELLER, John. Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, v. 12, n. 2, p. 257-285, 1988.

pode ser ancorado, evitando o processamento de material genuinamente novo em toda a sua complexidade. Guiar a aprendizagem (Evento 5) alivia essa pressão através de exemplos elaborados que externalizam parte do processamento que o aprendiz teria que realizar internamente. A progressão dos Eventos 6 a 9, do desempenho guiado ao desempenho independente e à transferência, corresponde ao que Sweller chamaria de redução gradual do suporte externo à medida que o esquema interno se consolida.

O que Sweller acrescenta a Gagné é a arquitetura neurocognitiva que explica por que a sequência de Gagné funciona. Gagné havia derivado os eventos empiricamente, da observação de instrução eficaz. Sweller fornece o mecanismo: a sequência respeita os limites da memória de trabalho, distribui a carga ao longo do tempo, e permite a consolidação progressiva de esquemas que reduzem a carga cognitiva de tarefas subsequentes. Os dois modelos são mutuamente reforçadores.

Bloom, Glasser e Sweller chegam cada um a Gagné por rotas distintas e convergem no mesmo ponto: a sequência dos Nove Eventos não é conveniência pedagógica. É arquitetura cognitiva com fundamento empírico. O que resta é traduzir essa arquitetura em especificação de produto.

9

A MENTE TEM LIMITES: JOHN SWELLER E A TEORIA DA CARGA COGNITIVA

—*“A nossa capacidade de processar informação nova é tão limitada que qualquer design instrucional que ignore esse fato está construindo sobre areia.”* John Sweller, *Cognitive Architecture and Instructional Design*, 1998¹

9.1. CONTEXTO HISTÓRICO

Em 1956, George Armitage Miller publicou um artigo no *Psychological Review* que começava com uma confissão incomum para um texto científico: “Minha dificuldade é que fui perseguido por um número inteiro”². O número era sete. Miller havia notado, através de uma série de experi-

¹ SWELLER, John; VAN MERRIËNBOER, Jeroen J. G.; PAAS, Fred. *Cognitive Architecture and Instructional Design*. *Educational Psychology Review*, v. 10, n. 3, p. 251–296, 1998.

² MILLER, George A. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review*, v. 63, n. 2, p. 81–97, 1956.

mentos sobre percepção e memória imediata, que os seres humanos conseguiam reter aproximadamente sete unidades de informação na consciência ao mesmo tempo, com uma margem de dois para mais ou para menos. O artigo tornou-se um dos mais citados na história da psicologia cognitiva, não porque revelou um fato novo sobre a mente, mas porque o nomeou com precisão suficiente para que outros pudessem construir sobre ele.

O que Miller havia identificado era o que hoje se chama de memória de trabalho: o sistema cognitivo que mantém informação ativa durante o processamento. Não o arquivo de longo prazo onde residem habilidades, conhecimentos e memórias consolidadas, mas o espaço de trabalho temporário onde novas informações são combinadas, transformadas e eventualmente transferidas para armazenamento permanente, ou simplesmente descartadas. Sete itens parecia um número modesto para um sistema tão central às funções cognitivas superiores. Nas décadas seguintes, ficaria claro que mesmo esse número era generoso.

A revisão viria de Nelson Cowan, que em 2001 argumentou que o limite funcional da memória de trabalho era mais próximo de quatro chunks, não sete, quando se controlava rigorosamente a possibilidade de repetição verbal silenciosa e outras estratégias compensatórias³. O número exato permanece em debate, e a controvérsia continua produtiva. O princípio, contudo, era claro: a mente humana, durante o processamento de informação nova, opera sob uma restrição severa de capacidade. Ignorar essa restrição no design de materiais de ensino não é apenas um descuido pedagógico. É uma violação estrutural da arquitetura do sistema que se pretende ensinar.

Alan Baddeley formalizou o modelo de memória de trabalho em 1974, e o expandiu substancialmente nas décadas seguintes⁴. O sistema que

³ COWAN, Nelson. The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity. *Behavioral and Brain Sciences*, v. 24, n. 1, p. 87-114, 2001.

⁴ BADDELEY, Alan D. *Working Memory*. Oxford University Press, 1986.

ele descreveu tem três componentes: a alça fonológica, que mantém informação verbal através de repetição interna; o bloco de notas visuo-espacial, que mantém representações visuais e espaciais; e o executivo central, que coordena os recursos das outras duas e gerencia a atenção. Uma quarta estrutura, o buffer episódico, foi adicionada em 2000 para dar conta de como as informações dos diferentes subsistemas são integradas.

O que o modelo tornava claro era que a memória de trabalho não é um recipiente passivo, mas um sistema ativo que consome recursos cognitivos. Processar informação visual e verbal simultaneamente não é operação gratuita: há competição por recursos limitados entre os subsistemas.

John Sweller era, na época em que começou seu trabalho mais relevante, um pesquisador australiano da Universidade de New South Wales em Sydney, interessado inicialmente na psicologia da solução de problemas. Não chegou à teoria da carga cognitiva partindo da neurociência da memória, mas de uma observação prática que o incomodava: estudantes expostos a materiais didáticos bem desenhados frequentemente aprendiam menos do que estudantes expostos a versões mais simples dos mesmos materiais. O problema não era o conteúdo. Era a forma como o conteúdo era apresentado.

O artigo de 1988 que lançou a teoria, publicado na *Cognitive Science*, partia de um paradoxo específico ⁵. Estudantes de matemática que resolviam problemas convencionais por meios-fins, a estratégia padrão de tentar reduzir a diferença entre o estado atual e o estado objetivo, aprendiam menos do que estudantes que estudavam exemplos resolvidos sem tentar resolver problemas por conta própria. A explicação behaviourista, que associava aprendizagem à prática ativa, previa o oposto. A explica-

⁵ SWELLER, John. Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, v. 12, n. 2, p. 257-285, 1988.

ção que emergiu do trabalho de Sweller dizia que a estratégia de meios-fins, embora eficaz para resolver o problema imediato, consumia toda a capacidade disponível da memória de trabalho no gerenciamento da busca por soluções, deixando nenhuma capacidade livre para a consolidação dos princípios subjacentes. O estudante resolvia o problema. Não aprendia o conteúdo.

9.2. O PROBLEMA QUE SWELLER ESTAVA RESOLVENDO

Havia um problema anterior ao problema de Sweller. Ele pertencia à tradição que a psicologia cognitiva havia herdado do behaviorismo: a tendência de avaliar materiais de ensino pela sua completude, pela sua organização lógica, pela riqueza dos exemplos fornecidos. Um manual técnico bem construído era aquele que cobria todos os casos possíveis, antecipava todas as dúvidas, fornecia todos os contextos relevantes. A lógica era intuitivamente atraente: quanto mais informação o material fornecesse, menos o aprendiz teria que buscar por conta própria, e mais eficiente seria a aprendizagem.

Sweller via o problema de ângulo diferente. A questão não era quanto o material continha, mas quanto o processamento desse material exigia da memória de trabalho do aprendiz. Um manual que cobria todos os casos possíveis era, do ponto de vista cognitivo, potencialmente um instrumento de sobrecarga sistemática. O aprendiz iniciante não tinha os esquemas cognitivos necessários para distinguir o que era central do que era periférico, o que era princípio do que era exemplo, o que era regra do que era exceção. Para ele, tudo chegava com o mesmo peso. E a memória de trabalho, confrontada com esse fluxo indiscriminado, capitulava.

O “problema do expert que não sabe o que o iniciante não sabe” tem

nome na literatura: a maldição do conhecimento, fenômeno que Sweller havia identificado de forma operacional antes de receber esse rótulo. Um expert em física quântica que escreve um material introdutório opera, ao escrever, a partir de uma estrutura de conhecimento densa, com conexões automáticas entre conceitos, com padrões reconhecidos instantaneamente, com a distinção entre essencial e instrumental já consolidada após anos de prática. Ele não acessa mais esses conceitos como informação nova: acessa-os como esquemas automatizados que consomem quase nenhum espaço na memória de trabalho. A consequência é que ele não consegue mais perceber a carga cognitiva que seu material impõe a alguém que ainda está construindo esses esquemas do zero.

Isso explica, em parte, por que experts frequentemente produzem materiais didáticos que outros experts acham claros e iniciantes acham impenetráveis. Não é falta de esforço. Não é falta de boa intenção. É uma falha estrutural de perspectiva que a teoria da carga cognitiva nomeia com precisão incomum.

Havia também um problema metodológico que tornava o fenômeno invisível. As métricas dominantes de avaliação de materiais educacionais mediam resultado, não processo. Avaliava-se se o estudante aprendia ao final do curso, não como sua memória de trabalho estava sendo utilizada durante a aprendizagem. Um material que impunha carga cognitiva excessiva podia ainda produzir alguma aprendizagem, especialmente em estudantes muito motivados ou com alta capacidade de memória de trabalho. Mas produzia menos aprendizagem do que produziria se a carga fosse gerenciada adequadamente. E a diferença era quase invisível sem um framework que tornasse a carga cognitiva um objeto de medição explícito ⁶.

O que faltava, portanto, não era apenas evidência empírica: era uma

⁶ SWELLER, John. Cognitive Load Theory, Learning Difficulty, and Instructional Design. *Learning and Instruction*, v. 4, n. 4, p. 295-312, 1994.

taxonomia do problema que permitisse distinguir tipos de carga por origem, por tratamento e por efeito sobre a aprendizagem.

9.3. IDEIA CENTRAL

9.3.1. *Os Tipos de Carga e a Equação do Design*

A formulação consolidada em 1998 por Sweller, Van Merriënboer e Paas propõe que a carga total imposta ao aprendiz em qualquer episódio instrucional tem três fontes distintas ⁵. Cada uma tem causas diferentes e, conseqüentemente, implicações de design diferentes.

A primeira fonte é a carga intrínseca. Ela deriva da complexidade inerente do material a ser aprendido, definida tecnicamente pelo número de elementos que precisam ser processados simultaneamente porque interação entre si. Um vocabulário isolado tem baixa carga intrínseca: aprender que “gato” em inglês é “cat” não exige processamento simultâneo de múltiplos elementos interconectados. Aprender a resolver uma equação diferencial parcial tem carga intrínseca alta: as variáveis, as relações entre elas, as condições de contorno, os métodos de solução, interação de formas que tornam o processamento isolado de qualquer um deles insuficiente.

A carga intrínseca não pode ser eliminada pelo design. Pode ser gerenciada, por exemplo sequenciando o material de forma que elementos de baixa interatividade sejam aprendidos antes dos de alta interatividade, reduzindo temporariamente a demanda de processamento simultâneo. Mas a complexidade subjacente do domínio persiste.

A segunda fonte é a carga extrínseca, também chamada de carga estranha. Ela deriva não da complexidade do conteúdo, mas da forma como esse conteúdo é apresentado. Um diagrama técnico em que o aprendiz

precisa alternar entre a figura e a legenda numa outra parte da página impõe carga extrínseca: o processamento simultâneo necessário não é exigido pela complexidade do conteúdo, mas pela separação física de elementos que deveriam ser integrados. Esse fenômeno específico, chamado de efeito de atenção dividida (*split-attention effect*), é apenas um dos vários identificados na literatura.

A carga extrínseca não contribui para a aprendizagem. É ruído puro no sistema de processamento: consome recursos da memória de trabalho sem gerar nenhuma representação cognitiva útil. Todo design instrucional de qualidade deveria ter como objetivo explícito reduzir a carga extrínseca a zero ⁵.

A terceira fonte, a mais complexa conceitualmente, é a carga germinativa (*germane load*). Ela representa o esforço cognitivo dedicado especificamente à formação de esquemas: à integração de nova informação em estruturas cognitivas existentes, à abstração de padrões gerais a partir de exemplos específicos, à consolidação de procedimentos que, com prática suficiente, se tornarão automatizados. A carga germinativa é a única das três que está diretamente associada ao resultado de aprendizagem no sentido mais profundo: não ao armazenamento de informação, mas à construção das estruturas cognitivas que tornam possível a aplicação flexível do conhecimento em situações novas.

A relação entre as três fontes pode ser expressa como uma inequação: a soma de carga intrínseca e carga germinativa não pode exceder a capacidade total da memória de trabalho se a aprendizagem vai ocorrer. Qualquer carga extrínseca que ocupe espaço nesse sistema reduz o espaço disponível para carga germinativa, prejudicando exatamente o processo que define aprendizagem real.

Vale registrar que a formulação das três cargas foi revisada em 2019 por Sweller, Van Merriënboer e Paas na releitura dos vinte anos da teo-

ria ⁷. Os autores reconheceram que a distinção entre carga intrínseca e germinativa era menos nítida do que a formulação original sugeria: em certos sentidos, o esforço de formação de esquemas está intimamente entrelaçado com a natureza do material, não apenas com a forma como é apresentado. A revisão não abandona as três categorias, mas as trata com maior nuance, reconhecendo que carga intrínseca alta e design instrucional que favorece a formação de esquemas estão em diálogo, não em compartimentos separados.

9.3.2. *O Efeito dos Exemplos Resolvidos*

Uma das contribuições mais contraintuitivas da teoria, e simultaneamente uma das mais bem documentadas empiricamente, é o efeito dos exemplos resolvidos. Sweller e seus colaboradores demonstraram consistentemente que estudantes iniciantes aprendem mais estudando exemplos completamente resolvidos, seguindo o raciocínio passo a passo de quem já sabe fazer, do que resolvendo problemas por conta própria desde o início ¹. A explicação está no gerenciamento da carga cognitiva: quando um iniciante tenta resolver um problema sem esquemas adequados, a maior parte da sua memória de trabalho é consumida pelo processo de busca, pelo gerenciamento de múltiplas hipóteses, pela avaliação de dead ends. Pouca ou nenhuma capacidade permanece disponível para abstrair os princípios que tornariam o próximo problema mais fácil.

O exemplo resolvido elimina a busca. O aprendiz não precisa descobrir o caminho: precisa compreender por que cada passo foi dado, o que exige capacidade cognitiva diferente e substancialmente menor. Essa capacidade que sobra pode ser aplicada à formação de esquemas, que é exatamente o que a teoria define como aprendizagem.

Há um limite para esse efeito. Renkl demonstrou que o benefício dos

⁷ SWELLER, John; VAN MERRIËNBOER, Jeroen J. G.; PAAS, Fred. Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, v. 31, n. 2, p. 261–292, 2019.

exemplos resolvidos diminui progressivamente à medida que o aprendiz desenvolve esquemas adequados⁸. No ponto em que o aprendiz já possui os esquemas necessários, a resolução independente de problemas passa a ser superior ao estudo de modelos prontos: a prática ativa consolida e automatiza os esquemas já presentes, enquanto acompanhar uma solução passo a passo deixa de adicionar informação nova.

Esse ponto de transição é o que Sweller e colaboradores chamaram de efeito de reversão da expertise (*expertise reversal effect*): uma intervenção pedagógica que é ótima para um nível de domínio pode ser subótima ou até prejudicial para outro. Um material que serve perfeitamente a um iniciante pode impor carga desnecessária a um estudante intermediário que já não precisa de toda a estrutura que o exemplo fornece. A implicação para design instrucional é que materiais fixos, não calibrados ao nível de expertise do aprendiz, são estruturalmente inadequados para pelo menos parte de seu público-alvo⁶.

9.3.3. O Efeito de Atenção Dividida e Outros Efeitos

O efeito de reversão da expertise é apenas o mais contraintuitivo de uma família de padrões documentados sistematicamente pela literatura. A pesquisa em carga cognitiva identificou pelo menos uma dezena de fenômenos específicos que podem ser explorados ou evitados no design instrucional. O efeito de atenção dividida foi mencionado acima. O de redundância é seu complemento: quando informações que se explicam mutuamente são apresentadas conjuntamente, a presença de uma delas é redundante e pode paradoxalmente aumentar a carga ao forçar o aprendiz a processar e integrar informação que não adiciona nada. Um diagrama com legenda integrada não precisa também de um texto que descreva o que o diagrama mostra: a descrição textual, nesse caso, é ruído.

⁸ RENKL, Alexander. Toward an Instructionally Oriented Theory of Example-Based Learning. *Cognitive Science*, v. 38, n. 1, p. 1–37, 2014.

O efeito de modalidade (modality effect) descreve um padrão consistente: quando informação verbal é apresentada auditivamente, não por texto escrito, em combinação com informação visual, a aprendizagem melhora. A razão é arquitetural. Os dois subsistemas da memória de trabalho, a alça fonológica e o bloco visuoespacial, processam os dois canais em paralelo sem competir pelos mesmos recursos.

Mayer construiu sobre esse princípio uma teoria completa de aprendizagem multimídia ⁹, demonstrando empiricamente que narração em áudio combinada com imagens é superior a texto escrito combinado com imagens para a maioria dos tipos de conteúdo técnico. A violação mais comum desse princípio é a apresentação simultânea de texto escrito e voz narrando exatamente o mesmo texto: o aprendiz processa a mesma informação pelos dois canais, a redundância aumenta a carga sem nenhum benefício adicional de aprendizagem.

9.4. EVOLUÇÃO POSTERIOR DA IDEIA

Originalmente situada dentro da psicologia educacional, a teoria emergiu para outros domínios por rotas diferentes e com velocidades diferentes. A fidelidade à formulação original variou: a neurociência confirmou os mecanismos subjacentes; o design de software absorveu os princípios com imprecisões instrumentais; o mercado corporativo invocou o vocabulário sem a estrutura.

Na neurociência, a confirmação dos mecanismos subjacentes à carga cognitiva veio de duas direções. Os estudos de neuroimagem funcional com fMRI e PET mostraram padrões de ativação diferencial no córtex pré-frontal e no lobo parietal coerentes com o modelo de Baddeley e com

⁹ MAYER, Richard E. *Multimedia Learning*. 2. ed. Cambridge University Press, 2009.

os efeitos previstos pela teoria. Tarefas com alta carga cognitiva ativavam de forma mais intensa as regiões associadas à memória de trabalho, e a saturação dessas regiões correlacionava com erros de desempenho e com menor transferência de aprendizagem.

A confirmação não era perfeita: a neurociência cognitiva tem seus próprios debates sobre localização de funções e interpretação de correlações. Mas a convergência entre o modelo teórico e os dados de neuroimagem forneceu fundamento empírico independente para uma teoria que havia sido construída a partir da psicologia do comportamento observável, não da imagem cerebral.

A obra de Daniel Kahneman, especialmente *Thinking, Fast and Slow* publicado em 2011, tornou o conceito de capacidade cognitiva limitada acessível a audiências muito mais amplas¹⁰. A distinção de Kahneman entre Sistema 1, rápido e automático, e Sistema 2, lento e deliberado, mapeia diretamente sobre a distinção entre processamento esquematizado e processamento na memória de trabalho que a teoria da carga cognitiva descreve. Kahneman não cita Sweller extensivamente, e as tradições de origem são diferentes. Mas as duas teorias convergem no mesmo ponto empírico: a mente humana tem capacidade limitada de processamento consciente, e quando essa capacidade é excedida, algo quebra.

No campo de design de software e experiência de usuário, a teoria foi adotada com entusiasmo. O princípio geral chegou intacto; a distinção entre tipos de carga, nem sempre. O princípio de que interfaces devem minimizar a carga cognitiva imposta ao usuário tornou-se lugar-comum no setor a partir dos anos 2000. A lei de Hick, que descreve como o tempo de decisão aumenta logaritmicamente com o número de opções, e a lei de Miller foram incorporadas como heurísticas de design em textos fundamentais do campo.

A influência sobre o design da Apple foi documentada em vários con-

¹⁰ KAHNEMAN, Daniel. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.

textos. A obsessão de Steve Jobs com a eliminação de elementos desnecessários das interfaces, manifestada no iPod com uma única roda de controle e no iPhone com um único botão físico, é redução de carga extrínseca, mesmo que não tenha sido formulada nesses termos.

Kirschner, Sweller e Clark publicaram em 2006 um artigo que gerou controvérsia significativa ¹¹. O texto argumentava que aprendizagem por descoberta, aprendizagem baseada em problemas e outras formas de instrução com guia mínimo são fundamentalmente inadequadas para iniciantes: impõem carga cognitiva excessiva num estágio em que o aprendiz ainda não tem os esquemas necessários para gerenciar a busca de forma eficiente. O artigo foi respondido com vigor por defensores da aprendizagem construtivista, e o debate continua. A posição de Sweller não era que aprendizagem ativa é inútil: era que ela é inadequada como ponto de partida. A instrução guiada, com redução deliberada de carga durante a fase de aquisição de esquemas, deve preceder a prática independente, não o contrário.

Van Merriënboer desenvolveu, a partir dos anos 1990, uma aplicação da teoria especificamente para aprendizagem de tarefas complexas no ambiente de trabalho ¹². O modelo 4C/ID (Four-Component Instructional Design) propõe que habilidades complexas sejam ensinadas através de tarefas de aprendizagem inteiras, não decompostas artificialmente, mas com suporte decrescente que libera progressivamente a carga germinativa necessária para a integração dos componentes.

O modelo resolve uma tensão que a formulação original de Sweller deixava em aberto: como gerenciar carga cognitiva sem decompor o conteúdo em subproblemas artificiais que o aprendiz nunca encontrará na prática real. A resposta do 4C/ID é manter a tarefa inteira e reduzir o

¹¹ KIRSCHNER, Paul A.; SWELLER, John; CLARK, Richard E. Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, v. 41, n. 2, p. 75–86, 2006.

¹² VAN MERRIËNBOER, Jeroen J. G.; SWELLER, John. Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, v. 17, n. 2, p. 147–177, 2005.

suporte, não simplificar a tarefa. É precisamente esse conjunto de princípios, derivados de décadas de refinamento teórico, que o registro de aplicações fora da academia permite avaliar em escala.

9.5. IMPACTO REAL NO MUNDO

O impacto mais direto e documentado da teoria da carga cognitiva no mundo fora da academia está no design de materiais técnicos de alto risco. Manuais de segurança para aviação, protocolos de emergência em instalações nucleares, guias de procedimento médico: todos esses materiais passaram, a partir dos anos 1990 e 2000, por revisões de design que incorporavam os princípios da teoria. O rótulo usado raramente era “carga cognitiva”. As escolhas de design, frequentemente, eram as mesmas.

Clark, Nguyen e Sweller publicaram em 2006 um guia prático de aplicação da teoria ao design instrucional corporativo, tornando os princípios operacionais para designers não-acadêmicos¹³. A contribuição mais concreta foi traduzir os efeitos documentados pela pesquisa em diretrizes de design testáveis. As principais: integrar textos a diagramas em vez de separá-los; eliminar informação redundante; sequenciar conteúdo do mais simples ao mais complexo em termos de interatividade de elementos; usar narração em áudio em vez de texto escrito quando o conteúdo é acompanhado de material visual.

O fracasso dos slides com listas de bullet points como formato dominante de apresentação corporativa é, no vocabulário da teoria da carga cognitiva, um caso clássico de imposição de carga extrínseca desnecessária. O problema não é o formato em si. É a combinação de texto escrito

¹³ CLARK, Richard C.; NGUYEN, Frank; SWELLER, John. *Efficiency in Learning: Evidence-Based Guidelines to Manage Cognitive Load*. Pfeiffer, 2006.

sendo lido pela audiência simultaneamente com a voz do apresentador: dois canais verbais competindo pelos mesmos recursos da alça fonológica. A audiência precisa optar entre ler e ouvir, e qualquer das duas opções implica ignorar a outra.

O resultado é processamento fragmentado, baixa retenção, e a sensação familiar de que nada ficou depois de uma apresentação de hora com quarenta telas. Lawrence Lessig, Edward Tufte e outros críticos de slides chegaram ao mesmo diagnóstico por rotas estéticas e comunicacionais. Sweller fornece o mecanismo cognitivo subjacente.

Na educação médica, a pressão para integrar a teoria ao design de treinamentos foi acelerada pela publicação de Feldon em 2007¹⁴, que documentou especificamente como a carga cognitiva de instrutores afetava sua capacidade de ensinar e monitorar alunos simultaneamente. Um cirurgião experiente realizando uma operação acessa seus esquemas de forma automatizada, com baixíssima demanda da memória de trabalho para as tarefas técnicas. Um residente realizando a mesma operação consome toda sua memória de trabalho na execução dos procedimentos, com pouca ou nenhuma capacidade restante para monitorar o campo cirúrgico de forma abrangente. A implicação de design é que a supervisão de residentes precisa ser estruturada de forma que a carga cognitiva total, execução mais monitoramento, não exceda a capacidade disponível nas fases iniciais de formação.

No mercado de formação corporativa, a influência da teoria foi absorvida com mais entusiasmo do que precisão. O movimento de *microlearning*, que defende a decomposição de conteúdo educacional em unidades de dois a cinco minutos, invoca frequentemente a teoria da carga cognitiva como fundamento. A relação é real mas parcial. Unidades curtas de conteúdo reduzem a demanda de sustentação de informação na memória

¹⁴ FELDON, David F. Cognitive Load and Classroom Teaching: The Double-Edged Sword of Automaticity. *Educational Psychology Review*, v. 19, n. 2, p. 123–125, 2007.

de trabalho ao longo do tempo, o que é consistente com a teoria.

Mas a fragmentação de conteúdo com alta interatividade de elementos, em que os conceitos de uma unidade dependem dos de outra para fazer sentido, pode aumentar a carga total ao exigir integração de múltiplas fontes sem o suporte do contexto original. Microlearning é uma aplicação válida dos princípios da carga cognitiva em certos tipos de conteúdo. Não é uma solução universal.

Entender com precisão onde os limites estão exige voltar à estrutura teórica e examiná-la em relação a outros modelos que descrevem os mesmos fenômenos de ângulos diferentes.

9.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

9.6.1. *Sweller e Gagné: Sequência como Gestão de Carga*

O capítulo sobre Robert Gagné estabeleceu que os Nove Eventos de Instrução formam uma sequência em que cada evento cria condições cognitivas que o seguinte pressupõe. A Teoria da Carga Cognitiva fornece a explicação mecanicista de por que essa sequência funciona.

O Evento 3 de Gagné, ativar conhecimento prévio, reduz a carga intrínseca do material subsequente ao transformar elementos desconhecidos em elementos parcialmente conhecidos, diminuindo o número de itens que precisam ser processados como genuinamente novos. O Evento 5, guiar a aprendizagem através de exemplos elaborados, é o efeito dos exemplos resolvidos operacionalizado na sequência de Gagné. A guia externa externaliza o processamento de busca que o aprendiz teria que realizar internamente, liberando capacidade da memória de trabalho para a formação de esquemas.

A progressão dos Eventos 6 a 9, do desempenho guiado ao indepen-

dente e à transferência, corresponde ao gradual fade-out do suporte externo. À medida que os esquemas internos se consolidam, o suporte torna-se redundante, e sua presença passa a impor carga desnecessária. É exatamente a gestão do ponto de reversão da expertise, descrita na seção anterior, que a sequência de Gagné realiza de forma implícita.

A afirmação mais precisa da relação entre os dois modelos é que Gagné construiu empiricamente uma sequência instrucional ótima e Sweller forneceu a arquitetura cognitiva que explica por que ela é ótima⁵¹⁵. Os eventos de instrução respeitam a capacidade da memória de trabalho não porque Gagné estava trabalhando com esse construto, que não estava completamente desenvolvido na sua época, mas porque ele destilou a observação de instrução eficaz em padrões que, por necessidade, eram aqueles que o sistema cognitivo humano conseguia processar.

9.6.2. Sweller e Bloom: A Taxonomia como Gradiente de Carga

A taxonomia de Bloom organiza objetivos educacionais em hierarquia de complexidade crescente, de Lembrar a Criar. A Teoria da Carga Cognitiva oferece uma leitura precisa de por que essa hierarquia existe e por que deve ser respeitada na sequência instrucional.

Objetivos no nível de Lembrar podem ser atingidos com processamento que envolve poucos elementos e baixa interatividade entre eles: memorizar um fato não exige a coordenação simultânea de múltiplos esquemas. À medida que se sobe na taxonomia, a complexidade cognitiva aumenta de forma não-linear. Aplicar um conceito a uma situação nova exige não apenas o conceito, mas a identificação das características da situação que tornam o conceito aplicável, a seleção entre múltiplas aplicações possíveis, e a execução do procedimento correto, tudo em simultâneo. Analisar e Avaliar exigem mais ainda: esquemas múltiplos

¹⁵ GAGNÉ, Robert M. *The Conditions of Learning and Theory of Instruction*. 4. ed. Holt, Rinehart and Winston, 1985.

precisam ser ativados e colocados em relação, com seus pontos de tensão identificados, e um julgamento de qualidade produzido.

A implicação é direta: progredir na taxonomia de Bloom sem garantir que os esquemas dos níveis anteriores estão suficientemente consolidados para operar de forma automatizada é impor ao aprendiz a soma de cargas de múltiplos níveis simultaneamente. O aprendiz que ainda não automatizou a aplicação de um conceito não tem capacidade disponível para analisá-lo criticamente. Bloom descreveu o destino da progressão. Sweller descreve as condições sob as quais a progressão é possível ⁵¹⁶.

9.6.3. Sweller e Glasser: Sobrecarga como Mecanismo da Não-retenção

O modelo de retenção de Glasser, que descreve como diferentes modalidades de engajamento produzem diferentes níveis de retenção, é uma descrição fenomenológica do que a teoria da carga cognitiva explica mecanisticamente.

Glasser observou que modalidades passivas, ouvir e ler, produzem retenção substancialmente menor do que as ativas, demonstrar e ensinar. A teoria da carga cognitiva fornece a explicação: formas passivas de engajamento não ativam os processos de recuperação e geração que são responsáveis pela consolidação de esquemas. Um aprendiz que ouve uma explicação processa a informação como input para a memória de trabalho, mas sem o processamento de recuperação que transfere representações para a memória de longo prazo de forma robusta. A modalidade ativa, especialmente ensinar, impõe carga germinativa alta porque força a organização e externalização do conhecimento, o que consolida esquemas de forma qualitativa que o processamento passivo não alcança.

A sobreposição entre os dois modelos vai além: quando Glasser ob-

¹⁶ BLOOM, Benjamin S. et al. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. David McKay Company, 1956.

serva que modalidades ativas produzem maior retenção, está, no vocabulário de Sweller, descrevendo contextos em que a carga germinativa é alta e a extrínseca está sob controle. Quando o sistema satura, seja por material excessivamente denso, seja por design instrucional que gera carga extrínseca elevada, o espaço disponível para carga germinativa colapsa, e mesmo as modalidades ativas tornam-se ineficazes. A saturação cognitiva não é apenas um problema de quantidade de informação: é o mecanismo pelo qual o processo de formação de esquemas é sistematicamente interrompido antes de produzir representações duráveis. Identificado o mecanismo, a pergunta que se coloca é de autoria: quem produz o material que impõe esse colapso, e por que não consegue perceber que o está fazendo.

PARTE III

Consciencia, Persuasao, Decisao

10

O MAPA DA MENTE DO COMPRADOR: EUGENE SCHWARTZ E OS NÍVEIS DE CONSCIÊNCIA

10.1. CONTEXTO HISTÓRICO

Nova York, 1960. A cidade carregava a aceleração do pós-guerra, o dinheiro circulava em velocidade que as décadas anteriores não tinham conhecido, e Madison Avenue havia se tornado sinônimo de uma ciência nova, ambígua, tecnicamente imprecisa e economicamente decisiva: a arte de fazer alguém querer comprar algo que, até segundos atrás, não sabia que precisava. A agência de publicidade funcionava menos como escritório criativo e mais como laboratório improvisado de psicologia aplicada. Os experimentos não tinham grupos de controle. Os resultados eram medidos em cartas de resposta, assinaturas canceladas e cheques enviados pelo correio.

Eugene Schwartz entrou nesse ambiente como menino de recados,

formou-se em jornalismo na Universidade de Washington e aprendeu a profissão na prática de fazer e errar. Não era acadêmico, e esse fato importa: o que o tornaria um dos copywriters mais bem pagos do século ¹ não foi formação teórica, mas a capacidade de observar sistematicamente por que certos textos funcionavam para certos públicos e não funcionavam para outros, e de extrair dessa observação algo que se aproximava de uma teoria.

A pressão que moldou esse pensamento veio de uma transformação técnica do setor. O crescimento do marketing de resposta direta por mala-direta havia criado uma condição inédita: mensurabilidade. Diferente da publicidade em massa dos anos 1950, que operava em televisão e rádio com critérios de sucesso nebulosos, o marketing de resposta direta era contável. Você sabia exatamente quantas cartas foram enviadas, quantas respostas chegaram, quanto cada resposta havia custado. Esse ambiente de dados empíricos brutos exercia pressão seletiva implacável: os que não produziam resultados consistentes eram rapidamente descartados ². Schwartz sobreviveu e prosperou. O mapa que ele construiu nasceu da necessidade de entender por que ele ganhava onde outros perdiam.

A tensão central que o cercava era esta: os profissionais da época acreditavam que a eficácia de um anúncio dependia da qualidade do produto combinada com a qualidade da cópia. Produto bom, cópia boa, vendas acontecem. Era uma lógica causal simples, elegante e amplamente equivocada. O produto excelente, apresentado com mensagem excelente, para o público errado, fracassava. O produto mediano, com mensagem mediana, para o público certo, vendia. Esse padrão se repetia com regularidade suficiente para exigir explicação, e a explicação não estava na cópia nem no produto. Estava em algo anterior a ambos: o estado

¹ SCHWARTZ, Eugene M. *Breakthrough Advertising: How to Write Ads That Shatter Traditions and Sales Records*. Boardroom Classics, 2004. (Edição comemorativa; original publicado em 1966.)

² CAPLES, John. *Tested Advertising Methods*. 5. ed. Prentice Hall, 1997.

mental do comprador no momento de recepção da mensagem.

10.2. O PROBLEMA QUE SCHWARTZ ESTAVA RESOLVENDO

Identificar o estado mental do comprador como variável decisiva era o primeiro passo. O segundo, mais difícil, era sistematizá-lo em algo operacionalizável.

O impasse técnico era o seguinte: mensagens bem escritas falhavam de modo sistemático quando direcionadas para pessoas que não estavam, mentalmente, no lugar certo para recebê-las. Um anúncio brilhante para um tratamento de saúde não convertia quando enviado para pessoas que ainda não sabiam que tinham o problema que o tratamento resolvia. Uma oferta irresistível para uma solução de produtividade não funcionava quando enviada para pessoas que já conheciam aquela solução e a tinham testado sem resultado. A mensagem certa chegava ao lugar errado e morria ali, sem efeito ³.

O problema filosófico subjacente era mais perturbador: a maioria dos profissionais da época tratava o comprador como um estado fixo. Alguém compra ou não compra. Está interessado ou não está. Tem o perfil ou não tem. Essa categorização binária tornava invisível o mecanismo real de decisão: o comprador não era um estado, era um processo. Ele se movia ao longo de uma trajetória cognitiva, e em cada ponto dessa trajetória precisava de uma mensagem diferente. Enviar a mensagem correta para o ponto errado não era apenas ineficaz: era ativamente prejudicial. Criava confusão, frustração e, eventualmente, rejeição permanente da marca ou da categoria de solução.

Havia ainda um conflito latente com a tradição retórica ocidental. De

³ SUGARMAN, Joseph. *The Adweek Copywriting Handbook: The Ultimate Guide to Writing Powerful Advertising and Marketing Copy from One of America's Top Copywriters*. John Wiley & Sons, 2007.

Aristóteles a Cícero, a retórica clássica assumia que o orador conhecia o estado do auditório e adaptava o discurso a ele. Mas essa adaptação era intuitiva, baseada em leitura de situação, não em taxonomia explícita ⁴. Schwartz precisava de algo mais operacional: um sistema ensinável, replicável e aplicável em escala por copywriters menos experientes do que ele. A intuição funciona para o gênio individual. Um sistema funciona para uma indústria.

O problema, portanto, era construir uma taxonomia do estado mental do comprador que fosse precisa o suficiente para prescrever a mensagem correta e simples o suficiente para ser ensinada. “Breakthrough Advertising” foi a resposta.

10.3. IDEIA CENTRAL

A contribuição central de Schwartz pode ser enunciada com precisão: o comprador não é uma categoria, é uma posição em uma sequência de estados cognitivos, e cada estado requer uma abordagem de comunicação radicalmente diferente. A esse mapa de estados ele deu o nome de “níveis de consciência do mercado”, e a obra que sistematizou essa teoria, publicada em 1966, é considerada até hoje a referência seminal na área ⁵.

Cinco níveis compõem a taxonomia.

No primeiro, o comprador é não consciente (“Unaware”): não sabe que tem um problema, não reconhece o sintoma, não nomeia a dor, não conecta seu estado atual com nenhuma categoria de solução disponível. Comunicar uma solução para alguém que não reconhece o problema é, em termos de eficácia, equivalente a falar em uma língua que o interlo-

⁴ ARISTOTLE. *Rhetoric*. Traduzida por W. Rhys Roberts. Oxford University Press, 1924.

⁵ SCHWARTZ, Eugene M. *Breakthrough Advertising*. Op. cit., pp. 6–35.

cutor não compreende. A mensagem eficaz para esse nível não menciona produto nem solução: cria reconhecimento do problema, desperta consciência de um estado que a pessoa vive sem conseguir nomear. O mecanismo é mais próximo da narrativa do que da persuasão direta. A identificação precede a abertura para a oferta.

No segundo nível, o comprador é consciente do problema (“Problem Aware”): sabe que tem uma dor, pode descrevê-la com alguma precisão, mas não sabe que existe solução. Não procura porque não acredita que existe o que procurar. A comunicação eficaz reconhece e valida a dor, criando identificação imediata, e introduz a existência de uma categoria de solução. Não o produto específico: a possibilidade de solução. Apresentar o produto antes de estabelecer essa possibilidade é prematuro e cria resistência.

No terceiro nível, o comprador é consciente da solução (“Solution Aware”): sabe que o problema tem saída, mas não conhece o produto específico que está sendo oferecido. Conhece a categoria, talvez tenha experimentado outras abordagens, tem expectativas formadas e provavelmente carrega ceticismo em relação a promessas que não se realizaram. A mensagem eficaz não argumenta pela existência da solução, mas pela diferenciação: por que este produto, esta abordagem, este mecanismo específico é diferente do que você já tentou? O tom muda. A evidência técnica passa a ter mais peso do que o apelo emocional.

No quarto, o comprador é consciente do produto (“Product Aware”): conhece o produto, pesquisou, considerou a compra, mas não está convencido. Tem objeções latentes ou explícitas. A mensagem eficaz trabalha objeção por objeção, antecipa resistências, apresenta garantias e evidências sociais. A estrutura argumentativa se aproxima do debate formal: os contra-argumentos são conhecidos e endereçados diretamente.

No quinto, “Most Aware”, o comprador quer comprar. Tomou a deci-

são internamente. Conhece o produto, confia na solução, só precisa das condições: preço, facilidade de acesso, um empurrão final. Copywriting longo e elaborado aqui não é apenas desnecessário: é contraproducente, porque cria fricção no exato momento em que o comprador já estava pronto.

A implicação sistêmica do modelo vai além da descrição. É prescritiva e, em certo sentido, punitiva quando ignorada. Enviar uma oferta de compra para alguém no nível um gera rejeição e associação negativa com a marca. Apresentar diferenciação técnica para alguém no nível dois produz confusão. A mensagem certa no nível errado é, operacionalmente, pior do que nenhuma mensagem⁶. Esse é o argumento definitivo de Schwartz, e é o que distingue sua taxonomia de um simples funil de marketing: o modelo não apenas localiza o comprador, mas especifica o custo de ignorar essa localização.

10.4. A EVOLUÇÃO POSTERIOR DA IDEIA

“Breakthrough Advertising” foi publicado em tiragem limitada e por décadas circulou apenas entre os profissionais que o buscavam ativamente. Não havia versão digital, o livro raramente aparecia em livrarias convencionais, e seu prestígio se transmitia de forma quase oral na comunidade de copywriters. Uma edição comemorativa foi publicada em 2004, mas a difusão ampla viria por um caminho que Schwartz não poderia ter antecipado: a internet.

O crescimento do marketing digital nas décadas de 1990 e 2000 gerou demanda massiva por frameworks que explicassem por que certas campanhas funcionavam e outras não. O modelo de funil AIDA, formulado

⁶ KENNEDY, Dan S. *The Ultimate Sales Letter: Attract New Customers, Boost Your Sales*. 4. ed. Adams Media, 2011.

por Elias St. Elmo Lewis no final do século XIX ⁷, havia sido amplamente adotado, mas era vago o suficiente para ser inútil na prática. Os profissionais que encontravam Schwartz percebiam nele algo que o AIDA não oferecia: especificidade operacional. Os cinco níveis diziam exatamente onde o comprador estava e, conseqüentemente, o que fazer.

A fundação da HubSpot por Brian Halligan e Dharmesh Shah ⁸ operacionalizou uma versão adaptada do modelo em escala digital. A premissa central do inbound marketing, de que a marca deve produzir conteúdo calibrado ao estágio do comprador na jornada de decisão, é essencialmente uma aplicação da taxonomia schwartzeana. O conteúdo de topo de funil trabalha os níveis um e dois; o de meio de funil, os níveis três e quatro; o de fundo de funil, o nível cinco. A linguagem mudou. A estrutura subjacente é a mesma.

O conceito de “Jobs to be Done”, desenvolvido por Clayton Christensen ⁹, oferece perspectiva complementar. Schwartz mapeia o estado cognitivo do comprador em relação à existência do produto e da solução; Christensen mapeia a motivação funcional, emocional e social que move o comprador a agir. As abordagens não são incompatíveis: o nível de consciência descreve onde a pessoa está no processo; o jobs-to-be-done descreve o que a impulsiona a progredir. Usados em conjunto, constroem um mapa de decisão mais completo do que cada um oferece isoladamente.

A crítica mais estruturada ao modelo vem da psicologia cognitiva contemporânea. Daniel Kahneman demonstrou que a decisão de compra raramente obedece a uma progressão linear por estágios ¹⁰: emoções, vieses cognitivos e o contexto imediato de decisão interferem de forma

⁷ BARRY, Thomas E. The Development of the Hierarchy of Effects: An Historical Perspective. *Current Issues and Research in Advertising*, v. 10, n. 1-2, p. 251–295, 1987.

⁸ HALLIGAN, Brian; SHAH, Dharmesh. *Inbound Marketing: Attract, Engage, and Delight Customers Online*. Edição revisada. John Wiley & Sons, 2014.

⁹ CHRISTENSEN, Clayton M.; HALL, Taddy; DILLON, Karen; DUNCAN, David S. *Competing Against Luck: The Story of Innovation and Customer Choice*. HarperBusiness, 2016.

¹⁰ KAHNEMAN, Daniel. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.

não sequencial. Um comprador no nível quatro, com objeções racionais aparentemente mapeadas, pode ser movido à compra não por um argumento que as resolve, mas por uma mudança no contexto emocional, como um depoimento que cria identificação ou uma perda iminente que cria urgência. O modelo de Schwartz descreve a estrutura cognitiva, mas não captura plenamente a dinâmica emocional do processo. É uma simplificação operacionalmente útil, não uma teoria completa do comportamento de compra. Essa limitação não invalida o mapa: define seu escopo de aplicação.

10.5. IMPACTO REAL NO MUNDO

A evidência do impacto de Schwartz é difusa precisamente porque a ideia foi absorvida de forma tão ampla que se tornou pressuposto invisível de grande parte do marketing contemporâneo. Poucos profissionais citam a fonte porque, em muitos casos, sequer sabem que existe uma.

A estrutura de “topo, meio e fundo de funil” adotada pela esmagadora maioria das equipes de marketing é uma tradução simplificada dos cinco níveis. A diferença é relevante: Schwartz fornecia precisão conceitual que a linguagem do funil perdeu na tradução. O que exatamente distingue alguém de “topo de funil” de alguém de “meio de funil”? A taxonomia schwartzeana responde com especificidade: o estado cognitivo em relação à existência do problema e à existência de solução. Empresas que atingiram valuations acima de um bilhão de dólares em tecnologia de consumo, como HubSpot, Salesforce e Mailchimp, construíram estratégias de aquisição que segmentam comunicação por estágio de consciência de forma explícita ¹¹.

¹¹ SHAH, Dharmesh; HALLIGAN, Brian. *Inbound Marketing*. Op. cit., cap. 4–6.

Pesquisas sobre abandono de funil conduzidas pelo MarketingSherpa em 2012 e 2015 identificaram que a principal causa de baixa conversão em campanhas B2B não era qualidade do produto nem preço: era timing inadequado da mensagem, especificamente o envio de oferta de venda para leads ainda em estágios iniciais de consciência ¹². A quantificação específica varia por setor, mas a direção é consistente: ignorar os níveis de consciência é a causa mais frequente de fracasso em lançamentos com produto e cópia tecnicamente competentes. O problema não estava onde os profissionais procuravam.

O blog como instrumento de marketing, o e-mail de nutrição como sequência de educação progressiva, a página de vendas como documento de conversão para o comprador no nível cinco: cada um desses formatos corresponde a um nível específico do mapa. Toda estratégia de conteúdo educativo contemporânea, mesmo quando seus praticantes desconhecem Schwartz, é schwartzeana na arquitetura.

10.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

A convergência entre Schwartz e Neil Rackham é, operacionalmente, a mais rica para quem trabalha com vendas consultivas. O SPIN Selling, sistematizado por Rackham a partir de doze anos de pesquisa em mais de 35.000 chamadas de venda ¹³, funciona como um protocolo de entrevista que move o comprador ao longo dos níveis de consciência em tempo real.

A pergunta de Situação diagnostica onde o comprador está. A pergunta de Problema cria ou amplia consciência da dor, movendo do nível um para o dois. A pergunta de Implicação aprofunda a percepção de ur-

¹² MARKETING SHERPA. *2012 Lead Generation Benchmark Report*. MECLABS Institute, 2012.

¹³ RACKHAM, Neil. *SPIN Selling*. McGraw-Hill, 1988.

gência, trabalhando dentro do nível dois. A pergunta de Necessidade de Solução move o comprador para o nível três ao criar antecipação de um estado melhor possível, antes de qualquer menção ao produto. O SPIN, nessa leitura, é uma aplicação conversacional do modelo de Schwartz: o vendedor que conduz as perguntas está diagnosticando o nível e aplicando as intervenções sequenciais corretas para mover o comprador. Dois frameworks desenvolvidos independentemente, em contextos distintos, descrevendo o mesmo mecanismo.

A conexão com Lincoln Murphy introduz uma consequência que o próprio Schwartz não explorou: o que acontece depois da venda quando o comprador fechou antes de completar o processo de consciência. Murphy demonstrou que o cliente que compra sem entender claramente o que o produto pode entregar chega ao onboarding com gaps de expectativa que o customer success não consegue preencher retrospectivamente ¹⁴. A venda que aparentemente “ganhou” ao comprimir etapas no mapa de consciência é, com frequência, o cliente que cancela no segundo mês. O nível de consciência no momento da compra determina, em medida significativa, o LTV futuro. Essa é a extensão natural do argumento de Schwartz: o mapa não descreve apenas o processo de aquisição, mas o pré-requisito cognitivo para a entrega de valor.

A convergência com Wiggins e McTighe translada a lógica para o design instrucional. O modelo de “Backward Design” ¹⁵ propõe que o currículo parta do resultado desejado em direção ao conteúdo. O que o modelo não explicita, mas o framework de Schwartz completa, é que o resultado desejado só pode ser comunicado com eficácia se calibrado ao nível de consciência do aprendiz no momento de ingresso. Um curso sobre gestão financeira pessoal que assume que o aluno já reconhece sua situação como um problema estrutural perde os participantes do nível

¹⁴ MURPHY, Lincoln. *Customer Success: How Innovative Companies Are Reducing Churn and Growing Recurring Revenue*. Wiley, 2016.

¹⁵ WIGGINS, Grant; MCTIGHE, Jay. *Understanding by Design*. 2. ed. ASCD, 2005.

um: aqueles que sentem que algo não está bem mas ainda não conseguem nomear por quê. A clareza sobre onde se quer chegar não substitui o diagnóstico de onde o aluno começa.

11

A ARTE DE PERGUNTAR PARA VENDER: NEIL RACKHAM E O SPIN SELLING

11.1. CONTEXTO HISTÓRICO

Há uma estranheza particular na origem da pesquisa de Neil Rackham: o maior estudo empírico já conduzido sobre comportamento de vendas foi financiado, em parte, por empresas que esperavam confirmar o que já acreditavam. Não foi o que aconteceu.

Era o início da década de 1970. O treinamento de vendas havia se tornado uma indústria considerável nos Estados Unidos e na Europa, ancorado em técnicas que remontavam ao início do século XX: criar rapport, identificar objeções, pressionar por fechamento, usar perguntas fechadas para conduzir o prospect a um “sim” acumulado. O manual era relativamente estável, e a premissa que o sustentava também. Vender era uma habilidade de persuasão, e persuasão era uma habilidade de persuasão controlada. Com o estímulo certo, na intensidade certa, o prospect

cederia ¹.

Rackham chegou a esse campo por uma rota incomum. Pesquisador comportamental com formação em psicologia experimental na Universidade de Sheffield, foi contratado pela Huthwaite International para conduzir um estudo sobre o que, exatamente, diferenciava vendedores de alto desempenho dos demais. A pergunta parecia óbvia. A resposta, não era.

O que se seguiu foi um programa de pesquisa em escala sem precedentes. Ao longo de doze anos, a equipe da Huthwaite observou mais de 35.000 chamadas de venda em 23 países, codificando comportamentos verbais em tempo real, registrando quais técnicas eram empregadas em quais momentos e correlacionando essas escolhas com os resultados obtidos ². O protocolo era rigoroso para os padrões da pesquisa comportamental aplicada: observadores treinados acompanhavam reuniões ao vivo ou gravadas, categorizavam cada comportamento do vendedor em esquemas de codificação pré-definidos e só marcavam uma venda como “bem-sucedida” quando havia evidência objetiva de progresso ou fechamento, não quando o próprio vendedor avaliava a reunião como positiva.

O rigor importava sobretudo porque o mercado que circundava a pesquisa havia mudado em uma direção que os patrocinadores do estudo não queriam ver. A profissionalização das vendas B2B havia acelerado nos anos 1960 com o crescimento das empresas de tecnologia, serviços financeiros e consultoria: setores nos quais o ciclo de venda durava meses, envolvia múltiplos decisores e dependia de confiança técnica que o modelo clássico de persuasão não conseguia sustentar. IBM, Xerox, NCR e as grandes empresas de equipamentos corporativos da época investiam quantias crescentes em treinamento de vendedores. Os treinamentos

¹ STRONG, Edward K. *The Psychology of Selling and Advertising*. McGraw-Hill, 1925.

² RACKHAM, Neil. *SPIN Selling*. McGraw-Hill, 1988, pp. 1–21.

não estavam funcionando³. O mercado havia mudado estruturalmente; as premissas de venda, não.

11.2. O PROBLEMA QUE RACKHAM ESTAVA RESOLVENDO

A primeira descoberta da pesquisa foi inconveniente para todos que a financiaram.

As técnicas tradicionais funcionavam. Mas apenas em um tipo específico de transação: baixo valor, ciclo curto, decisão individual. Em vendas de alto valor, com ciclos longos e múltiplos decisores envolvidos, as mesmas técnicas que aumentavam conversão em vendas simples reduziam ativamente as chances de sucesso⁴. O problema não era falta de treinamento. Era excesso do treinamento errado.

A lógica subjacente ao paradoxo era precisa. Nas técnicas clássicas de fechamento, a pressão funcionava porque a decisão de compra era reversível, o valor em jogo era baixo e o prospect, ao resistir inicialmente, estava muitas vezes apenas testando o vendedor. Empurrado na direção certa, comprava. Em uma venda de consultoria de seis dígitos, ou em uma implementação de software que comprometeria toda a operação da empresa por anos, a pressão não superava a resistência: a atirava. O prospect conduzido para um fechamento antes de estar genuinamente convencido não fechava; recuava, adicionava aprovações, solicitava mais reuniões, desaparecia. Ou fechava e arrependia-se, o que, em termos de valor de longo prazo, era o pior resultado possível.

A “ilusão do bom vendedor” era o nome implícito da pesquisa para um conjunto de habilidades funcionalmente prejudiciais em contextos específicos. Rapport excessivo criava suspeita. Perguntas fechadas acu-

³ RACKHAM, Neil; DeVINCENTIS, John. *Rethinking the Sales Force: Redefining Selling to Create and Capture Customer Value*. McGraw-Hill, 1999.

⁴ RACKHAM, Neil. *SPIN Selling*. Op. cit., pp. 23–48.

mulativas, projetadas para produzir uma série de “sins” que conduzissem ao fechamento, eram percebidas como manipulação por compradores sofisticados ⁵. O fechamento prematuro sinalizava que o vendedor estava mais interessado na comissão do que no problema do cliente. Em compras de alto valor, onde o comprador assumia risco real e visível, essa percepção destruía a transação mais rapidamente do que qualquer objeção racional.

O impasse teórico era preciso: como conduzir uma conversa que produzisse, no comprador, convicção genuína de que a compra era do seu próprio interesse, sem que essa convicção fosse plantada por pressão ou manipulação? A literatura disponível não oferecia resposta. Os modelos dominantes tratavam venda como persuasão unidirecional: o vendedor convencia o comprador. A inversão que Rackham buscava, um modelo no qual o vendedor estruturava condições para que o comprador se convencesse, não tinha precedente formalizado no campo ⁶. Era preciso construí-lo a partir dos dados.

11.3. IDEIA CENTRAL

A resposta de Rackham não foi uma técnica. Foi uma taxonomia do comportamento verbal do vendedor correlacionada empiricamente com resultados em vendas complexas.

O SPIN não é um acrônimo de marketing. É a descrição de quatro categorias de perguntas que os vendedores mais eficazes em vendas de alto valor faziam de forma consistente, identificadas pela análise dos dados comportamentais coletados ao longo de doze anos. A sequência observada nos vendedores de alto desempenho seguia uma progressão especí-

⁵ CIALDINI, Robert B. *Influence: The Psychology of Persuasion*. Harper Business, 1984.

⁶ RACKHAM, Neil. *SPIN Selling*. Op. cit., pp. 49–72.

fica: Situação, Problema, Implicação, Necessidade de Pagamento ⁷.

Perguntas de Situação estabelecem o contexto do comprador: quantos funcionários, qual o processo atual, quais sistemas estão em uso. São perguntas de coleta de informação, e a tentação é fazer muitas delas. A pesquisa mostrou o inverso: há correlação negativa entre o número de perguntas de Situação e o sucesso da venda. Compradores sofisticados percebem o vendedor que as multiplica como alguém que não fez o dever de casa. A recomendação é que sejam mínimas e cirúrgicas, limitadas ao que não pode ser obtido por preparação prévia ⁸.

Perguntas de Problema revelam dores: “Qual é o maior problema que você enfrenta com o processo atual?” ou “Existe alguma parte do sistema que cria dificuldades recorrentes?”. A correlação com sucesso é positiva em vendas de alto valor, embora funcione com menor impacto em vendas simples. A distinção importa porque em transações de baixo valor, identificar o problema é suficiente para motivar a ação. Em transações de alto valor, o comprador já sabe que tem o problema. O problema não é a questão.

O que importa em vendas complexas é a Implicação. Essa é, segundo os dados de Rackham, a categoria mais difícil de formular e a que possui maior correlação com sucesso em grandes vendas ⁹. Perguntas de Implicação expandem as consequências do problema não resolvido: “O que acontece quando esse processo falha?”, “Qual é o impacto disso na equipe de produção?”, “Se esse problema continuar pelos próximos dois anos, o que muda no custo operacional?”. O mecanismo não é manipulação: é ampliação genuína da perspectiva do comprador sobre algo que ele provavelmente subestimava. O problema que parecia gerenciável torna-se, quando suas implicações são articuladas, urgente.

A sequência culmina na Necessidade de Pagamento, ou Need-Payoff:

⁷ Ibid., pp. 73–110.

⁸ Ibid., pp. 73–82.

⁹ Ibid., pp. 95–110.

perguntas que convidam o comprador a articular, com suas próprias palavras, o valor da solução. “Se você conseguisse resolver esse problema de fluxo, o que isso mudaria para a equipe?” “Quanto você estima que esse retrabalho custa por mês?” A inversão é decisiva: o vendedor não diz ao comprador qual é o valor do produto. O comprador diz ao vendedor. Quando o comprador articula o valor com suas próprias palavras, a resistência à compra cai porque não há nada contra o qual resistir. O argumento de compra veio do próprio comprador.

Isso não é trivial.

O modelo clássico pressupõe que o comprador resiste porque não viu valor suficiente, e a tarefa do vendedor é demonstrá-lo com argumentos convincentes. O modelo SPIN pressupõe algo diferente: o comprador tem os componentes cognitivos para chegar ao valor por conta própria, mas opera com perspectiva incompleta sobre as implicações do problema. A tarefa do vendedor não é demonstrar valor: é criar as condições conversacionais para que o comprador chegue a ele por indução própria. A convicção resultante é qualitativamente diferente. Interna, não externa. Resistente a objeções posteriores porque o comprador não está respondendo a um argumento alheio: está operando a partir de uma conclusão que considera sua ¹⁰.

A pesquisa documentou também o que os vendedores de alto desempenho não faziam: não multiplicavam perguntas fechadas em sequência, não usavam técnicas de fechamento agressivo, não respondiam a objeções com contrapontos imediatos, mas com perguntas que aprofundavam a objeção. E não apresentavam características do produto antes de o comprador ter articulado necessidades explícitas. Apresentar benefícios antes dessa verbalização produzia, consistentemente, mais objeções, não menos ¹¹.

¹⁰ Ibid., pp. 111–130.

¹¹ Ibid., pp. 131–152.

11.4. A EVOLUÇÃO POSTERIOR DA IDEIA

“SPIN Selling” foi publicado em 1988 e encontrou um mercado que já estava, de certa forma, esperando por ele. O crescimento das vendas B2B enterprise havia tornado o modelo clássico de persuasão claramente inadequado. O livro vendeu mais de 150.000 cópias no primeiro ano e eventualmente superaria 500.000, tornando-se um dos textos de vendas mais influentes do século XX ¹².

As primeiras derivações vieram de dentro da própria Huthwaite. Michael Bosworth, treinado na metodologia de Rackham, desenvolveu o “Solution Selling” nos anos 1980, publicado formalmente em 1994 ¹³. A premissa é similar: em vendas de produtos e serviços complexos, o vendedor deve funcionar como diagnóstico antes de apresentar soluções. A diferença está na natureza de cada framework: o SPIN é descritivo, mapeia o que os melhores vendedores fazem; o Solution Selling é prescritivo, define um processo com etapas nomeadas. Os dois coexistiram e se complementaram nos treinamentos corporativos das décadas seguintes.

A evolução mais crítica veio com “The Challenger Sale”, publicado por Matthew Dixon e Brent Adamson em 2011 ¹⁴. A pesquisa da Corporate Executive Board identificou cinco perfis de vendedor e concluiu que o “Challenger”, aquele que instrui, personaliza e assume controle da conversa comercial, superava os demais em vendas complexas. A tensão com Rackham é genuína, embora frequentemente superestimada. O Challenger Sale não refuta o SPIN: propõe que a sequência de perguntas deve ser precedida pela capacidade do vendedor de ensinar o comprador

¹² HUTHWAITE INTERNATIONAL. *25 Years of SPIN: A Retrospective on the Research and Its Impact*. Huthwaite, 2013.

¹³ BOSWORTH, Michael. *Solution Selling: Creating Buyers in Difficult Selling Markets*. McGraw-Hill, 1994.

¹⁴ DIXON, Matthew; ADAMSON, Brent. *The Challenger Sale: Taking Control of the Customer Conversation*. Portfolio/Penguin, 2011.

algo que ele não sabia sobre seu próprio negócio. Provocação intelectual como antecedente da escuta diagnóstica. As duas ideias são sequencialmente compatíveis.

No campo da tecnologia enterprise, o SPIN encontrou sua aplicação mais sistemática nos programas de treinamento de SAP, Oracle e, mais tarde, Salesforce. O ciclo de venda de software enterprise, com múltiplas camadas de aprovação, RFPs elaborados e stakeholders técnicos e executivos, reproduzia quase em laboratório as condições que a pesquisa havia descrito como ambiente natural do SPIN. A pergunta de Implicação tornou-se peça central do playbook enterprise porque é na implicação que o CTO e o CFO encontram linguagens comuns: impacto operacional e impacto financeiro ¹⁵.

A crítica estrutural mais consistente ao SPIN diz respeito à sua dependência de acesso prolongado ao comprador. A proliferação de canais digitais de venda e a compressão dos ciclos em vários setores criaram ambientes nos quais a sequência completa simplesmente não cabe. Em muitos contextos, o comprador chega ao vendedor já nos níveis três ou quatro de consciência (usando a taxonomia de Schwartz), tendo feito a maior parte do diagnóstico por conta própria via pesquisa online. As perguntas de Situação, nesse cenário, são ainda menos toleradas do que Rackham documentou, e as perguntas de Implicação precisam ser reposicionadas para operar sobre um estado de consciência mais avançado. Isso modifica a técnica sem invalidar a lógica que a sustenta.

¹⁵ THULL, Jeff. *Mastering the Complex Sale: How to Compete and Win When the Stakes Are High*. 2. ed. John Wiley & Sons, 2010.

11.5. IMPACTO REAL NO MUNDO

O impacto do SPIN é difícil de quantificar com precisão porque, como ocorreu com o modelo de Schwartz, foi tão amplamente absorvido que grande parte dos profissionais de vendas opera com versões derivadas sem saber da fonte.

Os dados mais robustos vêm dos estudos da própria Huthwaite, o que exige cautela interpretativa. Uma meta-análise de programas de treinamento baseados em SPIN aplicados em empresas como Motorola, Kodak e IBM entre as décadas de 1980 e 1990 registrou incrementos de 17% a 30% em taxas de fechamento em vendas de alto valor após a implementação, com maior impacto nos primeiros doze meses¹⁶. Os números são de fonte interessada, mas a direção é consistente com o que estudos independentes sobre vendas consultivas documentaram em outros contextos.

O impacto mais duradouro talvez não seja estatístico, mas conceitual. O SPIN estabeleceu no campo uma distinção que antes não tinha nome: “necessidades implícitas” versus “necessidades explícitas”. Necessidade implícita é o problema reconhecido mas ainda não articulado como urgência de solução. Necessidade explícita é a necessidade verbalizada pelo próprio comprador como demanda de ação. O papel do vendedor é criar as condições conversacionais para que necessidades implícitas se tornem explícitas. Toda a estrutura do inbound marketing contemporâneo, da jornada do comprador à qualificação de leads, opera sobre essa distinção, frequentemente sem reconhecê-la como herança de Rackham¹⁷.

Uma análise da Gartner de 2019 sobre o comportamento de compradores B2B em ciclos longos revelou que mais de 75% do processo de decisão ocorria antes do primeiro contato com um representante de vendas,

¹⁶ HUTHWAITE INTERNATIONAL. *Measuring the Impact of SPIN Selling Training: A Multi-Company Analysis*. Huthwaite Research Group, 1993.

¹⁷ HALLIGAN, Brian; SHAH, Dharmesh. *Inbound Marketing: Attract, Engage, and Delight Customers Online*. Edição revisada. John Wiley & Sons, 2014.

e que compradores com necessidades já explicitadas fechavam em ciclos mais curtos e com menor resistência a preço ¹⁸. Lido à luz de Rackham, esse dado reposiciona o ponto de entrada da lógica SPIN: a tarefa do vendedor moderno é cada vez menos criar consciência do problema e cada vez mais aprofundar implicações já conhecidas, movendo o comprador da consciência da solução para a articulação do valor específico do que está sendo oferecido.

11.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

A relação entre SPIN e os níveis de consciência de Schwartz não é apenas complementar. É estruturalmente isomórfica.

Schwartz descreve onde o comprador está em um mapa de estados cognitivos. Rackham prescreve as intervenções verbais que o movem ao longo desse mapa em tempo real de conversa. A pergunta de Situação diagnostica o ponto de entrada: estabelece em qual nível da trajetória o comprador se encontra antes de qualquer intervenção. A pergunta de Problema trabalha com o comprador no nível dois, aprofundando a consciência da dor e criando terreno para o próximo movimento. A pergunta de Implicação é a intervenção crítica para quem está no nível dois sem urgência suficiente para agir: não inventa um problema, mas expande as ramificações de um já reconhecido, movendo o comprador do nível dois para o três. A Necessidade de Pagamento opera na transição do nível três para o quatro, criando as condições para que o comprador construa, com as próprias palavras, o caso de compra ¹⁹.

O que Schwartz descreveu como mapa de estados cognitivos, Rackham tornou um protocolo dinâmico de movimento. Juntos, os dois frameworks

¹⁸ GARTNER. *The New B2B Buying Journey and Its Implication for Sales*. Gartner Research Report, 2019.

¹⁹ SCHWARTZ, Eugene M. *Breakthrough Advertising*. Boardroom Classics, 2004. (Original: 1966.)

respondem perguntas que nenhum responde sozinho: Schwartz localiza o comprador; Rackham move.

A convergência com Lincoln Murphy introduz a dimensão pós-venda. O argumento central de Murphy em *Customer Success* é que o cliente que compra sem compreensão genuína do que está contratando cria, desde o primeiro dia, as condições para o cancelamento²⁰. A venda SPIN trabalha estruturalmente contra esse resultado: ao exigir que o comprador articule o valor com suas próprias palavras antes de fechar, produz alinhamento de expectativas que o próprio comprador validou. O cliente que chegou à conclusão de que deveria comprar tem comprometimento qualitativamente diferente do cliente que foi persuadido a comprar. Esse comprometimento determina o comportamento no onboarding, a profundidade de adoção da solução e, eventualmente, a renovação. O SPIN não é apenas técnica de fechamento: é pré-requisito estrutural para o sucesso pós-venda.

A conexão com Michael Polanyi introduz uma tensão produtiva. O argumento de Polanyi sobre conhecimento tácito postula que grande parte do que os especialistas sabem não pode ser completamente articulada em linguagem explícita: é incorporada, situacional, operada por reconhecimento de padrões que resistem à verbalização²¹. Aplicado ao SPIN, isso cria um problema prático preciso: o vendedor que genuinamente domina seu domínio tem conhecimento sobre o problema do cliente que o cliente ainda não consegue articular. A pergunta de Implicação é, nessa leitura, uma tentativa de externalizar o conhecimento tácito do vendedor para o processo de decisão do comprador, sem declaração explícita. “O que acontece com sua operação se esse problema continuar por mais um ano?” não é uma pergunta cuja resposta o vendedor desconhece: é uma pergunta cuja resposta o vendedor conhece melhor do que o comprador,

²⁰ MURPHY, Lincoln. *Customer Success: How Innovative Companies Are Reducing Churn and Growing Recurring Revenue*. Wiley, 2016.

²¹ POLANYI, Michael. *The Tacit Dimension*. Doubleday, 1966.

e cuja função é criar as condições para que o comprador chegue independentemente à mesma conclusão. O conhecimento tácito do especialista é operacionalizado pela estrutura da pergunta, não pela declaração direta. O paradoxo, como o especialista usa o que sabe sem revelar o que sabe mantendo a percepção de que a conclusão pertence ao comprador, encontra aqui sua resolução operacional.

PARTE IV

Retencao, Valor, Expansao

12

O CLIENTE COMO RESULTADO: LINCOLN MURPHY E O CUSTOMER SUCCESS

12.1. CONTEXTO HISTÓRICO

O modelo de negócio SaaS (Software as a Service) chegou ao início dos anos 2000 carregando uma promessa de eficiência que escondia uma armadilha estrutural. Vender não é o mesmo que entregar valor, e o mercado levaria uma geração inteira de empreendedores a aprender isso da forma mais cara possível.

A lógica do software como serviço era sedutora. Em vez de vender licenças perpétuas com ciclo de receita irregular e custos de distribuição elevados, as empresas passariam a cobrar mensalmente por acesso contínuo. O modelo de assinatura resolvia o problema de fluxo de caixa, democratizava o acesso ao software e, ao menos em tese, alinhava os incentivos da empresa com a continuidade do uso pelo cliente. Se o cliente pagava todo mês, a empresa tinha interesse genuíno em que ele

continuasse usando. A lógica era impecável no papel.

Empresas de SaaS começaram a crescer em base de clientes e encolher em receita recorrente líquida. O fenômeno tinha nome: churn. Clientes assinavam, usavam por um período, e cancelavam. Às vezes em volume suficiente para anular completamente o crescimento de novas aquisições. Uma empresa que adquiria 200 novos clientes por mês e perdia 190 por cancelamentos estava, na prática, estagnada, a despeito de investir massivamente em marketing e vendas. O ciclo vicioso era perverso: quanto mais se vendia para compensar o churn, mais capital era imobilizado em aquisição, mais o custo de aquisição subia. E o problema de fundo permanecia intacto ¹.

Lincoln Murphy era um dos consultores que as empresas de SaaS chamavam quando o problema chegava ao limite. Sua formação combinava experiência operacional em empresas de software com uma capacidade analítica incomum para identificar onde, exatamente, o valor estava se perdendo. O diagnóstico que ele articulava ao longo da primeira metade dos anos 2010, consolidado no livro “Customer Success” coescrito com Nick Mehta e Dan Steinman em 2016 ², não era uma crítica ao produto nem às vendas nem ao marketing. Era uma crítica a uma premissa que havia permanecido invisível por tempo demais: a premissa de que a responsabilidade da empresa terminava no momento da venda.

O contexto intelectual do período era dominado por frameworks de growth hacking, funis de aquisição e otimização de conversão. A energia do ecossistema de startups apontava para cima do funil: como atrair mais prospects, como converter mais trials, como reduzir o custo por aquisição. O que acontecia depois que o cliente pagava pela primeira vez recebia atenção marginal, geralmente delegada a equipes de suporte

¹ SKOK, David. SaaS Metrics 2.0: A Guide to Measuring and Improving What Matters. For *Entrepreneurs Blog*, 2013. Disponível em: <https://www.forentrepreneurs.com/saas-metrics-2>.

² MEHTA, Nick; STEINMAN, Dan; MURPHY, Lincoln. *Customer Success: How Innovative Companies Are Reducing Churn and Growing Recurring Revenue*. Wiley, 2016.

cuja missão era reativa, não proativa. Resolver problemas quando o cliente reclamava, não criar condições para que o cliente chegasse aos resultados que justificavam a assinatura ³.

Murphy foi um dos primeiros a nomear isso como falha estrutural, não como ineficiência operacional. A distinção importa: ineficiência pode ser corrigida com processo; falha estrutural exige reconfiguração da premissa.

12.2. O PROBLEMA QUE MURPHY ESTAVA RESOLVENDO

A confusão central que Murphy encontrava nas empresas que consultava era entre dois conceitos que pareciam sinônimos mas não eram: satisfação do cliente e sucesso do cliente.

Satisfação mede a experiência subjetiva da interação. Um cliente pode estar satisfeito com o atendimento, com a interface do produto, com a velocidade de resposta do suporte, e ainda assim cancelar. Porque satisfação não é resultado. E o que retém um cliente em uma relação de assinatura não é a qualidade da experiência de uso: é a percepção de que o produto está contribuindo para um resultado que importa para ele ⁴.

Isso parece óbvio dito assim. No ambiente operacional das empresas de SaaS dos anos 2000, não era.

O diagnóstico mais perturbador de Murphy não era sobre métricas de satisfação. Uma parcela significativa dos cancelamentos não tinha como causa a insatisfação com o produto: tinha como causa o fato de que nunca havia sido definido, desde o início da relação, o que seria “sucesso” para aquele cliente específico. O cliente que assina uma plataforma de automação de marketing sem ter sido conduzido a definir uma

³ BLISSFULLY. *Annual SaaS Trends Report*. Blissfully Research, 2019.

⁴ MEHTA, Nick; STEINMAN, Dan; MURPHY, Lincoln. *Customer Success*. Op. cit., pp. 41–62.

meta concreta de resultado não tem critério para avaliar se o produto está funcionando ou não. Navega no produto, usa algumas funcionalidades, não observa mudança perceptível nos indicadores que importam para ele, porque nunca os definiu como critério de avaliação, e cancela quando o custo da assinatura supera o valor percebido, que era baixo desde o início porque nunca havia sido articulado ⁵.

O impasse era preciso: as empresas de SaaS estavam construindo estruturas inteiras de geração de demanda, vendas e onboarding para entregar o cliente a um estado de uso que não era suficiente para garantir renovação. O produto podia ser excelente, a venda impecável, o onboarding eficiente. O cliente ainda cancelava porque nunca havia sido definido o que, concretamente, o produto precisava fazer para que a assinatura valesse a pena.

12.3. IDEIA CENTRAL

A resposta de Murphy não foi uma métrica ou um processo. Foi uma redefinição do que o produto precisa fazer para merecer renovação, expressa num conceito de precisão técnica que muitas traduções simplificam ao ponto de deformar: o Desired Outcome.

Desired Outcome não é o objetivo do cliente formulado em linguagem vaga. Não é “crescer”, “ser mais produtivo” ou “economizar tempo”. É uma construção de dois componentes com funções distintas e complementares: Required Outcome mais Appropriate Experience ⁶.

O Required Outcome é o resultado que o cliente precisa alcançar para que o produto tenha valido. Não o que o cliente quer vagamente, mas o que ele precisa objetivamente, definido com especificidade suficiente

⁵ Ibid., pp. 63–89.

⁶ Ibid., pp. 90–114.

para ser mensurado. Uma empresa que usa uma plataforma de CRM tem como Required Outcome não “melhorar o relacionamento com clientes” (que é vago demais para ser operacionalizável), mas algo como “reduzir o ciclo de vendas em 20% nos próximos seis meses” ou “aumentar a taxa de conversão de leads qualificados de 12% para 18% no primeiro trimestre”. A especificidade é funcional, não estética: sem ela, não há como avaliar se o produto está cumprindo sua função, e sem essa avaliação, a renovação depende de fatores que a empresa não controla.

O Appropriate Experience é o componente frequentemente negligenciado, e é aqui que a formulação de Murphy rompe com os frameworks anteriores. Mesmo que o cliente alcance o Required Outcome, a renovação pode falhar se a experiência durante o processo foi inadequada para o perfil desse cliente. Uma empresa enterprise que alcança os resultados técnicos esperados de uma plataforma, mas cujo time foi forçado a lidar com um processo de implementação excessivamente técnico sem suporte adequado, pode decidir não renovar por razões que têm a ver com o custo interno de manutenção da relação, não com a eficácia do produto. A experiência adequada não é a experiência máxima possível (o que seria custo proibitivo): é a experiência que esse perfil de cliente específico precisa ter para perceber o Required Outcome como valioso e a relação como sustentável ⁷.

A combinação dos dois componentes gera o comportamento que o modelo de receita recorrente precisa para funcionar: o cliente que alcança seu Desired Outcome não apenas renova. Expande. Recomenda. Compra mais funcionalidades, mais assentos, mais módulos ⁸. O churn cai não porque o suporte melhorou, mas porque o produto passou a ser percebido como componente de resultado, não como custo de funcionamento.

⁷ Ibid., pp. 115–138.

⁸ MURPHY, Lincoln. *Customer Success: The Definitive Guide*. Gainsight, 2014. (White paper.)

A implicação para o design é profunda. Em vez de otimizar para engajamento, quantas sessões por semana, quantas funcionalidades usadas, qual o tempo médio de sessão, o produto deve ser otimizado para resultado definido pelo Desired Outcome do cliente, não pelas métricas internas que a empresa escolheu monitorar. Essa inversão tem consequências em toda a organização: no produto, nas métricas de sucesso, na estrutura de onboarding, no modelo de contrato, na forma como a equipe de vendas qualifica prospects.

O instrumento operacional que traduz essa lógica em gestão contínua é o Customer Health Score: uma composição ponderada de indicadores que permite estimar, em tempo real, se um cliente está ou não no caminho do Desired Outcome. O Health Score não mede satisfação: mede progresso em direção ao resultado definido. Um cliente com score de saúde deteriorando não é necessariamente um cliente insatisfeito. É um cliente em risco de não alcançar o Required Outcome, aproximando-se de um cancelamento que pode ainda não ser consciente. A intervenção proativa, antes que o cliente identifique o problema, é o mecanismo central do Customer Success como função estratégica ⁹.

12.4. A EVOLUÇÃO POSTERIOR DA IDEIA

O impacto do framework de Murphy foi amplificado pelo momento em que chegou. As empresas de SaaS estavam, entre 2012 e 2016, atingindo escala suficiente para que o problema do churn se tornasse existencial. Uma empresa com US\$10 milhões de receita recorrente e churn anual de 20% precisava gerar US\$2 milhões em novas assinaturas apenas para manter o patamar atual. Com US\$100 milhões de receita recorrente e o

⁹ MEHTA, Nick; STEINMAN, Dan; MURPHY, Lincoln. *Customer Success*. Op. cit., pp. 139–165.

mesmo percentual de churn, o problema multiplicava por dez. A aritmética era implacável.

O Customer Success migrou de conceito para função organizacional com velocidade impressionante. Empresas como Salesforce, HubSpot, Zendesk e Gainsight (que se tornaria a plataforma de referência de CS) estruturaram departamentos de Customer Success com orçamentos comparáveis aos de vendas ¹⁰. O cargo de Chief Customer Officer, praticamente inexistente antes de 2010, passou a aparecer nos organogramas das empresas de SaaS mais bem capitalizadas. Numa economia de assinaturas, a receita de hoje depende da performance de CS de ontem, e a receita de amanhã depende da performance de CS de hoje.

O movimento de Product-Led Growth (PLG), desenvolvido pelos fundadores da OpenView Ventures e popularizado por Wes Bush ¹¹ ao longo da segunda metade dos anos 2010, pode ser lido como uma tentativa de operacionalizar o Customer Success dentro do próprio produto, sem depender de interação humana. A premissa do PLG é que o produto deve ser capaz de conduzir o usuário ao momento de valor (o chamado “aha moment”) de forma autônoma, instrumentando o onboarding, a ativação e a expansão sem necessidade de equipe dedicada. Em startups de baixo ticket com base massiva de usuários, isso faz sentido econômico. A tensão com o framework de Murphy é real: PLG tende a otimizar para comportamento de produto (uso, adoção de funcionalidades), enquanto Murphy argumenta que o que importa é resultado externo ao produto. Um produto bem adotado que não gera o Required Outcome do cliente ainda produz churn. Engajamento não é resultado.

A migração do conceito para o campo da educação é particularmente relevante para a APE. O paralelo entre customer success e learning success foi articulado por vários autores de instructional design a partir de

¹⁰ GAINSIGHT. *The State of Customer Success*. Gainsight Research, 2018.

¹¹ BUSH, Wes. *Product-Led Growth: How to Build a Product That Sells Itself*. Product-Led Institute, 2019.

meados dos anos 2010. A taxa de conclusão de cursos online, que é o equivalente educacional do engajamento de produto, mostrou-se um proxy inadequado de sucesso do aluno, da mesma forma que o uso de funcionalidades é um proxy inadequado de resultado do cliente ¹². Um aluno que conclui 100% de um curso mas não aplica o que aprendeu, não muda comportamento, não avança em direção ao resultado que motivou a matrícula, não teve sucesso educacional. Teve conclusão de conteúdo. A distinção é estruturalmente análoga à que Murphy traçou entre satisfação e sucesso, e tem as mesmas implicações para o design do produto, seja ele de software ou pedagógico.

12.5. IMPACTO REAL NO MUNDO

A quantificação do impacto do Customer Success no modelo de negócio SaaS passou, ao longo da segunda metade dos anos 2010, de argumento conceitual para dado mensurável.

Uma análise da Gainsight de 2018 ¹³, baseada em dados de mais de 500 empresas de SaaS, demonstrou que empresas com programas de Customer Success estruturados apresentavam churn anual médio de 5%, comparado com 14% em empresas sem programa equivalente. A diferença se traduzia diretamente em múltiplo de valuation: empresas com churn baixo eram avaliadas, em média, a múltiplos 3 a 4 vezes maiores do que empresas com churn elevado no mesmo segmento e nível de receita. O raciocínio dos investidores era simples: receita recorrente com baixo churn é previsível, e previsibilidade tem valor de mercado.

O caso da Salesforce é emblemático. A empresa foi uma das primei-

¹² MEANS, Barbara; TOYAMA, Yukie; MURPHY, Robert; BAKI, Marianne; JONES, Karla. *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*. U.S. Department of Education, 2010.

¹³ GAINSIGHT. *The State of Customer Success*. Op. cit.

ras a estruturar uma função formal de Customer Success, já no início dos anos 2000, quando o modelo SaaS era ainda uma aposta de alto risco. Benioff argumentava que o sucesso da Salesforce dependia do sucesso de seus clientes: se os clientes crescessem, a Salesforce cresceria com eles¹⁴. Essa premissa tinha implicações operacionais concretas: equipes de CS com metas de expansão de receita, não apenas de retenção; contratos vinculados à demonstração de resultado; sistemas de health score para intervenção proativa antes que o cliente chegasse à decisão de cancelamento. Uma empresa que estruturou a função de CS quando o modelo SaaS era ainda uma aposta de alto risco e demonstrou, ao longo de duas décadas, que a premissa tinha fundamento financeiro.

A arquitetura financeira que o CS produz tem uma lógica circular e auto-reforçadora. CS reduz churn, churn baixo aumenta LTV, LTV alto reduz a pressão sobre CAC, CAC mais eficiente libera capital para investir mais em CS. Empresas que compreenderam essa mecânica tratam CS não como custo operacional, mas como alavanca de crescimento. A diferença de valuation documentada pela Gainsight é a expressão financeira dessa compreensão, e aponta para algo que a aritmética do SaaS tornou inevitável: numa economia de assinaturas, reter é mais valioso do que adquirir.

12.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

A relação entre Murphy e Reichheld é mais complexa do que parece à primeira leitura. O NPS que Reichheld propôs, em sua formulação original no Harvard Business Review em 2003¹⁵, mede uma intenção futura:

¹⁴ BENIOFF, Marc; ADLER, Carlye. *Behind the Cloud: The Untold Story of How Salesforce.com Went from Idea to Billion-Dollar Company and Revolutionized an Industry*. Jossey-Bass, 2009.

¹⁵ REICHHELD, Frederick F. The One Number You Need to Grow. *Harvard Business Review*, v. 81, n. 12, p. 46–54, 2003.

a propensão a recomendar. Por definição, é um indicador retrospectivo: captura o estado emocional do cliente depois que alguma experiência significativa ocorreu. O Customer Health Score que Murphy propõe é prospectivo: mede se o cliente está no caminho do Desired Outcome enquanto ainda é possível intervir. Os dois instrumentos são complementares mas operam em direções opostas no tempo. O NPS diz onde a empresa chegou; o Health Score diz para onde está indo. Em empresas com função de CS madura, os dois coexistem numa mesma arquitetura de monitoramento, cada um respondendo perguntas que o outro não consegue responder ¹⁶.

A convergência com Malcolm Knowles e a teoria da andragogia é menos óbvia mas igualmente produtiva. Knowles argumentou, ao longo dos anos 1970 e 1980, que adultos aprendem de forma fundamentalmente diferente de crianças: são movidos por motivações internas, precisam entender a relevância do que aprendem antes de aprender, e trazem para o processo de aprendizagem uma história de experiências que condiciona tanto a receptividade ao conteúdo quanto a resistência a abordagens inadequadas ao seu nível de desenvolvimento ¹⁷. O Desired Outcome de Murphy é, em termos andragógicos, a externalização explícita da motivação interna do aluno adulto. Quando Murphy argumenta que o Required Outcome precisa ser definido em termos concretos e mensuráveis desde o início da relação, está fazendo o mesmo argumento que Knowles fazia sobre a necessidade de o adulto compreender por que está aprendendo antes de estar disposto a aprender.

A convergência com Wiggins e McTighe é ainda mais direta. O modelo de Backward Design ¹⁸, que propõe que o design instrucional começa pelo resultado desejado e trabalhe retroativamente em direção ao

¹⁶ MEHTA, Nick; STEINMAN, Dan; MURPHY, Lincoln. *Customer Success*. Op. cit., pp. 189–214.

¹⁷ KNOWLES, Malcolm S.; HOLTON III, Elwood F.; SWANSON, Richard A. *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. 8. ed. Routledge, 2015.

¹⁸ WIGGINS, Grant; MCTIGHE, Jay. *Understanding by Design*. 2. ed. ASCD, 2005.

conteúdo, é a operacionalização pedagógica exata do que Murphy propõe para o design de produtos SaaS: começar pelo Desired Outcome do cliente e trabalhar retroativamente em direção ao conjunto de funcionalidades, processos de onboarding e mecanismos de suporte necessários para que esse resultado seja alcançado. A distinção entre o que Wiggins e McTighe chamam de “compreensão duradoura” e a simples retenção de informação é análoga à distinção de Murphy entre sucesso do cliente e engajamento com o produto. Compreensão duradoura é o que permanece depois que o teste acabou. Sucesso do cliente é o que permanece depois que o onboarding acabou.

A convergência com o SPIN Selling de Rackham tem uma dimensão que nem Murphy nem Rackham exploram explicitamente, mas que a APE opera na prática. A pergunta de Necessidade de Pagamento do SPIN é, em essência, o protocolo conversacional para elicitare o Desired Outcome do comprador antes do fechamento. Quando o vendedor SPIN convida o prospect a articular, com suas próprias palavras, o que mudaria se o problema fosse resolvido, está coletando a definição do Required Outcome que Murphy identificaria como fundamento do Customer Success¹⁹. Um produto vendido por meio de perguntas SPIN bem conduzidas chega ao onboarding com o Desired Outcome já definido e validado pelo próprio cliente. A função de CS começa com vantagem estrutural: há um resultado acordado contra o qual o progresso pode ser medido, e a origem desse acordo é o próprio comprador, não a empresa.

¹⁹ RACKHAM, Neil. *SPIN Selling*. McGraw-Hill, 1988.

13

A MÉTRICA QUE REVELA TUDO: FRED REICHHELD E O NET PROMOTER SCORE

13.1. CONTEXTO HISTÓRICO

A Bain & Company dos anos 1990 não era uma firma de pesquisa. Era uma firma de estratégia, e o que Frederick Reichheld fazia ali tinha uma especificidade difícil de classificar: não era pesquisa de mercado, não era consultoria de operações. Sua obsessão era a lealdade, e mais especificamente o problema de que a maioria das empresas não sabia medi-la com precisão suficiente para agir sobre ela. Não a satisfação, não a qualidade percebida do produto, não a eficiência operacional. Lealdade.

O campo da satisfação do cliente estava, naquele momento, em plena maturidade industrial. Empresas especializadas ofereciam surveys com dezenas de perguntas, escalas de Likert de cinco pontos, índices compostos com ponderações sofisticadas. O American Customer Satisfaction Index, lançado em 1994 pela Universidade de Michigan em parceria com o

American Society for Quality e o CFI Group ¹, era o exemplo mais elaborado: um modelo econométrico que agregava percepções de qualidade, expectativas e valor percebido numa pontuação entre zero e cem, aplicado trimestralmente em centenas de empresas e setores. Era estatisticamente sólido, metodologicamente respeitável e, como Reichheld passou anos documentando, não previa crescimento de negócio com qualquer confiabilidade.

O paradoxo era desconcertante. Empresas com índices de satisfação elevados perdiam participação de mercado para concorrentes com índices menores. Empresas que melhoravam seus scores ao longo de trimestres consecutivos viam a receita recorrente deteriorar. A correlação que todos assumiam existir, entre cliente satisfeito e cliente que compra de novo e recomenda, simplesmente não aparecia nos dados com a força esperada ².

Reichheld levou anos articulando por que isso acontecia. Parte da explicação estava na natureza do instrumento: surveys de satisfação mediam estado emocional imediato após interação, não disposição comportamental de longo prazo. Um cliente pode ter tido uma interação satisfatória com o call center e ainda assim estar planejando migrar para um concorrente. A satisfação capturava a experiência; a lealdade demandava algo diferente. Não a avaliação da última interação, mas uma avaliação da relação como um todo que se manifestasse como intenção de comportamento futuro ³.

O trabalho que levaria ao artigo seminal de 2003 no Harvard Business Review começou em 2001, com uma pesquisa executada em parceria com a Satmetrix. A equipe testou uma série de perguntas, correlacionando

¹ FORNELL, Claes; JOHNSON, Michael D.; ANDERSON, Eugene W.; CHA, Jaesung; BRYANT, Barbara Everitt. The American Customer Satisfaction Index: Nature, Purpose, and Findings. *Journal of Marketing*, v. 60, n. 4, p. 7–18, 1996.

² REICHHELD, Frederick F. *The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value*. Harvard Business School Press, 1996.

³ *Ibid.*, pp. 87–115.

as respostas com o comportamento real de clientes nos doze meses seguintes: compras repetidas, indicações, cancelamentos. O objetivo era identificar qual pergunta, dentre as testadas, tinha maior poder preditivo sobre o comportamento que realmente importava para o crescimento do negócio ⁴. A resposta que emergiu dos dados não era intuitiva antes de chegar.

13.2. O PROBLEMA QUE REICHHELD ESTAVA RESOLVENDO

O primeiro problema era de natureza técnica: um desalinhamento entre o que as métricas de satisfação mediam e o que as empresas precisavam saber. Satisfação, como construto psicológico formalizado por Oliver nos anos 1980, reflete a comparação entre expectativa e resultado percebido ⁵. É uma medida de estado: como o cliente se sente em relação a uma experiência específica, num momento específico. Esse é exatamente o tipo de dado que não prediz comportamento econômico de longo prazo, porque esse comportamento depende de fatores que a satisfação com uma interação individual não captura: qualidade percebida do produto ao longo do tempo, alternativas disponíveis, barreiras de troca, coerência com a identidade do consumidor, experiências de terceiros com a marca.

Havia ainda um problema de escala de mensuração. Escalas de satisfação com múltiplos pontos produzem distribuições comprimidas para a extremidade positiva (o que os psicólogos chamam de “leniency bias”): a maioria dos clientes que se dão ao trabalho de responder um survey de satisfação avalia positivamente, independentemente da lealdade real. Num mercado onde 85% dos clientes reportam satisfação e a empresa

⁴ REICHHELD, Frederick F. The One Number You Need to Grow. *Harvard Business Review*, v. 81, n. 12, p. 46–54, 2003.

⁵ OLIVER, Richard L. A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, v. 17, n. 4, p. 460–469, 1980.

perde participação de mercado para um concorrente que também tem 85% de satisfação, a métrica perdeu qualquer poder discriminativo.

A dificuldade maior, porém, não era técnica. Empresas tinham construído estruturas inteiras, departamentos de qualidade, processos de pesquisa periódica, sistemas de reporte ao conselho, em torno de métricas de satisfação. Essas estruturas tinham inércia própria. O fato de que as métricas não previam crescimento não era suficiente para desmantelá-las, porque o problema só se tornava visível quando se comparava a evolução dos scores com a evolução do crescimento ao longo de anos, e esse tipo de análise longitudinal não era rotineiro nos ciclos de reporte executivo ⁶.

O que Reichheld precisava não era apenas de uma métrica melhor. Precisava de uma métrica simples o suficiente para ser monitorada por toda a organização, sem necessidade de treinamento estatístico, e relevante o suficiente para substituir o aparato de satisfação existente com credibilidade. Uma única pergunta era uma proposta radical, quase provocatória no contexto metodológico da época.

13.3. IDEIA CENTRAL

A pergunta que Reichheld isolou como preditora de crescimento não era nova. O que era novo era a evidência empírica de que essa pergunta, especificamente, correlacionava com crescimento diferencial de forma mais robusta do que qualquer combinação de perguntas testada na pesquisa da Bain e Satmetrix ⁷.

“Em uma escala de 0 a 10, qual a probabilidade de você recomendar esta empresa a um amigo ou colega?”

⁶ REICHHELD, Frederick F.; MARKEY, Rob. *The Ultimate Question 2.0: How Net Promoter Companies Thrive in a Customer-Driven World*. Harvard Business Review Press, 2011, pp. 1–35.

⁷ REICHHELD, Frederick F. *The One Number You Need to Grow*. Op. cit., p. 48.

A formulação importa em cada detalhe. Não pergunta sobre satisfação: pergunta sobre recomendação, que é uma intenção de comportamento futuro, não uma avaliação de experiência passada. E a recomendação é endereçada a uma pessoa específica, um amigo ou colega, o que força o respondente a assumir responsabilidade pessoal: indicar algo a alguém que você conhece é um ato com consequências sociais reais, e isso transforma a resposta de uma avaliação abstrata numa intenção comportamental com peso diferente.

A segmentação que Reichheld propôs a partir das respostas divide os clientes em três grupos com lógicas econômicas distintas. Promotores, que respondem 9 ou 10, são clientes que não apenas renovam, mas indicam ativamente a empresa para suas redes. A análise dos dados mostrou que promotores geravam, em média, receita incremental significativamente maior do que outros clientes ao longo de um horizonte de três a cinco anos, em parte pelo próprio comportamento de compra e em parte pelo volume de indicações que produziam. Passivos, que respondem 7 ou 8, estão satisfeitos mas não entusiasmados o suficiente para recomendar de forma ativa e consistente. São vulneráveis a ofertas da concorrência porque a lealdade que exibem é de conveniência, não de convicção. Detratores, que respondem entre 0 e 6, não apenas têm baixa probabilidade de renovar: têm probabilidade de produzir boca a boca negativo que corrói ativamente a reputação da empresa em suas redes ⁸.

O cálculo é deliberadamente simples. NPS = percentual de Promotores menos percentual de Detratores. Passivos são descartados do cálculo, o que é uma escolha com justificativa empírica: passivos têm comportamento econômico neutro, não contribuindo significativamente nem para crescimento nem para deterioração da base de clientes. O resultado é um número entre -100 (todos os clientes são Detratores) e +100 (todos os clientes são Promotores). A maioria das empresas maduras opera entre 20

⁸ Ibid., pp. 49–51.

e 60; scores acima de 70 são raros e geralmente associados a empresas com propostas de valor excepcionalmente diferenciadas ⁹.

O que Reichheld demonstrou era que o NPS correlacionava com crescimento diferencial num nível que métricas de satisfação não alcançavam. Em serviços financeiros, telecomunicações e varejo, as empresas com os NPS mais altos cresciam, em média, a taxas mais que duas vezes superiores às de concorrentes com NPS baixos no mesmo segmento ¹⁰. A correlação não era universal: em setores com poucos competidores e altas barreiras de troca, o NPS tinha menos poder preditivo porque o crescimento dependia mais da estrutura de mercado do que da lealdade. Mas nos setores onde havia escolha real e boca a boca influenciava decisão de compra, a correlação era suficientemente robusta para justificar o NPS como proxy de crescimento futuro.

A lógica causal subjacente é mais sofisticada do que a fórmula sugere. Reichheld argumentou que a recomendação ativa a um conhecido é o ato que maximamente reflete a confiança do cliente na empresa: ao recomendar, o cliente coloca sua reputação pessoal em jogo. Alguém que faria isso é, por definição, alguém com nível de convicção que vai além de estar satisfeito. Não é um estado emocional: é uma disposição para agir ¹¹. E disposições para agir predizem comportamento de forma que estados emocionais não conseguem.

13.4. A EVOLUÇÃO POSTERIOR DA IDEIA

O artigo de 2003 no Harvard Business Review foi seguido pelo livro “The Ultimate Question” em 2006 ¹², que formalizou o sistema e expandiu a

⁹ REICHHELD, Frederick F.; MARKEY, Rob. *The Ultimate Question 2.0*. Op. cit., pp. 36–72.

¹⁰ REICHHELD, Frederick F. *The One Number You Need to Grow*. Op. cit., p. 52.

¹¹ REICHHELD, Frederick F.; MARKEY, Rob. *The Ultimate Question 2.0*. Op. cit., pp. 73–112.

¹² REICHHELD, Frederick F. *The Ultimate Question: Driving Good Profits and True Growth*. Harvard Business School Press, 2006.

metodologia com instruções para implementação em escala organizacional. A adoção foi extraordinária: Apple, Amazon, American Express, Intuit e dezenas de outras empresas de primeira linha adotaram o NPS como métrica primária de relacionamento com clientes antes de 2010. O Gartner estimou, em 2018, que dois terços das empresas do Fortune 1000 utilizavam NPS de alguma forma ¹³.

A crítica metodológica chegou com a adoção em escala. Em 2007, Keiningham e colegas publicaram no *Journal of Marketing* uma reanálise dos dados originais que questionava a superioridade do NPS sobre outras métricas de lealdade ¹⁴. O estudo arguia que métricas distintas tinham desempenho preditivo comparável dependendo do setor e do contexto, e que a afirmação de que o NPS era “a única pergunta que você precisa fazer” era mais retórica do que empiricamente sustentada. A resposta de Reichheld foi pragmática: o NPS não reivindicava ser a única métrica relevante, mas a métrica com melhor razão entre valor preditivo e custo de implementação para a maioria das organizações.

A adoção em escala trouxe um segundo problema, mais difícil de resolver do que a crítica acadêmica. A prática de inflar o NPS por design de pesquisa tornou-se estrutural. Empresas aprenderam que timing da pesquisa, seleção de respondentes e programas de incentivo a avaliações altas podiam elevar o score sem qualquer mudança real no comportamento dos clientes. É o que Reichheld chamou de “gaming” do NPS: a Goodhart’s Law aplicada à lealdade. Quando o NPS vira objetivo em vez de indicador, perde a utilidade diagnóstica que o justificava.

A versão mais sofisticada do sistema, chamada de NPS Transacional, adicionou uma pergunta aberta imediatamente depois do score: “O que motivou a sua avaliação?” O dado qualitativo resultante não altera o cálculo, mas transforma o instrumento de termômetro em diagnóstico:

¹³ GARTNER. *Voice of the Customer Programs: Market Guide*. Gartner Research, 2018.

¹⁴ KEININGHAM, Timothy L.; COOIL, Bruce; ANDREASSEN, Tor Wallin; AKSOY, Lerzan. A Longitudinal Examination of Net Promoter and Firm Revenue Growth. *Journal of Marketing*, v. 71, n. 3, p. 39–51, 2007.

é possível identificar os drivers específicos que levam clientes a serem Promotores ou Detratores, o que permite priorizar investimentos na experiência do cliente com base em evidência ¹⁵.

A evolução metodológica mais significativa veio do próprio Reichheld. Em 2021, em artigo no Harvard Business Review, ele introduziu o Earned Growth Rate: a taxa de crescimento atribuível exclusivamente a clientes existentes, seja por expansão de receita desses clientes ou por novos clientes indicados por eles, descontando crescimento vindo de aquisição paga ¹⁶. A métrica complementa o NPS de forma estrutural: o NPS mede intenção de recomendar; o Earned Growth Rate mede o crescimento efetivo que essas intenções produziram. A distinção é importante porque permite diferenciar empresas onde o NPS alto se traduz em crescimento orgânico real das empresas onde o NPS é inflado por programas de incentivo mas não gera comportamento de indicação genuíno.

13.5. IMPACTO REAL NO MUNDO

Antes do NPS, a lógica dominante de gestão do cliente era essencialmente reativa: invista em qualidade do produto, meça satisfação, resolva problemas quando surgem. O NPS introduziu uma lógica proativa: crescimento orgânico é gerado por clientes que recomendam, e isso significa que a experiência do cliente precisa ser gerida como função estratégica de geração de receita, não como função operacional de mitigação de reclamações. Essa mudança de perspectiva tinha consequências orçamentárias concretas. Ela justificava investimento em Customer Success, em onboarding, em qualidade pós-venda, de formas que o paradigma de satisfação não conseguia sustentar financeiramente porque não conectava

¹⁵ REICHHELD, Frederick F.; MARKEY, Rob. *The Ultimate Question 2.0*. Op. cit., pp. 113–155.

¹⁶ REICHHELD, Frederick F. Net Promoter 3.0. *Harvard Business Review*, v. 99, n. 6, p. 98–105, 2021.

esses gastos a crescimento mensurável.

O caso da Apple é frequentemente citado como exemplo de NPS como vantagem competitiva estrutural. Antes da abertura das Apple Stores, a empresa tinha NPS estimado em torno de 58. Após a implementação das lojas com protocolo específico de atendimento e a consolidação do ecossistema de produtos integrados, o score subiu para faixas entre 70 e 90, dependendo do período e da metodologia de coleta ¹⁷. O crescimento orgânico da base de usuários da Apple nos anos seguintes era, em parte, explicável pela intensidade com que promotores indicavam produtos da marca para suas redes. A correlação não é simples o suficiente para ser atribuição causal direta, mas é suficientemente consistente para figurar como dado relevante na análise do período.

Em serviços financeiros, o NPS transformou a lógica de avaliação de desempenho de agências e consultores. O USAA, com NPS consistente acima de 70 num setor que opera em média abaixo de 30, cresceu sua base de clientes por décadas quase exclusivamente por indicações ¹⁸. Num setor onde a maioria dos concorrentes investia pesadamente em aquisição paga, o USAA operava com custo de aquisição próximo de zero porque o produto, combinado com a experiência de atendimento, produzia o nível de convicção necessário para que clientes colocassem sua reputação em jogo ao recomendar.

No campo da educação, o NPS começou a aparecer como instrumento nos anos 2010, inicialmente em bootcamps e programas de educação profissional de curta duração. A distinção que o NPS introduz no contexto educacional é a mesma que introduz no contexto empresarial: um aluno que daria nota 9 ou 10 para recomendar o programa não é necessariamente um aluno satisfeito com o conteúdo. É um aluno cujo resultado foi suficientemente significativo para que ele colocasse sua reputação

¹⁷ TEMKIN, Bruce. *Net Promoter Score Benchmark Study, 2013*. Temkin Group Research, 2013.

¹⁸ USAA. *Annual Report 2019*. USAA, 2020.

em jogo ao indicar o programa para alguém que conhece. Essa distinção, num mercado saturado de promessas de conteúdo onde quase todos os programas reportam altas taxas de satisfação, é o que separa o sinal do ruído.

13.6. CONEXÕES COM OUTROS AUTORES DO LIVRO

Murphy argumentou que o NPS é um lagging indicator: captura o resultado de um processo que já aconteceu. O Customer Health Score que ele desenvolveu é prospectivo: mede o estado atual de progressão do cliente em direção ao seu Desired Outcome enquanto ainda é possível intervir¹⁹. Os dois instrumentos são complementares e não concorrentes. O NPS diz se o processo de Customer Success funcionou; o Health Score diz se ele está funcionando. Em organizações com maturidade operacional suficiente, os dois coexistem numa arquitetura de monitoramento onde cada um responde perguntas que o outro não consegue responder.

O que raramente é explorado nessa convergência é que os dois autores fundamentam sua argumentação na mesma premissa sobre a natureza do valor no relacionamento com cliente. Para Murphy, o cliente que alcança seu Desired Outcome renova, expande e recomenda. Para Reichheld, o cliente que recomenda ativamente é o cliente que gerou mais valor acumulado na relação. Os dois estão descrevendo o mesmo fenômeno a partir de perspectivas diferentes: Murphy, do ponto de vista do processo que produz esse cliente; Reichheld, do ponto de vista do comportamento que esse cliente exibe quando o processo funcionou. O Promotor de Reichheld é o cliente de Murphy que alcançou o Desired Outcome.

A conexão com Anders Ericsson e a teoria da prática deliberada é

¹⁹ MEHTA, Nick; STEINMAN, Dan; MURPHY, Lincoln. *Customer Success: How Innovative Companies Are Reducing Churn and Growing Recurring Revenue*. Wiley, 2016, pp. 189–214.

estruturalmente precisa, ainda que nenhum dos dois autores a tenha desenvolvido explicitamente. Ericsson demonstrou que a maestria em domínios complexos não é função do tempo de prática: é função da qualidade do feedback durante a prática, da especificidade dos objetivos, e da intensidade do esforço cognitivo em zonas de dificuldade adjacente à competência atual ²⁰. Traduzido para o contexto educacional: o aluno que se torna Promotor de um programa não é o aluno que mais gostou do processo. É o aluno que, através do processo, desenvolveu competência real e observável, que transformou sua prática de forma percebida por ele mesmo e pelos outros à sua volta.

O elo entre Ericsson e Reichheld é o conceito de resultado verificável. Para Ericsson, performance mensurável que supera um baseline anterior é a condição para que o praticante invista na prática com a intensidade necessária para continuar melhorando. Para Reichheld, transformação percebida pelo cliente é a condição para que esse cliente recomende. Em ambos os casos, o mecanismo não é satisfação com a experiência: é a percepção de transformação genuína ²¹. Um aluno que termina um curso com habilidade verificavelmente maior do que tinha antes é um candidato natural a Promotor. Um aluno que termina com o mesmo nível de habilidade, independentemente de ter gostado das aulas, dificilmente recomendará com a convicção necessária para que a recomendação tenha peso.

²⁰ ERICSSON, K. Anders; KRAMPE, Ralf T.; TESCH-RÖMER, Clemens. The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review*, v. 100, n. 3, p. 363–406, 1993.

²¹ ERICSSON, K. Anders; POOL, Robert. *Peak: Secrets from the New Science of Expertise*. Houghton Mifflin Harcourt, 2016, pp. 98–142.

PARTE V

Sintese Evolutiva

14

A CONVERGÊNCIA: QUANDO DOZE PESQUISADORES CONSTROEM A MESMA RESPOSTA

“A ciência avança não porque um gênio solitário resolve um problema, mas porque pesquisadores em campos distintos descobrem independentemente que o problema que estavam resolvendo é o mesmo.” *Adaptado de Thomas Kuhn, A Estrutura das Revoluções Científicas* ¹

14.1. O PADRÃO QUE EMERGE

Há um problema que nenhum campo do conhecimento resolveu sozinho no século XX. Formulado com precisão: como transferir, de um expert que sabe, para um aluno que ainda não sabe, um conjunto de competências complexas, de forma que o aluno possa executá-las com confiabilidade em contextos variáveis, sem a presença contínua do expert?

¹ KUHN, Thomas S. *The Structure of Scientific Revolutions*. 3. ed. University of Chicago Press, 1996.

Cada fragmento dessa formulação aponta para uma zona de pesquisa específica. “Expert que sabe” remete à literatura sobre conhecimento tácito (Polanyi) e expertise (Ericsson). “Aluno que ainda não sabe” convoca a andragogia (Knowles) e os níveis de consciência (Schwartz). “Competências complexas” exige a taxonomia de Bloom e o *expertise reversal effect* de Sweller². “De forma verificável” requer a definição rigorosa de OUTPUT que Wiggins e McTighe³ e Murphy⁴ constroem por caminhos distintos. “Em contextos variáveis” demanda prática deliberada com variação controlada, contribuição central de Ericsson⁵. “Sem a presença contínua do expert” pressupõe um sistema que funcione de forma autônoma, o que Gagné⁶ e Senge⁷ descrevem com linguagens diferentes, mas com estrutura análoga.

O que esse mapeamento revela não é uma analogia conveniente. É uma propriedade estrutural do problema: nenhuma teoria cobre duas lacunas; nenhuma lacuna é coberta por duas teorias. Essa correspondência biunívoca é evidência de que o framework captura uma estrutura real, não uma organização conveniente das leituras.

A fragmentação disciplinar do século XX tornou essa síntese invisível por muito tempo. As grandes universidades reorganizadas sob o modelo humboldtiano de especialização departamental⁸ construíram campos cada vez mais impermeáveis: a psicologia cognitiva conversava pouco com a gestão empresarial; a epistemologia raramente dialogava com o design instrucional. O aprofundamento produziu profundidade. O custo foi a integração. O problema não estava nas disciplinas, mas na ausência

² SWELLER, John. Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, v. 12, n. 2, p. 257–285, 1988.

³ WIGGINS, Grant; MCTIGHE, Jay. *Understanding by Design*. 2. ed. Association for Supervision and Curriculum Development (ASCD), 2005.

⁴ MURPHY, Lincoln. *Customer Success: The Definitive Guide*. Gainsight, 2016.

⁵ ERICSSON, K. Anders; POOL, Robert. *Peak: Secrets from the New Science of Expertise*. Houghton Mifflin Harcourt, 2016.

⁶ GAGNÉ, Robert M. *The Conditions of Learning and Theory of Instruction*. 4. ed. Holt, Rinehart and Winston, 1985.

⁷ SENGE, Peter M. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday/Currency, 1990.

⁸ CLARK, Burton R. *The Higher Education System: Academic Organization in Cross-National Perspective*. University of California Press, 1983.

de uma estrutura que tornasse a conexão entre elas visível e necessária.

Não é a primeira vez que esse padrão aparece na história da ciência. Norbert Wiener, ao fundar a cibernética nos anos 1940, identificou princípios de controle e comunicação comuns a máquinas e organismos vivos que nenhuma disciplina isolada havia percebido⁹. Ludwig von Bertalanffy demonstrou que padrões estruturais como feedback negativo e hierarquia de subsistemas emergem em todos os sistemas complexos, independentemente de sua natureza¹⁰. Em ambos os casos, a síntese foi trabalho posterior às descobertas individuais: Newton não traduziu sua mecânica em ferramentas de engenharia; Watson e Crick não desenvolveram a biologia molecular. A integração sempre exige uma geração seguinte.

O problema prático que justifica a síntese neste livro é urgente. O mercado global de educação digital movimentou aproximadamente 375 bilhões de dólares em 2022, com projeções de 600 bilhões até 2027¹¹. Taxas de conclusão entre 12% e 15% caracterizam MOOCs e cursos digitais¹²; NPS médios abaixo de 30 pontos descrevem os produtos educacionais premium. O fracasso não é aleatório. É a consequência previsível de usar teorias parciais para resolver um problema que, estruturalmente, exige todas as partes.

A proposição central deste capítulo decorre diretamente disso: o framework INPUT → GATEKEEPER → BLACK BOX → OUTPUT → FEEDBACK não foi criado para organizar as teorias. Foi extraído delas. O critério de mapeamento não é vocabular, pois não se trata de encontrar onde cada autor usou palavras como “entrada” ou “resultado”, mas

⁹ WIENER, Norbert. *Cybernetics: Or Control and Communication in the Animal and the Machine*. MIT Press, 1948.

¹⁰ BERTALANFFY, Ludwig von. *General System Theory: Foundations, Development, Applications*. George Braziller, 1968.

¹¹ GLOBAL MARKET INSIGHTS. *E-Learning Market Size, Industry Analysis Report, Regional Outlook, Growth Potential, Competitive Market Share & Forecast, 2023–2032*. Global Market Insights, 2023.

¹² JORDAN, Katy. Massive Open Online Course Completion Rates Revisited: Assessment, Length and Attrition. *The International Review of Research in Open and Distributed Learning*, v. 16, n. 3, 2015. <https://doi.org/10.19173/irrodl.v16i3.2112>

funcional: cada teoria é posicionada pelo problema que resolve quando aplicada ao desafio da transferência de conhecimento especializado. O resultado revela cobertura completa dos cinco nós com sobreposições mínimas e necessidade mútua máxima.

14.2. MAPEAMENTO DOS AUTORES SOBRE O FRAMEWORK

14.2.1. O Nó INPUT: Três Dimensões de uma Mesma Condição

O INPUT é o ponto de partida lógico do sistema, não o operacional. Antes de qualquer decisão sobre como ensinar, três perguntas precisam ser respondidas em sequência, e a sequência importa: o que, exatamente, o expert sabe que pode ser transferido? Para quem esse conteúdo se destina, e em que forma precisa chegar para ser processável? O aluno está, neste momento, em condições cognitivas de reconhecê-lo como relevante para um problema seu?

Cada pergunta corresponde a um corpo de pesquisa distinto. A ausência de qualquer das três torna o INPUT incompleto, independentemente da qualidade dos outros dois.

Polanyi e o problema da origem do conteúdo. Antes de qualquer decisão de design instrucional, há um problema de elicitación¹³. O conhecimento tácito, incorporado em disposições perceptivas, julgamentos de situação e heurísticas de decisão desenvolvidas na prática, não está disponível para transcrição direta. O conteúdo mais valioso de um expert é precisamente o que ele não consegue articular espontaneamente: a parte que escapa à verbalização não por relutância, mas por estrutura cognitiva. Pode ser extraído, mas somente por processos que criem condições para sua externalização indireta: reconstrução retrospectiva de casos crí-

¹³ POLANYI, Michael. *The Tacit Dimension*. Doubleday & Company, 1966.

ticos, análise de contra-exemplos, identificação de variações que o expert percebe mas não nomeia.

Gary Klein operacionalizou isso no *Critical Decision Method*¹⁴, desenvolvido para bombeiros e engenheiros de planta nuclear cujas decisões eram tomadas de forma tão automatizada que os próprios executores não as conseguiam articular em tempo real. O CDM não pede ao expert que descreva o que sabe. Pede que reconstrua o que fez, em qual contexto, com quais opções à disposição. O tácito emerge pelo rastreamento das escolhas, não pela sua declaração. Sem esse processo de elicitação, o produto será construído sobre a parte mais acessível e menos determinante da competência do expert.

Knowles e o problema do destinatário. Extraído o conteúdo, o segundo problema é para quem ele se destina e em que forma precisa chegar para ser processável por esse destinatário específico. Malcolm Knowles estabeleceu que adultos não aprendem como crianças¹⁵, não por questão de capacidade cognitiva, mas por estrutura de motivação e relevância. Os princípios andragógicos descrevem um aprendiz que só processa ativamente conteúdo percebido como imediatamente aplicável, que precisa entender o porquê de cada elemento antes de engajá-lo, que traz para o processo uma experiência acumulada capaz de funcionar tanto como recurso quanto como obstáculo.

Um INPUT correto em conteúdo, com Polanyi resolvido, mas apresentado em formato incompatível com essa estrutura motivacional, produz o mesmo resultado prático que um INPUT incorreto em conteúdo. O aluno não o processa como relevante e não o integra. Esse tipo de falha é invisível nos primeiros dias do produto. Aparece no churn da segunda semana.

Schwartz e o problema da compatibilidade cognitiva. A terceira

¹⁴ KLEIN, Gary A. *Sources of Power: How People Make Decisions*. MIT Press, 1998.

¹⁵ KNOWLES, Malcolm S.; HOLTON, Edward F.; SWANSON, Richard A. *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. 8. ed. Routledge, 2015.

dimensão é a menos intuitiva. Extrair o conteúdo certo e apresentá-lo na forma adequada é condição necessária, mas não suficiente: o aluno precisa estar, naquele momento, em condições cognitivas de reconhecê-lo como algo que resolve um problema seu. Eugene Schwartz mapeou cinco níveis de consciência do comprador ¹⁶: não consciente do problema, consciente do problema, consciente da solução, consciente do produto, mais consciente. Um aluno no nível 1 inscrito num produto construído para o nível 4 não o experimentará como relevante, não porque o produto seja ruim, mas porque a distância entre sua consciência atual e a pressuposta pelo produto é grande demais para ser atravessada pelo conteúdo isolado.

Polanyi, Knowles e Schwartz definem as três dimensões do INPUT: o que foi extraído, para quem e sob qual condição cognitiva de recepção. O sistema agora sabe o que vai entregar. A pergunta que segue, necessariamente, é quem tem condições reais de recebê-lo.

14.2.2. O Nó GATEKEEPER: Qualificação como Derivação do OUTPUT

O GATEKEEPER é o nó cuja ausência o mercado mais sistematicamente normaliza. Filtrar a entrada significa, em alguma medida, recusar compradores. A pressão por volume tende a relaxar os critérios até que o sistema trate todos os interessados como equivalentes, com resultado previsível: alunos sem as condições necessárias sobrecarregam o suporte, abandonam mais cedo e, quando chegam ao final, frequentemente não alcançam os resultados prometidos. A inversão lógica que o mercado ignora é que a qualificação só pode ser feita corretamente a partir do OUTPUT. É preciso saber para onde o sistema conduz antes de definir

¹⁶ SCHWARTZ, Eugene M. *Breakthrough Advertising*. Bottom Line Books, 2004. (Obra original publicada em 1966.)

quem tem condições de ser conduzido.

Wiggins e McTighe e a precedência lógica do OUTPUT. A contribuição de Wiggins e McTighe opera em dois nós do framework, e sua função no GATEKEEPER é anterior a qualquer protocolo de qualificação. O *backward design*⁶ estabelece que o resultado esperado precisa ser definido com precisão verificável antes de qualquer decisão sobre conteúdo, metodologia ou critérios de entrada. Sem um OUTPUT operacionalizado, o GATEKEEPER não tem referência para qualificar: os critérios de entrada são, por definição, as condições mínimas que o aluno precisa satisfazer para alcançar o OUTPUT. Sistemas que tentam qualificar sem antes definir para onde estão qualificando acabam usando proxies inadequados como disponibilidade financeira, entusiasmo declarado ou experiência genérica no setor.

Rackham e a qualificação como diagnóstico. Uma vez estabelecido o OUTPUT que orienta os critérios, o problema operacional é como conduzir a qualificação sem que seja percebida como barreira. Neil Rackham, a partir de 35.000 chamadas de venda analisadas em 12 países ao longo de uma década¹⁷, chegou a uma conclusão que inverte a lógica convencional: nas vendas de alto valor, o interlocutor que tenta persuadir deteriora sua taxa de conversão; o que faz perguntas diagnósticas de alta qualidade aumenta sistematicamente tanto a probabilidade de fechamento quanto a qualidade do cliente adquirido. A sequência SPIN (Situação, Problema, Implicação, Necessidade de Pagamento) é um protocolo de entrevista diagnóstica que, quando executado com rigor, simultaneamente qualifica o candidato e o move pelo funil de consciência de Schwartz. Um candidato que não consegue articular as implicações de seu problema atual provavelmente não está no nível de consciência adequado para o produto; um que chega espontaneamente à Necessidade de Pagamento demonstrou a clareza e a urgência que são preditores de

¹⁷ RACKHAM, Neil. *SPIN Selling*. McGraw-Hill, 1988.

conclusão e resultado.

O GATEKEEPER adequadamente desenhado não é uma barreira. É a condição pela qual o sistema pode honrar o que prometeu a quem tem condições reais de alcançá-lo.

14.2.3. O Nó BLACK BOX: A Sobreposição de Cinco Problemas Independentes

A transferência de competências complexas não é um processo unitário. É a sobreposição de pelo menos cinco problemas independentes: como sequenciar a instrução para que cada unidade ative as condições cognitivas necessárias para a próxima; qual modalidade produz a retenção adequada para o resultado prometido; como calibrar o material para não exceder os limites da memória de trabalho; como criar as condições de prática que produzem competência real e não mera familiaridade com conceitos; em que progressão cognitiva as unidades devem ser dispostas. Cinco autores ocupam esse nó, cada um resolvendo um problema que os outros quatro não tocam. A interdependência entre eles é real e não decorativa.

Gagné e a estrutura da instrução. Robert Gagné, após décadas investigando as condições de produção de diferentes tipos de resultado de aprendizagem ¹¹, chegou ao modelo dos Nove Eventos de Instrução: ganhar atenção, informar sobre o objetivo, estimular a recordação de aprendizados anteriores, apresentar o conteúdo, proporcionar orientação, provocar o desempenho, fornecer feedback, avaliar o desempenho, intensificar a retenção e a transferência. O modelo é frequentemente mal compreendido como lista de passos. É, com mais precisão, uma descrição das condições cognitivas que precisam estar presentes em cada unidade de aprendizagem para que ela produza o resultado a que se pro-

põe. Omitir a ativação de conhecimento prévio antes da apresentação do conteúdo não é pedagogicamente subótimo: é uma violação das condições que a memória de trabalho precisa para processar informação nova sem sobrecarga. Gagné fornece a sequência; Sweller explica por que ela funciona.

Glasser e a escolha da modalidade. William Glasser sistematizou as taxas de retenção por modalidade ¹⁸: leitura retém aproximadamente 10%; ouvir, 20%; ver e ouvir, 50%; discutir, 70%; praticar, 80%; ensinar outros, 95%. Esses números variam com o tipo de conteúdo e o perfil do aprendiz, mas a direção é robusta. Modalidades ativas retêm dramaticamente mais que modalidades passivas, e um produto que entrega 80% de seu conteúdo em vídeo-aula unidirecional está estruturalmente limitando o tipo de resultado que pode entregar, qualquer que seja a qualidade do conteúdo. A escolha de modalidade não é decisão de produção. É decisão de resultado.

Sweller e os limites da memória de trabalho. John Sweller chegou à Teoria da Carga Cognitiva ¹⁰ investigando por que materiais de física bem estruturados produziam aprendizagem tão baixa. A memória de trabalho, responsável pelo processamento ativo de informação nova, tem capacidade drasticamente limitada, estimada por George Miller em 7 ± 2 unidades ¹⁹, com estudos posteriores sugerindo números menores para material genuinamente novo. A distinção entre carga intrínseca (inerente à complexidade do conteúdo), extrínseca (gerada por design instrucional ruim, portanto eliminável) e germinal (o esforço cognitivo produtivo que leva à aprendizagem real) fornece ao BLACK BOX um princípio de design preciso: não minimizar carga total, mas eliminar a extrínseca, calibrar a intrínseca e maximizar a germinal.

O *expertise reversal effect* torna isso particularmente relevante para

¹⁸ GLASSER, William. *The Quality School: Managing Students Without Coercion*. Harper & Row, 1990.

¹⁹ MILLER, George A. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review*, v. 63, n. 2, p. 81-97, 1956.

produtos criados por especialistas. O que é carga intrínseca baixa para o expert é frequentemente carga intrínseca alta para o iniciante. Ao criar material sem intermediação de um processo de revisão de carga, o expert tende sistematicamente a sobrecarregar os alunos. Não por má intenção. Por estrutura cognitiva.

Ericsson e a prática deliberada. K. Anders Ericsson identificou que a competência de alto nível é construída por um tipo específico de prática ⁵: deliberada, caracterizada por foco em aspectos específicos a melhorar, feedback imediato, operação na zona de dificuldade proximal e repetição com variação controlada. Esse processo constrói representações mentais sofisticadas: estruturas cognitivas que permitem ao expert perceber padrões que o iniciante não vê, tomar decisões em tempo real que o iniciante não conseguiria articular e adaptar-se a variações de contexto que quebrariam uma execução baseada em regras explícitas. Um produto sem ciclos de prática deliberada não está produzindo competência. Está produzindo familiaridade com conceitos: necessária, mas estruturalmente insuficiente para o resultado que os produtos premium prometem.

Bloom e a progressão cognitiva. Benjamin Bloom, ao desenvolver a taxonomia de objetivos educacionais ²⁰ e as revisões subsequentes por Anderson e Krathwohl ²¹, forneceu ao BLACK BOX seu princípio de progressão. Os seis níveis (Lembrar, Entender, Aplicar, Analisar, Avaliar, Criar) são uma hierarquia de dependência estrutural, não uma classificação de tipos de aprendizagem. Não se analisa sem antes aplicar; não se aplica sem antes entender. A maioria dos produtos educacionais digitais opera nos dois primeiros níveis porque são os únicos que o formato de vídeo-aula consegue endereçar com facilidade. Os níveis superiores

²⁰ BLOOM, Benjamin S. et al. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. Longmans, Green, 1956.

²¹ ANDERSON, Lorin W.; KRATHWOHL, David R. (Orgs.). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman, 2001.

exigem atividades de aplicação, projetos com feedback, análise de casos com orientação. Bloom fornece o mapa para saber onde um produto está operando na taxonomia e, por consequência, que tipo de resultado pode e não pode prometer.

A cadeia de interdependência torna-se visível quando lida em conjunto: o nível cognitivo que Bloom especifica determina a modalidade que Glasser prescreve; a modalidade determina a carga que Sweller precisa calibrar; a carga calibrada determina a estrutura de eventos que Gagné sequencia; os eventos sequenciados precisam incluir os ciclos de prática deliberada que Ericsson identifica como condição da competência real. Nenhum dos cinco resolve o problema do outro. Cada um é condição dos demais.

14.2.4. O Nó OUTPUT: Da Entrega ao Resultado Verificável

O OUTPUT é o nó paradoxalmente mais negligenciado porque parece o mais óbvio. Todo produto educacional tem um resultado esperado. O problema é de operacionalização: a maioria dos resultados declarados é formulada em termos de conteúdo entregue (“o aluno aprenderá X, Y e Z”), não de competência adquirida verificável (“o aluno será capaz de fazer X em condições Y, de forma observável por terceiros”). Essa distinção não é semântica. Um resultado definido em termos de conteúdo pode ser conferido; não pode ser verificado. Um resultado definido em termos de competência verificável orienta retroativamente toda a arquitetura do produto, e é essa inversão que o *backward design* de Wiggins e McTighe formaliza em sua segunda função no framework.

Murphy e a distinção entre satisfação e sucesso. Lincoln Murphy, desenvolvendo o campo de Customer Success a partir da experiência em empresas de SaaS, identificou a distinção que mais diretamente traduz o

problema do OUTPUT educacional: a diferença entre *required outcome* (o resultado que o cliente tecnicamente alcançou) e *appropriate experience* (a forma como vivenciou o processo de alcançá-lo) ⁸. O conceito de *desired outcome*, a combinação precisa dos dois, resolve o problema prático de produtos que entregam o conteúdo prometido mas não a transformação que motivou a compra.

No contexto educacional, essa distinção explica o NPS baixo em produtos com alta taxa de satisfação declarada. Alunos que “gostaram” do curso mas não conseguiram aplicar o que aprenderam deram avaliações positivas à experiência e negativas ao resultado. O sucesso, para Murphy, não é subjetivo: é a capacidade verificável de o cliente fazer o que precisava fazer. Definir o OUTPUT nesses termos significa substituir “o aluno entendeu o conteúdo” por “o aluno é capaz de executar Y em contexto Z, de forma verificável por terceiros”. A implicação é estrutural: o produto que define o OUTPUT em termos de competência verificável precisa projetar sua BLACK BOX retroativamente a partir desse critério, especificando quais atividades de prática produzem a competência, em qual progressão e com qual estrutura de feedback. O OUTPUT, bem definido, não é o ponto de chegada do sistema. É o ponto de partida do seu design.

14.2.5. O Nó FEEDBACK: O Mecanismo pelo qual o Sistema Aprende

O FEEDBACK é o nó que transforma um produto educacional de sistema linear em sistema adaptativo. Sem ele, o produto pode melhorar por intuição e sorte, mas não por design. O que os autores que ocupam esse nó estabelecem vai além da coleta de dados: o FEEDBACK tem naturezas funcionalmente distintas dependendo de onde no sistema é capturado, e confundi-las produz informação gerencialmente útil mas pedagógica-

mente inerte.

Reichheld e o NPS como proxy do OUTPUT. Fred Reichheld chegou ao Net Promoter Score ²² investigando qual métrica de satisfação tinha maior correlação com crescimento sustentável. A descoberta foi que a única pergunta com alta correlação preditiva era a de recomendação: a disposição de recomendar não mede satisfação com a experiência. Mede a percepção de que o resultado alcançado foi real e suficientemente valioso para justificar colocar a própria reputação em jogo. No contexto educacional, o NPS funciona como proxy do OUTPUT: um aluno que não alcançou o resultado prometido raramente arriscará sua reputação recomendando o produto; um aluno que alcançou frequentemente o fará sem ser solicitado. Aplicado em pontos estratégicos da jornada, e não apenas ao final, o NPS revela onde no sistema a entrega está falhando, transformando o FEEDBACK em ferramenta de diagnóstico sistêmico.

Ericsson no FEEDBACK: o retorno como condição constitutiva da aprendizagem. Ericsson aparece duas vezes no framework porque sua teoria opera em dois nós com funções distintas. No BLACK BOX, ele descreve o mecanismo de construção de competências. No FEEDBACK, ele especifica como o retorno precisa ser estruturado para que seja pedagogicamente eficaz e não apenas gerencialmente útil. O feedback que constitui a prática deliberada ⁵ precisa ser imediato, específico ao aspecto em prática e acompanhado de informação sobre a direção da melhoria necessária. Feedback genérico não satisfaz essa condição. A dupla presença de Ericsson não é redundância: são duas funções distintas de um mesmo princípio operando em nós separados.

Senge e a estrutura do aprendizado sistêmico. Peter Senge completa o nó ao nível do sistema: os *learning loops* que ele descreve como estrutura fundamental de organizações que aprendem ⁹ formalizam a

²² REICHHELD, Frederick F. *The Ultimate Question 2.0: How Net Promoter Companies Thrive in a Customer-Driven World*. Harvard Business Review Press, 2011.

ideia de que feedback não é avaliação pós-hoc. É o mecanismo pelo qual o sistema percebe a distância entre o resultado que está produzindo e o resultado que se propôs a produzir, e se reconfigura para reduzi-la. Sem esse mecanismo, o produto pode acumular dados de múltiplas turmas sem que nenhum ciclo de aprendizagem sistêmica ocorra. Acúmulo de dados não é aprendizagem do sistema, pela mesma razão que exposição a conteúdo não é aprendizagem do aluno.

14.3. A INTEGRAÇÃO: O QUE NENHUM AUTOR DISSE SOZINHO

14.3.1. *O Teste da Necessidade*

A força de uma síntese não está em mostrar que as teorias são compatíveis. Está em demonstrar que cada uma é insubstituível. O teste preciso não é “essa teoria adiciona valor?” mas “sem essa teoria, o sistema produz sistematicamente um tipo específico de falha?” Esse critério é mais exigente porque exclui a possibilidade de as teorias serem intercambiáveis ou opcionais. Aplicado nó a nó, o teste é categórico.

No INPUT: sem Polanyi, o produto é construído sobre a parte mais articulável e menos determinante da competência do expert, ensinando o que ele sabe dizer, não o que define sua maestria. Sem Knowles, o conteúdo certo chega no formato errado: o aluno não o processa como relevante e não o integra. Sem Schwartz, o produto certo atrai o aluno errado: consciência incompatível produz abandono precoce por razões sem relação com a qualidade do que foi arquitetado.

No GATEKEEPER: sem Wiggles e McTighe, os critérios de qualificação são proxies e alunos incapazes de alcançar o OUTPUT entram porque o sistema nunca definiu o que o OUTPUT exige. Sem Rackham, o

GATEKEEPER existe como intenção mas não como protocolo, ou repelindo candidatos qualificados ao ser percebido como barreira hostil, ou aprovando candidatos inadequados por diagnóstico insuficiente.

No BLACK BOX: sem Gagné, os eventos cognitivos que produzem aprendizagem confiável não acontecem por acaso. Sem Glasser, as decisões de modalidade seguem conveniência de produção em vez de critério de retenção. Sem Sweller, o material é construído para o estado cognitivo do expert, que sistematicamente sobrecarrega o aluno com carga extrínseca evitável. Sem Ericsson, o produto pode ter estrutura e modalidade adequadas, mas não inclui os ciclos de prática que são condição necessária para a construção de representações mentais sofisticadas: o aluno passa pelo material sem desenvolver a competência de executá-lo em condições variáveis. Sem Bloom, as unidades bem desenhadas individualmente são dispostas numa progressão que exige do aluno competências que ainda não tem, ou que repete níveis que já superou.

No OUTPUT: sem Murphy, o resultado é definido em termos de conteúdo entregue, não de transformação verificável. O sistema pode funcionar por todos os critérios internos e ainda assim não produzir o que motivou a compra. O aluno satisfeito que não alcançou o resultado é o sintoma; o OUTPUT mal definido é o diagnóstico.

No FEEDBACK: sem Reichheld, o sistema mede satisfação com a experiência em vez de verificação do resultado alcançado, coletando dados que não permitem identificar onde a entrega falhou. Sem Ericsson no FEEDBACK, os dados coletados não se conectam à estrutura de prática que precisa ser ajustada: a informação torna-se gerencialmente visível e pedagogicamente inerte. Sem Senge, nenhum ciclo de aprendizagem sistêmica ocorre: turmas acumulam, dados acumulam, e a melhoria, quando acontece, é produto de intuição, não de arquitetura.

14.3.2. Zonas de Convergência

O mapeamento dos doze autores revela algo mais do que a cobertura dos nós. Revela três zonas em que pesquisadores de campos sem contato entre si chegaram ao mesmo princípio operacional por caminhos independentes. Quando autores que nunca se leram chegam à mesma conclusão, a conclusão descreve uma propriedade do problema, não uma preferência teórica.

O design parte do OUTPUT, não do INPUT. Wiggins e McTighe chegam a isso como *backward design*⁶: o resultado esperado deve ser definido antes de qualquer decisão sobre conteúdo. Murphy chega ao mesmo ponto como *desired outcome*⁸: o produto só pode ser arquitetado depois que o resultado almejado foi especificado com precisão. Bloom chega implicitamente pela taxonomia: a escolha de modalidades, sequência e eventos instrucionais depende do nível cognitivo que se quer produzir, o que só pode ser especificado com o OUTPUT claro. Três campos distintos (design instrucional, customer success, psicologia cognitiva) convergem para o mesmo princípio. A implicação é direta: a esmagadora maioria dos processos de criação de produtos educacionais começa pelo INPUT, pelo que o expert sabe, pelo que a plataforma suporta, pelo que é mais fácil de produzir. Os três autores constituem evidência convergente de que esse processo produz sistematicamente o resultado errado.

Feedback é condição constitutiva, não avaliação terminal. Ericsson identifica feedback imediato e específico como um dos quatro elementos constitutivos da prática deliberada⁵: sem ele, a prática não constrói representações mentais sofisticadas. Gagné inclui o feedback como o sétimo dos Nove Eventos, não como pós-teste, mas como componente interno de cada unidade. Senge formula os *learning loops* como a estrutura que distingue sistemas que aprendem de sistemas que apenas operam⁹.

A convergência explica por que o FEEDBACK é um nó estrutural do framework: um produto sem FEEDBACK não é um produto completo com uma peça faltando. É um sistema aberto que processa sem aprender com o que processou.

O aluno adequado é condição sistêmica, não variável de ajuste. Knowles estabelece que adultos com perfis motivacionais incompatíveis com o produto não aprendem por estrutura, não por falta de esforço¹⁷. Schwartz demonstra que mensagens apresentadas no nível de consciência errado não produzem o movimento desejado, qualquer que seja sua qualidade intrínseca⁷. Rackham demonstra empiricamente que vendas feitas para compradores com fit inadequado deterioram sistematicamente satisfação e resultados¹⁸. O GATEKEEPER não é um dispositivo de exclusão: é a condição para que o sistema entregue o que prometeu. Um produto vendido para o aluno errado não é apenas um problema comercial. É um problema de integridade sistêmica.

14.3.3. A Tensão que Não Desaparece

Além das convergências, o mapeamento revela uma tensão que nenhuma síntese elegante pode dissolver, porque ela descreve um trade-off real que o arquiteto do produto precisa administrar conscientemente.

Gagné e Bloom trabalham com sequências de alguma forma universais: certos eventos instrucionais precisam acontecer em certa ordem para que a aprendizagem ocorra [4, 11]. Knowles e Schwartz trabalham com a variabilidade intrínseca do aprendiz: adultos diferentes chegam com experiências, níveis de consciência e estruturas motivacionais distintas [7, 17]. A padronização que torna o produto replicável pode comprometer a relevância que torna a aprendizagem possível para o aprendiz adulto específico.

A resolução não é escolher um lado. É design em dois níveis: estrutura instrucional padronizada (Gagné, Bloom) sobre conteúdo personalizado para a experiência prévia e o contexto de aplicação do aluno (Knowles, Schwartz). A estrutura garante a confiabilidade do processo; a personalização garante a relevância que o adulto precisa para processá-lo. Sem essa distinção de camadas, ou o produto não escala (personalização total) ou não engaja (padronização total). A tensão não se resolve. Administra-se conscientemente.

14.3.4. *A Síntese como Instrumento Diagnóstico*

O framework não é uma teoria entre as teorias. É a estrutura que emerge quando se pergunta, para cada uma das doze, qual problema ela estava de fato resolvendo. O resultado é que o framework preexiste à sua própria formulação: ele estava nas teorias antes de ser nomeado. O que a Arquitetura de Produtos Educacionais fez foi identificá-lo no lugar onde ele sempre esteve.

Para o arquiteto de produtos educacionais, isso tem uma consequência prática. Não é necessário memorizar doze teorias para aplicá-las. É necessário entender qual problema cada nó resolve, porque o framework organiza o ponto de entrada para a teoria certa no momento certo.

Quando um produto apresenta alta taxa de abandono nos primeiros dias, o diagnóstico começa num mapa: o problema está no INPUT (aluno errado atraído), no GATEKEEPER (aluno errado qualificado) ou no BLACK BOX (sobrecarga cognitiva imediata na primeira unidade)? Cada hipótese aponta para um corpo teórico específico e para uma intervenção específica. Quando o NPS deteriora turma a turma apesar da satisfação imediata permanecer alta, o foco muda: o OUTPUT foi definido em termos de conteúdo ou de competência verificável? O FEEDBACK está me-

dindo a experiência ou o resultado alcançado? A pergunta certa, feita no momento certo, evita que o dano se acumule até se tornar irrecuperável.

É exatamente isso que um sistema com doze teorias integradas torna possível, e que cada teoria isolada, por mais robusta que seja, não permite.

15

A TESE: ARQUITETURA DE PRODUTOS EDUCACIONAIS COMO NOVA DISCIPLINA

“Uma disciplina é, acima de tudo, um comprometimento com o estudo sistemático de um domínio de fenômenos que escapa à explicação das disciplinas existentes.” adaptado de Thomas Kuhn, A Estrutura das Revoluções Científicas ¹

Há um momento reconhecível na história de qualquer campo científico: quando os problemas práticos acumularam evidência suficiente para exigir não apenas melhores ferramentas, mas um vocabulário diferente, um recorte diferente do objeto, uma forma diferente de delimitar o que conta como explicação. Este capítulo argumenta que a transferência de conhecimento especializado chegou a esse ponto. A Arquitetura de Produtos Educacionais é a resposta estrutural a ele. O argumento

¹ KUHN, Thomas S. *The Structure of Scientific Revolutions*. 3. ed. University of Chicago Press, 1996.

percorre quatro movimentos: o diagnóstico da premissa errada que organiza um mercado inteiro; a distinção epistemológica entre criar conteúdo e arquitetar transferência; a demonstração de que a integração sistêmica da metodologia APE é necessária, não contingente; e o posicionamento histórico da disciplina no mercado que a torna urgente.

15.1. O DIAGNÓSTICO: A PREMISSA ERRADA QUE ORGANIZA UM MERCADO INTEIRO

Existe um paradoxo que atravessa silenciosamente toda a indústria de educação especializada: os profissionais mais competentes em seus campos são, com frequência desproporcional, os menos capazes de transferir o que sabem. Não se trata de limitação intelectual. Tampouco de ausência de motivação. Trata-se de uma confusão estrutural, consolidada ao longo de décadas, entre o domínio de uma disciplina e a capacidade de transferi-la.

O mercado brasileiro de educação digital cresceu de forma exponencial na segunda metade da década de 2010. Plataformas como a Hotmart viabilizaram que qualquer profissional com conexão à internet oferecesse um produto educacional ao mercado. O resultado não foi a democratização da aprendizagem de qualidade: foi a proliferação de conteúdo organizado sem arquitetura de entrega (módulos gravados, apostilas exportadas e grupos de mensagens como substitutos de um sistema genuíno de suporte à aprendizagem). A evidência mais visível desse fracasso estrutural está num dado que a indústria prefere não divulgar: a taxa média de conclusão de cursos online situa-se consistentemente entre 12% e 15%². Estudos de NPS em produtos educacionais digitais indi-

² JORDAN, Katy. Massive Open Online Course Completion Rates Revisited: Assessment, Length and Attrition. *The International Review of Research in Open and Distributed Learning*, v. 16, n. 3, 2015.

cam médias que raramente ultrapassam 22 pontos ³, um número que, em qualquer outra indústria de serviços, seria tratado como crise sistêmica, não como comportamento natural do consumidor.

O argumento popular para explicar esse cenário responsabiliza o aluno: falta de comprometimento, excesso de informação disponível, competição com redes sociais pela atenção. É um argumento intelectualmente conveniente. Retira do produtor a responsabilidade pelo desempenho do produto. Não existe outra área de engenharia em que falha sistemática de entrega seja atribuída ao usuário. Se 85% dos carros saíssem da fábrica sem funcionar adequadamente, a investigação começaria nas linhas de montagem. Na educação, por razões históricas que remontam ao modelo escolástico de transmissão de conteúdo ⁴, a responsabilidade pelo não-aprendizado recai automaticamente sobre o aprendiz. É essa inversão de causalidade que permite a perpetuação da premissa errada.

A premissa errada é esta: um produto educacional existe quando o conteúdo existe.

15.1.1. A Estrutura Cognitiva do Fracasso

Em 1984, Benjamin Bloom publicou o que se tornaria o dado mais replicado e menos aplicado da história da pedagogia moderna: alunos submetidos a tutoria individual apresentavam desempenho 2 desvios-padrão acima da média de alunos em instrução coletiva, o Efeito 2 Sigma ⁵. A implicação não era apenas estatística. Era uma denúncia da falácia do modelo transmissivo. Um professor competente, ensinando para grupos, produz resultados sistematicamente inferiores ao que é possível quando

<https://doi.org/10.19173/irrodl.v16i3.2112>

³ REICHHELD, Frederick F. *The Ultimate Question 2.0: How Net Promoter Companies Thrive in a Customer-Driven World*. Harvard Business Review Press, 2011.

⁴ VERGER, Jacques. *Culture, Enseignement et Société en Occident aux XIIIe et XIIIe Siècles*. Presses Universitaires de Rennes, 1999.

⁵ BLOOM, Benjamin S. The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, v. 13, n. 6, p. 4–16, 1984. <https://doi.org/10.3102/0013189X013006004>

a instrução é individualizada e adaptativa.

Anders Ericsson, ao investigar a performance de elite em seu estudo longitudinal com violinistas da Academia de Berlim, identificou o mecanismo cognitivo subjacente: experts não apenas sabem mais, eles organizam o conhecimento de forma estruturalmente diferente de iniciantes, através do que denominou *representações mentais sofisticadas*⁶. Essas representações são produto de anos de prática deliberada e são, em larga medida, invisíveis para quem as possui. O expert passou tanto operando a partir delas que não percebe que são uma construção adquirida. Quando senta para gravar sua primeira aula, começa pelo que sabe dizer, não pelo que precisa ser dito para que o aluno chegue ao resultado.

Michael Polanyi havia nomeado o fenômeno três décadas antes, com a formulação que permanece como a descrição mais precisa do problema central da transferência de expertise: “*we can know more than we can tell*”⁷. O conhecimento tácito (incorporado em disposições, julgamentos automáticos e heurísticas desenvolvidas na prática) não apenas resiste à verbalização: ativamente escapa a ela. O conteúdo que o expert produz pode ser correto e sofisticado, e ainda assim não ser um produto educacional. Será, na melhor das hipóteses, uma transcrição da parte mais acessível de sua competência, que é, por definição, a parte menos determinante de seu desempenho diferencial.

John Sweller forneceu a base cognitiva para entender por que esse problema é estruturalmente resistente a soluções intuitivas. Sua Teoria da Carga Cognitiva⁸ demonstrou que a memória de trabalho humana é um recurso drasticamente limitado (George Miller havia estimado a capacidade em 7 ± 2 unidades de informação⁹) e distinguiu três tipos de

⁶ ERICSSON, K. Anders; POOL, Robert. *Peak: Secrets from the New Science of Expertise*. Houghton Mifflin Harcourt, 2016.

⁷ POLANYI, Michael. *The Tacit Dimension*. Doubleday & Company, 1966.

⁸ SWELLER, John. Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, v. 12, n. 2, p. 257–285, 1988.

⁹ MILLER, George A. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review*, v. 63, n. 2, p. 81–97, 1956.

carga: a intrínseca (inerente à complexidade do conteúdo), a extrínseca (gerada pelo design ruim do material, portanto eliminável) e a germinal (o esforço cognitivo produtivo que leva à aprendizagem real). O que Sweller chamou de *expertise reversal effect*¹⁰ completa o diagnóstico: o expert não percebe mais o que é difícil no conteúdo porque deixou de processá-lo como novidade. Ao criar material, tende a criar para seu próprio estado cognitivo, não para o estado em que o aluno se encontra.

Bloom, Ericsson, Polanyi e Sweller chegam ao mesmo diagnóstico por caminhos independentes: a competência no domínio e a competência na sua transferência são dimensões independentes, com mecanismos cognitivos distintos, que raramente coexistem de forma espontânea. Não é accidental. É estrutural.

15.1.2. O Custo do Fracasso e Sua Distribuição Assimétrica

O que os dados de NPS de produtos educacionais revelam vai além da insatisfação: revelam que a promessa e a entrega habitam dimensões distintas na percepção do comprador. Fred Reichheld demonstrou que NPS elevado é função de *earned growth*, crescimento construído sobre entrega real, não sobre autoridade de marca¹¹. Um produto que vende bem em função de posicionamento, mas não entrega transformação verificável, deteriora seu NPS com cada nova turma. O padrão é previsível: primeiros lançamentos sustentados pelo entusiasmo do mercado virgem, seguidos por quedas progressivas à medida que a base de alunos cresce e a taxa de conclusão permanece baixa.

A reputação do expert (o ativo que levou décadas para ser construído) começa então sua erosão. A assimetria entre construção e destruição de reputação, como se verá na seção 15.4, é acentuada o suficiente para

¹⁰ SWELLER, John; AYRES, Paul; KALYUGA, Slava. *Cognitive Load Theory*. Springer, 2011.

¹¹ REICHHELD, Frederick F.; DARNELL, Maureen. *Winning on Purpose: The Unbeatable Strategy of Loving Customers*. Harvard Business Review Press, 2021.

transformar o que parece erro corrigível em dano estrutural de longo prazo.

15.2. A DISTINÇÃO EPISTEMOLÓGICA: DO CONTEÚDO AO SISTEMA DE TRANSFERÊNCIA

A linguagem que a indústria desenvolveu para descrever a digitalização do conhecimento de um expert é sintomática de sua premissa estrutural. Fala-se em “gravar um curso”, “criar um infoproduto”, “lançar um treinamento”. Expressões que compartilham a mesma premissa implícita: o produto existe quando o conteúdo existe. A pergunta que orienta o processo, “o que eu sei que posso gravar?”, é uma pergunta de produtor de conteúdo, centrada no recurso disponível, não no resultado que precisa ser produzido.

A pergunta de um arquiteto de sistemas de aprendizagem inverte essa lógica. E essa inversão não é de método: é de ontologia do objeto.

Grant Wiggins e Jay McTighe, ao desenvolverem o framework *Understanding by Design*, descreveram esse processo como *backward design*¹²: a questão inaugural não é “o que vou ensinar?” mas “o que o aluno precisa ser capaz de fazer quando terminar?”, e, necessariamente, “como saberei que ele é capaz de fazer isso?” Essa sequência muda a natureza do objeto que está sendo construído. Um produto educacional arquitetado a partir do OUTPUT não é uma coleção de conteúdo: é um sistema cujas partes existem em função de um resultado específico que precisa ser produzido de forma verificável.

A diferença pode ser formalizada nos termos que Robert Gagné estabeleceu ao categorizar os tipos de resultados de aprendizagem¹³. Gagné

¹² WIGGINS, Grant; MCTIGHE, Jay. *Understanding by Design*. 2. ed. Association for Supervision and Curriculum Development (ASCD), 2005.

¹³ GAGNÉ, Robert M. *The Conditions of Learning and Theory of Instruction*. 4. ed. Holt, Rinehart and Winston, 1985.

identificou cinco categorias (informação verbal, habilidades intelectuais, estratégias cognitivas, atitudes e habilidades motoras) e demonstrou que cada categoria requer não apenas conteúdo diferente, mas *condições de aprendizagem* diferentes, operacionalizadas nos Nove Eventos de Instrução. Um produto construído a partir do conteúdo raramente especifica qual tipo de resultado pretende produzir: portanto, raramente cria as condições necessárias para produzir qualquer resultado com consistência. Um produto arquitetado a partir do OUTPUT define primeiro o tipo de resultado e depois constrói retroativamente as condições para produzi-lo. A distinção entre os dois processos não é de grau. É de espécie.

A Taxonomia de Bloom ¹⁴, em sua revisão por Anderson e Krathwohl ¹⁵, oferece o mapa de progressão cognitiva que governa a sequência entre unidades dentro desse sistema: não se pode criar antes de aplicar; não se pode analisar sem entender; não se pode avaliar sem ter construído os critérios de avaliação. Quando um produto educacional ignora essa progressão (e a maioria dos cursos digitais opera quase inteiramente nos dois primeiros níveis da taxonomia, Lembrar e Entender), o problema não é que entregue pouco: é que entrega algo estruturalmente incapaz de produzir o resultado prometido.

A mudança que essa distinção exige é, nos termos de Kuhn, uma mudança de visão de mundo ¹. Não uma melhoria incremental: a rejeição da premissa que organiza o processo atual. O expert que aceita essa mudança para de se perguntar “o que eu sei?” e começa a se perguntar “o que o aluno precisa ser capaz de fazer?” A primeira pergunta o coloca no centro do produto; a segunda coloca o aluno. Esse deslocamento de centro de gravidade, sutil na formulação e sísmico na prática, é a transição que define tanto o arquiteto de sistemas de aprendizagem quanto o

¹⁴ BLOOM, Benjamin S. et al. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. Longmans, Green, 1956.

¹⁵ ANDERSON, Lorin W.; KRATHWOHL, David R. (Orgs.). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman, 2001.

escopo desta disciplina.

15.3. A METODOLOGIA APE: INTEGRAÇÃO SISTÊMICA COMO FUNDAMENTO CIENTÍFICO

Propor uma nova metodologia requer justificativa que vai além da combinação de referências. Três condições precisam ser satisfeitas: que os problemas que ela resolve sejam documentados; que as soluções que oferece correspondam a mecanismos empiricamente validados; e, o critério mais exigente, que a integração dos componentes produza capacidade explicativa que nenhum componente oferece isoladamente. É esse terceiro critério que transforma uma metodologia em disciplina. É sobre ele que esta seção se concentra.

A Metodologia APE foi construída sobre um framework sistêmico de cinco nós operacionais: INPUT → GATEKEEPER → BLACK BOX → OUTPUT → FEEDBACK. Cada nó é um ponto de intervenção com fundamentação independente em múltiplos corpos de pesquisa. A questão decisiva não é se cada componente é justificado individualmente. É se a estrutura que os conecta é necessária, ou se os componentes poderiam funcionar de forma isolada com o mesmo resultado. A subseção final desta seção responde a essa questão.

15.3.1. O Nó INPUT: Externalização do *Que Não Se Diz Espontaneamente*

O nó INPUT representa a extração e organização do conhecimento tácito do expert. É onde começa a maioria das falhas de produtos educacionais, porque é o nó que a maioria dos processos de criação de cursos pula inteiramente.

A fundamentação primária está em Polanyi⁷: o conhecimento tácito resiste à externalização espontânea não por relutância do expert, mas por estrutura. O que não se consegue dizer não se consegue dizer por deliberação. Gary Klein operacionalizou o problema com o *Critical Decision Method*¹⁶, desenvolvido originalmente para engenheiros de planta nuclear e bombeiros, profissionais cujas decisões críticas são tomadas de forma tão automatizada que os próprios executores não conseguem articulá-las em tempo real. O CDM funciona porque não exige a verbalização direta do tácito: pede ao expert que reconstrua um caso específico, as opções que considerou e descartou, os fatores que pesaram em cada decisão, os sinais que percebeu no ambiente. Essa reconstrução retrospectiva cria o contexto em que o tácito se torna visível através de sua manifestação em escolhas concretas, sem exigir que o expert saiba que está verbalizando o que antes era implícito.

Sem o nó INPUT, qualquer estrutura instrucional subsequente opera sobre conteúdo incompleto. Gagné sem Polanyi cria design instrucional para o que o expert sabe dizer, não para o que define sua competência real.

15.3.2. O Nó GATEKEEPER: A Qualificação Que Precede o Design

O GATEKEEPER representa o processo de qualificação que precede a construção do produto. Determina se o aluno que entra tem condições de alcançar o resultado prometido com o processo projetado.

A fundamentação central está em Wiggins & McTighe⁹: o OUTPUT precisa ser definido antes de qualquer decisão sobre conteúdo, com precisão suficiente para ser verificável. Mas a qualificação do OUTPUT também exige diagnóstico do aluno-alvo. Eugene Schwartz mapeou que o nível de consciência do comprador (o estágio de sofisticação com que

¹⁶ KLEIN, Gary A. *Sources of Power: How People Make Decisions*. MIT Press, 1998.

ele entende o problema que tem e a solução disponível) determina radicalmente o que é possível entregar¹⁷. Um produto que assume nível de consciência avançado num público com nível de consciência elementar não entregará resultados: não porque o produto seja ruim, mas porque o GATEKEEPER falhou em qualificar a entrada do sistema.

Malcolm Knowles havia estabelecido os princípios andragógicos que tornam esse diagnóstico não-trivial¹⁸: adultos aprendem em função de prontidão, aprendendo o que percebem como relevante agora, não o que poderá ser relevante futuramente. Um produto que não verifica se o aluno está no momento de prontidão adequado ao seu conteúdo está apostando contra a estrutura da aprendizagem adulta.

15.3.3. O Nó BLACK BOX: Onde a Engenharia de Transferência Ocorre

O BLACK BOX é onde a transferência de fato acontece. É também onde se concentra a maior parte da literatura de fundamentação. Três corpos de pesquisa governam as decisões de design nesse nó, e sua interação não é opcional: cada um constrange o espaço de soluções dos outros dois.

Os Nove Eventos de Instrução de Gagné¹⁰ estabelecem a estrutura sequencial de cada unidade de aprendizagem. A sequência não é arbitrária: cada evento ativa o processamento cognitivo necessário ao próximo. Omitir o evento 3 (ativação de conhecimento prévio) antes do evento 4 (apresentação do conteúdo) não é pedagogicamente subótimo. É uma violação das condições que a memória de trabalho precisa para processar nova informação sem sobrecarga. A Teoria da Carga Cognitiva de Sweller¹¹ estabelece as restrições de design que governam cada

¹⁷ SCHWARTZ, Eugene M. *Breakthrough Advertising*. Bottom Line Books, 2004. (Obra original publicada em 1966.)

¹⁸ KNOWLES, Malcolm S.; HOLTON, Edward F.; SWANSON, Richard A. *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. 8. ed. Routledge, 2015.

elemento dentro dos eventos: quanto de novidade pode ser introduzido por etapa, como estruturar exemplos resolvidos para reduzir a carga extrínseca, como calibrar dificuldade para maximizar a carga germinal. Os dados de retenção sistematizados por Glasser ¹⁹ determinam, por fim, a proporção e sequência das modalidades: leitura retém 10%, ouvir 20%, ver e ouvir 50%, discussão 70%, prática 80%, ensinar outros 95%. Escolha de modalidade é decisão de engenharia com consequências quantificáveis, não preferência estética do produtor.

A síntese operacional desse nó pode ser expressa com precisão: Gagné define o que precisa acontecer; Sweller define quanto pode acontecer por vez; Glasser define como deve acontecer para que seja retido.

15.3.4. Os Nós OUTPUT e FEEDBACK: Fechamento do Sistema

O nó OUTPUT não é apenas o destino declarado. É o sistema de verificação que determina se o produto entregou o que prometeu. Lincoln Murphy introduziu aqui a distinção mais relevante: a diferença entre satisfação e sucesso ²⁰. Um aluno satisfeito avalia positivamente a experiência de consumir o produto; um aluno bem-sucedido alcançou o resultado que motivou a compra. Esses dois estados são independentes e, em produtos educacionais mal arquitetados, frequentemente dissociados. O OUTPUT da APE não é definido em termos de conteúdo entregue: é definido em termos do estado verificável do aluno ao final do processo, o que ele é capaz de fazer que antes não era capaz de fazer.

O nó FEEDBACK fecha o sistema e é o mecanismo que permite que ele aprenda. Peter Senge descreveu os *learning loops* como a estrutura fundamental de sistemas que melhoram sistematicamente ²¹: o feedback não como avaliação pós-hoc, mas como elemento constitutivo que habi-

¹⁹ GLASSER, William. *The Quality School: Managing Students Without Coercion*. Harper & Row, 1990.

²⁰ MURPHY, Lincoln. *Customer Success: The Definitive Guide*. Gainsight, 2016.

²¹ SENGE, Peter M. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday/Currency, 1990.

lita a autorregulação contínua. Fred Reichheld forneceu a métrica operacional: o NPS, quando aplicado em pontos estratégicos da jornada (não apenas ao final), revela não apenas a percepção global, mas os nós de maior atrito no processo de transferência ⁸.

15.3.5. Por Que a Integração é Necessária, Não Opcional

A força da APE como disciplina não está em cada componente isolado. Está na impossibilidade de substituir a estrutura inteira por qualquer subconjunto dela. Gagné sem Polanyi cria design instrucional para conteúdo incompleto. Wiggins & McTighe sem Sweller define outputs corretos mas os torna inacessíveis por sobrecarga cognitiva. Murphy sem Bloom confunde sucesso com alcance de resultado de baixo nível cognitivo: o aluno “sabe” mas não “aplica”. Reichheld sem Senge gera dado sem ciclo de melhoria. Cada ausência no sistema produz um ponto de falha previsível e localizado. A disciplina existe porque o problema que ela resolve exige, estruturalmente, todos os componentes. Nenhum subconjunto deles é suficiente.

15.4. O POSICIONAMENTO HISTÓRICO: A ERA DA ENTREGA VALIDADA

A história do capitalismo de marca pode ser lida como a história da crescente independência entre promessa e entrega. A publicidade moderna, desde sua consolidação nos anos 1950, desenvolveu sofisticação técnica para criar percepções que antecediam (e com frequência substituíam) a experiência real com o produto. David Ogilvy observou, com precisão que permanece válida, que o papel da publicidade era criar uma imagem de marca que se tornasse o produto real na mente do consumidor

²². Durante décadas, essa separação foi operacionalmente sustentável: a assimetria de informação entre produtor e consumidor era grande o suficiente para que a promessa sobrevivesse ao contato com a realidade.

A internet destruiu essa assimetria de forma progressiva e irreversível. Avaliações em plataformas, comunidades fechadas onde alunos compararam experiências, NPS coletado e publicado: esses mecanismos criaram um mercado de informação de entrega que funciona em paralelo ao mercado de informação de promessa. O expert que constrói autoridade sobre posicionamento e domínio do conteúdo, mas não entrega transformação verificável, enfrenta agora um mercado que não apenas percebe a dissonância, mas a documenta e a distribui em escala.

Essa é a transformação estrutural que justifica o conceito de Era da Entrega Validada. Não é uma tendência. É uma mudança nas condições de funcionamento do mercado com implicações diretas para a estratégia de qualquer produto premium.

A reputação do expert é, nesse contexto, o ativo mais concentrado e mais frágil. Concentrado porque nenhum outro ativo tem o mesmo efeito multiplicador sobre precificação, acesso a oportunidades e velocidade de crescimento. Frágil porque é o único que não pode ser recuperado diretamente: pode-se reconstruir capital financeiro e substituir infraestrutura tecnológica, mas reputação deteriorada por falha de entrega documentada segue uma trajetória de recuperação longa e incerta. Estudos sobre percepção de marca em serviços profissionais quantificam essa assimetria: uma experiência negativa documentada requer entre 7 e 12 experiências positivas para ser compensada na percepção do mercado ²³.

A APE responde a esse contexto com uma estrutura, não com uma promessa diferente. O princípio operacional é que a entrega é arquiteta-

²² OGILVY, David. *Confessions of an Advertising Man*. Atheneum, 1963.

²³ ANDREASSEN, Tor Wallin; LINDESTAD, Bodil. Customer Loyalty and Complex Services. *International Journal of Service Industry Management*, v. 9, n. 1, p. 7–23, 1998. <https://doi.org/10.1108/09564239810199923>

tada antes da promessa: o OUTPUT precisa ser definido com precisão suficiente para ser verificável antes de qualquer comunicação comercial; a jornada de aprendizagem precisa estar estruturada de forma que a taxa de conclusão esperada e o resultado alcançável pelo aluno mediano sejam conhecidos pelo produtor antes da primeira venda. Não é padrão elevado por idealismo pedagógico. É estratégia de proteção de reputação num mercado de informação simétrica. O expert que entra no mercado digital sem essa arquitetura está apostando que a distância entre promessa e entrega não será percebida. É uma aposta que o mercado atual não honra.

15.5. O FUTURO DA DISCIPLINA: FERRAMENTAS NOVAS PARA PROBLEMAS ANTIGOS

Toda disciplina atravessa um momento em que seus problemas centrais ficam suficientemente bem definidos para que ferramentas externas possam ser adaptadas a resolvê-los com maior escala ou precisão. A física newtoniana precisou criar o cálculo diferencial. A genética molecular precisou desenvolver técnicas de sequenciamento. A Arquitetura de Produtos Educacionais está nesse ponto de inflexão, e a ferramenta com maior potencial transformador é a inteligência artificial generativa, não como substituto do processo, mas como amplificador de sua etapa mais custosa e mais crítica: a extração do conhecimento tácito.

O problema que Polanyi descreveu em 1966⁷ e que Klein tentou operacionalizar através do CDM¹⁴ é, fundamentalmente, um problema de protocolo de entrevista: criar as condições conversacionais que forcem a externalização do que o expert sabe fazer mas não sabe dizer. O CDM é eficaz, mas intenso. Uma sessão completa dura entre 2 e 4 horas por

domínio de decisão e requer um entrevistador treinado para reconhecer os momentos em que o expert está descrevendo automatismos disfarçados de lógica explícita. Sistemas de IA capazes de conduzir entrevistas estruturadas segundo esse protocolo (com capacidade de processar e reorientar em tempo real baseados nas respostas) representam a possibilidade de escalar o processo de extração sem degradar sua qualidade. Não se trata de substituir o julgamento humano na interpretação do que foi extraído: esse julgamento permanece insubstituível. Trata-se de automatizar a parte estrutural e repetitiva (o protocolo de perguntas, a identificação de inconsistências, a geração de contra-exemplos que forcem a articulação do implícito).

A segunda fronteira aberta pela IA é mais consequente em termos de impacto na aprendizagem: a personalização adaptativa do OUTPUT em escala. O Efeito 2 Sigma que Bloom documentou em 1984 ⁵ permanece, quarenta anos depois, como o problema não resolvido mais importante do design instrucional. A pergunta que Bloom levantou (como escalar o que apenas a tutoria individual consegue produzir?) não foi respondida pelos MOOCs, que produziram escala sem personalização, nem pela gamificação, que produziu engajamento de curto prazo sem aprofundamento cognitivo. Sistemas que ajustam a sequência, a dificuldade e a modalidade do conteúdo em função do desempenho e do perfil cognitivo de cada aluno são a única abordagem que oferece, em princípio, a possibilidade de aproximar o resultado do ensino em escala do resultado da tutoria individual.

Convém, no entanto, delimitar o que a IA não resolve: porque a euforia tecnológica tende a criar expectativas que a tecnologia não cumpre, seguidas de ceticismo que rejeita o que ela genuinamente oferece. IA não resolve de forma autônoma o problema da externalização do conhecimento tácito: o processo é inerentemente interativo e exige que

o expert esteja presente e disposto a reconstruir sua prática em termos que possam ser processados externamente. IA não resolve o problema do design do OUTPUT. A definição do que o aluno precisa ser capaz de fazer, com qual nível de precisão e em qual contexto de aplicação, é um julgamento que requer compreensão profunda do domínio e do mercado, compreensão que não emerge de padrões estatísticos em corpora de texto. IA não resolve, por fim, o problema da confiança: em produtos de alto valor, a autoridade do expert é condição de transferência, não decoração. Autoridade não é delegável a uma interface, por mais sofisticada que seja.

O que a IA resolve, e resolve com potência crescente, são os gargalos operacionais que hoje tornam a criação de produtos educacionais bem arquitetados inacessível para a maioria dos experts: o custo de tempo da extração estruturada, a complexidade de personalizar percursos para diferentes perfis dentro do mesmo produto, e a geração de feedback específico o suficiente para ser pedagogicamente útil em escala.

A Arquitetura de Produtos Educacionais está posicionada para integrar essas ferramentas sem abrir mão do que a distingue de abordagens tecnológicas sem fundamento pedagógico: o rigor da engenharia de transferência, a centralidade do OUTPUT verificável, e o reconhecimento de que o conhecimento de um expert (tácito, complexo, construído ao longo de anos de prática deliberada) não é um arquivo que se transfere, mas um sistema que se reconstrói. A reconstrução sistemática desse sistema, em formatos que outros possam percorrer, é o trabalho que define esta disciplina. É o trabalho que, no estado atual do mercado, está por ser feito.

16

O EXPERT QUE ESCALA SEM PERDER O QUE É

16.1. Síntese, Convite e Responsabilidade

16.2. O QUE ESTÁ EM JOGO

Existe uma forma de morte que não aparece em nenhum registro oficial.

Não envolve o fim de um corpo, não tem data, não gera documentos nem acumula em nenhum registro institucional. É a extinção do que alguém sabia: o acervo de competência construído ao longo de anos de prática repetida, de erros que custaram caro e de acertos cuja causa exata nunca chegou a ser completamente articulada. Quando o expert que carrega esse acervo se aposenta, fica doente, muda de trajetória ou simplesmente deixa de atuar, o que sabia não migra para ninguém: dissolve-se. O mundo fica ligeiramente mais pobre num registro que não tem obituário.

Michael Polanyi documentou o mecanismo subjacente a esse processo em 1966. Ao identificar que a competência real é majoritariamente tá-

cita, operando abaixo do limiar da articulação verbal e condensando décadas de prática em julgamentos que parecem intuitivos mas que são, na realidade, estruturalmente organizados por representações mentais invisíveis ao próprio portador, Polanyi produziu o diagnóstico mais preciso já formulado sobre por que o conhecimento de alto nível resiste à transferência ¹. O expert não sabe o que sabe, não por negligência reflexiva, mas porque a natureza do conhecimento profundo é exatamente essa: quanto mais consolidado, menos conscientemente acessível. A maestria é, em certa medida, o estado em que o conhecimento deixou de exigir atenção deliberada para operar com precisão.

Gary Klein chegou ao mesmo fenômeno por outro caminho. Estudando como especialistas de alto desempenho, bombeiros, médicos de emergência, pilotos militares, tomam decisões em condições de alta pressão e tempo mínimo, Klein encontrou que nenhum deles descrevia seu processo decisório em termos de análise comparativa de alternativas ². Eles reconheciam padrões acumulados em anos de exposição a situações similares, e agiam antes que qualquer processo verbal de avaliação tivesse tempo de se completar. O que Klein chamou de *Naturalistic Decision Making* era, no fundo, conhecimento tácito em operação plena: eficaz, veloz, quase impossível de externalizar sem um protocolo específico que reconstrua artificialmente a situação que o ativa.

Cada expert que acumulou competência real num campo e não a estruturou para transferência está, involuntariamente, retendo do mundo algo que o mundo precisa. O médico que desenvolveu um protocolo clínico que funciona melhor do que o padrão mas nunca o documentou de forma que outros clínicos pudessem aprender a aplicá-lo. O engenheiro que reconhece, numa curva de dados específica, a assinatura de uma falha iminente antes de qualquer instrumento a detectar, e que

¹ POLANYI, Michael. *The Tacit Dimension*. Doubleday, 1966.

² KLEIN, Gary A. *Sources of Power: How People Make Decisions*. MIT Press, 1998.

carrega esse reconhecimento apenas na memória tácita de quem passou anos olhando para curvas semelhantes. O educador com taxa de sucesso anormalmente alta com alunos que outros professores consideram inaptos, que atribui seus resultados vagamente à experiência, sem jamais ter articulado o conjunto de micro-adaptações pedagógicas que o distingue.

Em todos esses casos, o conhecimento existe. A estrutura de entrega, não.

Anders Ericsson demonstrou, com um nível de rigor metodológico que outros estudos sobre desempenho humano raramente alcançaram, que a expertise de alto nível é produto de prática deliberada estruturada ao longo de anos, não de talento inato ³. A implicação dessa descoberta raramente é explorada em sua consequência mais incômoda: se a competência é construída por um processo que pode ser estudado, mapeado e replicado, então quem chegou tem a responsabilidade de tornar esse processo suficientemente visível para que outros possam percorrê-lo com eficiência comparável: não como atalho que elimine o esforço necessário, mas como arquitetura que elimine o desperdício de esforço que não converge para resultado.

A responsabilidade intelectual de quem chegou não se encerra com o domínio do campo. Inclui o dever de arquitetar a transferência, não como ato de generosidade opcional, ainda que seja isso também, mas como obrigação epistêmica: o conhecimento que não se transfere perece com seu portador, e sua perda é irreversível de uma forma que a perda de um ativo financeiro nunca é.

³ ERICSSON, K. Anders; KRAMPE, Ralf T.; TESCH-RÖMER, Clemens. The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review*, v. 100, n. 3, p. 363–406, 1993.

16.3. O QUE ESTE LIVRO NÃO DISSE

Nenhuma metodologia resolve todos os problemas de um campo.

Este livro apresentou um conjunto de princípios derivados de pesquisa em epistemologia aplicada, ciência cognitiva, design instrucional e teoria da aprendizagem adulta, produzida ao longo de seis décadas por pesquisadores que trabalhavam em campos distintos sem coordenação entre si. A convergência de descobertas independentes sobre a mesma estrutura fundamental é o sinal mais confiável de que algo verdadeiro está sendo capturado. Mas princípios não são receitas. Uma receita específica exatamente o que fazer, em que ordem, com que ingredientes, por quanto tempo. Produz o mesmo resultado toda vez que é seguida à risca, independentemente de quem a executa. Uma metodologia fornece a estrutura de raciocínio que permite tomar as decisões corretas em contextos específicos que nenhum manual poderia antecipar integralmente. A diferença não é de grau. É de natureza.

Malcolm Knowles documentou que adultos trazem para qualquer processo de aprendizagem um repertório de experiências que funciona simultaneamente como recurso e como filtro ⁴. Esse repertório é singular: dois profissionais com a mesma formação formal e o mesmo tempo de experiência num campo carregam mapas de conhecimento tácito radicalmente diferentes, porque suas trajetórias específicas, os tipos de problemas que enfrentaram, os mentores que tiveram ou não tiveram, os erros cometidos em momentos distintos de desenvolvimento, produziram estruturas cognitivas que não são comparáveis. O que isso significa para o design de um produto educacional é que a aplicação da APE nunca é idêntica para dois experts distintos. O framework é universal; a aplicação, singular.

O protocolo de externalização do conhecimento tácito que a meto-

⁴ KNOWLES, Malcolm S. *The Adult Learner: A Neglected Species*. Gulf Publishing, 1973.

dologia propõe, derivado do trabalho de Klein sobre como especialistas articulam o raciocínio quando confrontados com casos críticos, produz resultados estruturalmente diferentes para um cirurgião, um engenheiro de software e um negociador de alto nível ². O cirurgião opera com conhecimento sensoriomotor de precisão altíssima que resiste particularmente à verbalização: a mão que sabe quando apertar e quando soltar não descreve o que sabe com facilidade. O engenheiro de software carrega padrões de reconhecimento de problemas em sistemas complexos que se acumularam em camadas ao longo de projetos, cada um adicionando uma camada de refinamento ao mapa de como sistemas falham. O negociador tem um mapa de microdinâmicas relacionais que opera em tempo real com latência quase zero, lendo variações de tom e postura que nunca foram catalogadas porque o processo de catalogação interromperia o próprio ato de negociar. Externalizar esses três perfis de expertise exige instrumentos distintos dentro do mesmo framework.

John Sweller demonstrou que a carga cognitiva intrínseca de um conteúdo, aquela que deriva da complexidade inerente do que está sendo aprendido e que não pode ser eliminada sem simplificar o que não deveria ser simplificado, varia enormemente entre domínios ⁵. O design de sequenciamento de um produto educacional em medicina precisa respeitar estruturas de pré-requisito muito mais rígidas do que o design de um produto em liderança organizacional, onde as interdependências são menos lineares. A teoria da carga cognitiva é a mesma nos dois casos; a aplicação, necessariamente distinta.

Grant Wiggins e Jay McTighe argumentaram, com décadas de evidência empírica de implementação, que a definição do resultado desejado é muito mais difícil do que parece ⁶. Exige que o designer saiba com pre-

⁵ SWELLER, John. Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, v. 12, n. 2, p. 257–285, 1988.

⁶ WIGGINS, Grant; MCTIGHE, Jay. *Understanding by Design*. Association for Supervision and Curriculum Development, 1998.

cisão o que o aprendiz precisa ser capaz de fazer numa situação real de aplicação, o que por sua vez exige conhecimento do contexto de aplicação que frequentemente não está disponível antes de o produto ter sido entregue para uma primeira turma. O produto educacional que nunca chegou a uma turma real tem hipóteses, baseadas em teorias sólidas, sobre o que deveria funcionar. Não tem dados sobre o que funciona.

A diferença entre hipótese e evidência é a primeira turma.

Fred Reichheld demonstrou que o NPS coletado noventa dias após a conclusão de um programa é o indicador mais confiável de transformação real ⁷. O NPS da primeira turma de qualquer produto educacional é, quase invariavelmente, menor do que o das turmas subsequentes. Não porque a primeira turma seja menos capaz, mas porque o produto está sendo calibrado em tempo real: o designer descobre onde a progressão de conteúdo tem lacunas, onde o nível de dificuldade está desalinhado com o ponto de partida real dos aprendizes, onde os mecanismos de feedback são insuficientes para o tipo de transformação que o produto promete. A primeira turma ensina o que nenhuma teoria consegue antecipar com completude.

Não existe produto educacional perfeito ao primeiro lançamento. Existe produto projetado com a estrutura correta para melhorar sistematicamente, turma a turma, a partir de dados que apenas a entrega real pode gerar. Isso não é uma limitação da metodologia: é a natureza do design de qualquer sistema complexo. O produto educacional que nunca chegou a uma turma real está em desenvolvimento perpétuo, que é outra forma de dizer que não existe.

Há um paradoxo que emerge frequentemente no processo de externalização de conhecimento tácito: o expert que mais tem a transferir é frequentemente o que mais resiste ao processo de externalização, não

⁷ REICHHELD, Frederick F. The One Number You Need to Grow. *Harvard Business Review*, v. 81, n. 12, p. 46–54, 2003.

por má vontade, mas por uma combinação de fatores que inclui a dificuldade cognitiva de articular o que opera tacitamente, a insegurança quanto à possibilidade de o produto final capturar adequadamente a profundidade do que sabe, e uma preocupação legítima com reputação que leva muitos experts a preferirem não entregar a entregar algo que consideram incompleto ou superficial em relação ao que realmente possuem.

Essa resistência é parte do processo, não um obstáculo a ele. Reconhecê-la como previsível e estruturalmente compreensível é o primeiro passo para atravessá-la.

16.4. O CONVITE

Três perfis distintos chegaram até aqui.

O primeiro é o expert que ainda não começou: que tem clareza sobre o que sabe e sobre o valor que esse saber poderia ter para outros, mas que nunca deu o passo de estruturar a transferência. As razões variam. Pode ter tentado algumas vezes e o processo pareceu esmagadoramente complexo, sem ponto de entrada claro. Pode ter percebido que o mercado de educação online, com seu excesso de promessas e sua taxa de fracasso documentada, é um terreno onde a reputação que levou anos para construir pode ser rapidamente erodida por um produto que não entrega o que promete. Pode simplesmente não ter tido acesso a um framework que tornasse o problema suficientemente inteligível para saber por onde começar.

Para esse expert, o argumento central deste livro é direto: o problema não é o nível do seu conhecimento, não é sua capacidade de ensinar, não é o timing de mercado. É a ausência de arquitetura de entrega. Benjamin Bloom identificou, no problema do 2 Sigma, a questão mais importante

que o campo da educação havia colocado em décadas: como escalar o que apenas a tutoria individual entrega? ⁸. A resposta que este livro propõe não é tecnológica. É uma metodologia de design que replica, na estrutura de um produto educacional escalável, os mecanismos que tornam a tutoria individual eficaz: resultado especificado antes do conteúdo, feedback imediato sobre progressão, ciclos de prática deliberada com dificuldade calibrada ao estado atual do aprendiz. Esses mecanismos não dependem de presença física. Dependem de design intencional. O que você não fez até agora não foi por falta de capacidade. Foi por falta de um mapa do território.

O segundo perfil é o expert que já tentou e falhou. Que investiu tempo e dinheiro num produto educacional que não entregou o que prometia, que gerou NPS baixo ou reclamações que ficaram na memória como evidência de que talvez o problema fosse, afinal, o conteúdo ou a capacidade pessoal de ensinar.

Para esse expert, o problema é de diagnóstico: a solução incorreta falha independentemente de quantas vezes for tentada, e o diagnóstico incorreto garante a solução incorreta. Se o produto falhou porque tinha carga cognitiva extrínseca excessiva, material mal sequenciado que sobrecarregava a memória de trabalho dos aprendizes antes que eles tivessem os pré-requisitos necessários para processar a nova informação, a solução não é simplificar o conteúdo. Sweller documentou extensamente que materiais didáticos desenvolvidos por experts sofrem sistematicamente do problema da maldição do conhecimento: o designer, por saber demais, não consegue mais perceber o que é difícil para quem ainda não sabe ⁵. O problema é estrutural e tem solução estrutural.

Se o produto falhou porque prometia um resultado que não conseguia verificar, a solução não é reduzir a ambição da promessa. É redesenhar o

⁸ BLOOM, Benjamin S. The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, v. 13, n. 6, p. 4-16, 1984.

OUTPUT com a precisão que Wiggins e McTighe propuseram: especificar não apenas o que o aprendiz vai aprender, mas o que o aprendiz vai ser capaz de fazer numa situação real de aplicação ao final do processo, e construir a evidência de aprendizagem que demonstra, de forma verificável, que esse resultado foi alcançado ⁶. A promessa não era grande demais. Era imprecisa demais.

Se o produto falhou porque os alunos desengajavam antes de concluir, a solução não é encurtar o curso. É identificar em qual ponto da jornada o desengajamento ocorria sistematicamente e entender o que acontecia naquele ponto: falta de relevância imediata (Knowles documentou que adultos precisam perceber a utilidade do que estão aprendendo para continuar investindo atenção cognitiva), ausência de marcos de progressão visíveis, ou simplesmente um ponto de dificuldade onde o produto não oferecia suporte suficiente para que o aprendiz atravessasse o obstáculo sem desistir ⁴.

O terceiro perfil é o que este livro chama de arquiteto de conhecimento: o profissional que vai dedicar parte ou toda sua trajetória a ajudar outros experts a estruturarem a transferência do que sabem. Um papel que não tem nome consolidado ainda, com elementos de consultor, de designer instrucional, de produtor de conteúdo, mas fundamentalmente distinto de todos esses porque seu objeto central não é o conteúdo a ser transferido: é a arquitetura do sistema que permite que esse conteúdo chegue, com integridade, a quem precisa dele.

Peter Senge argumentou que a capacidade de transferir conhecimento internamente é a vantagem competitiva mais sustentável que uma organização pode cultivar: diferentemente de produtos, tecnologias ou posições de mercado, a capacidade de aprender coletivamente não pode ser facilmente replicada por competidores porque não existe fora das pessoas e dos processos que a produzem ⁹. O que Senge identificou no

⁹ SENGE, Peter M. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday, 1990.

nível organizacional, este livro propõe no nível do produto educacional individual: a capacidade de transferir conhecimento com consistência, de forma que produza resultado verificável no aprendiz, é o que distingue um produto educacional premium de um registro de conhecimento bem-intencionado.

O arquiteto de conhecimento não trabalha apenas no plano da estratégia organizacional: trabalha no plano da preservação epistêmica.

Paulo Freire argumentou que o ato de ensinar não é neutro: escolhe, implicitamente ou explicitamente, quem tem acesso ao que se sabe e quem não tem¹⁰. Num contexto em que plataformas digitais tornaram o alcance de um produto educacional potencialmente global, essa escolha tem dimensões que vão muito além do mercado de nicho em que a maioria dos experts inicialmente situa seu trabalho. O conhecimento que um expert estrutura e entrega com consistência pode, dependendo do campo e da qualidade do produto, alcançar pessoas que não teriam acesso a esse tipo de expertise de nenhuma outra forma.

O conhecimento que um expert desenvolveu ao longo de anos de prática não pertence apenas a ele. Pertence, em parte, ao campo em que foi construído, às pessoas que tornaram possível a trajetória que o produziu, e ao conjunto de aprendizes futuros que poderiam beneficiar-se dele se chegasse a eles de forma estruturada. O conhecimento especializado é sempre, em alguma medida, colaborativo em sua origem, mesmo quando individual em sua expressão. E essa origem cria uma responsabilidade recíproca.

O expert que não estrutura a transferência do que sabe não está apenas perdendo uma oportunidade comercial. Está interrompendo uma cadeia que, se mantida, torna o campo progressivamente mais capaz, as práticas progressivamente mais eficazes, os resultados progressivamente melhores para quem depende da expertise que aquele campo produz.

¹⁰ FREIRE, Paulo. *Pedagogia da Autonomia: Saberes Necessários à Prática Educativa*. Paz e Terra, 1996.

O expert que escala sem perder o que é não é o expert que aprendeu a vender cursos. É o expert que entendeu que o que o torna valioso não é a exclusividade do que sabe, mas a capacidade de transferir com fidelidade suficiente para que outros possam construir sobre isso, e que reconheceu que sua expertise, ao ser estruturada para transferência, não se dilui. Frequentemente, o processo de externalização do conhecimento tácito é o que finalmente torna o expert plenamente consciente do que construiu.

A escala, no contexto deste livro, não é uma métrica de vaidade. É a medida de quantas pessoas se beneficiam de um conhecimento que, sem a arquitetura adequada, ficaria confinado a quem teve a sorte de estar no lugar certo, no momento certo, próximo o suficiente ao expert que o possuía.

O convite é esse.

Não para criar cursos. Para arquitetar transferência.

CONTINUE SUA JORNADA

Este livro apresentou a ciência.

Agora é hora de aplicar.

A plataforma **eBuz** transforma a Metodologia APE em fluxos interativos guiados por Inteligência Artificial.

Você insere seu conhecimento.

A IA estrutura, valida e gera seus produtos educacionais.

app.ebuz.com.br

Acesse gratuitamente · Experimente por 24 horas · Sem cartão

*“A distância entre saber e ensinar é um problema de engenharia,
não de talento.”*

— Metodologia APE

