



Para este número seleccionámos

Racionalidade e irracionalidade dos estudos comparativos internacionais

Christine Keitel e Jeremy Kilpatrick

A nossa escolha para este número é a tradução integral do texto *The Rationality and Irrationality of International Comparative Studies*, de que são autores Christine Keitel e Jeremy Kilpatrick e que constitui o capítulo 16 do livro *International comparisons in mathematics education (studies in mathematics education series 11, pp. 241-256)*, editado por G. Keiser, E. Luna e I. Huntley e publicado pela Falmer Press (Londres) em 1998.

Os autores são duas das mais prestigiadas figuras internacionais da educação matemática e conhecem particularmente bem a problemática abordada no texto, uma vez que foram ambos consultores do Third International Mathematics and Science Study — TIMSS. Neste artigo, muito actual e importante, eles questionam concepções muito generalizadas sobre os propósitos, o alcance e a validade dos grandes estudos comparativos internacionais do desempenho matemático dos alunos, bem como a utilização que é geralmente feita dos resultados desses estudos.

Introdução

Os estudos comparativos internacionais chegam a dominar o discurso educativo em muitos países, quando educadores, decisores educativos e políticos equacionam o que deve ser feito para melhorar o seu sistema educativo. Os resultados destes estudos são aceites, em muitos casos, como fornecendo evidências científicas indiscutíveis sobre o que sabem os alunos dos países comparados, assim como sobre o grau de qualidade dos currículos e da forma como os professores ensinam.

Neste artigo, vamos examinar algumas características dos estudos comparativos internacionais em Matemática. Consideraremos em primeiro lugar como é encarada nesses estudos a natureza problemática do currículo. Em seguida, focaremos o estudo comparativo mais recente e amplo, o TIMSS, considerando questões que se levantam pela sua grande amplitude e as respectivas consequências. Iremos olhar em particular para três estudos complementares que foram concebidos para permitir interpretações mais claras dos dados do teste, mas que, em vez disso, parecem conduzir a más interpretações e contradições. Concluiremos com várias questões

que os investigadores desta área parecem não ter levantado e com algumas reflexões sobre a natureza, aceite *a priori* sem discussão, dos estudos comparativos internacionais em geral.

Investigando os currículos de Matemática

O tratamento dos currículos da Matemática escolar em investigações comparativas internacionais é uma história de esforços crescentes para levar em conta aspectos da complexidade curricular. Contudo, também é uma história de fracasso contínuo na tentativa de aprofundar e de questionar os pressupostos do que se entende por currículo.

As investigações comparativas internacionais têm assumido duas formas. Uma delas procura examinar o modo como a compreensão matemática, a capacidade, o sucesso, o progresso ou, para utilizarmos um termo mais neutro, o desempenho dos alunos diverge entre sistemas educativos¹, no que se refere a um dado tópico curricular — por exemplo, problemas de palavras ou cálculos de adição e subtração de números inteiros. Nestes estudos, o currículo não é analisado em si mesmo. Em vez disso, assume-se que o tópico está

presente em todos os sistemas estudados e, portanto, pode ser considerado como uma constante não problemática.

A segunda forma de investigação, mais comum, procura verificar o sucesso ou progresso ou, mais geralmente, o desempenho, como um todo, apesar de quase sempre se referir somente a um dado nível etário ou ano de escolaridade. Os estudos deste segundo tipo são inevitavelmente confrontados com o problema de examinar, ou pelo menos caracterizar, os currículos dos sistemas educativos que estão a ser considerados.

A abordagem mais comum tem sido a de estabelecer uma espécie de estrutura para analisar os currículos. Desenvolvem-se então medidas de desempenho, para encaixarem nessa estrutura. Nas investigações mais antigas, era típica a utilização de matrizes de processos-conteúdos, para construir e classificar itens de desempenho. Reconhecendo que as variáveis de "processo" não pareciam operar através dos conteúdos, de forma a fazerem ressaltar diferenças interpretáveis e que a colocação de um item numa posição única vai violentar o que acontece, tanto na avaliação como no currículo, investigadores recentes (Lapointe, Mead e



Phillips, 1989; *Survey of Mathematics and Science Opportunities*, 1993, Apêndice B) substituíram o processo por dimensões como a *expectativa de desempenho* e o *nível de proficiência*, acrescentaram dimensões como a *perspectiva* para enquadrarem objectivos curriculares não cognitivos e descreveram os itens de desempenho utilizando mais do que uma categoria para cada dimensão.

Uma segunda tentativa para lidar com a complexidade curricular foi a de distinguir entre o currículo planeado e o currículo implementado. Uma distinção entre o *currículo planeado* (tal como está representado em documentos oficiais, manuais, ou em ambos) e o *currículo implementado* (normalmente medido através de questionários aos professores) foi feita no *Second International Mathematics Study* — SIMS (Travers e Westbury, 1989). A distinção já tinha sido antecipada no *First International Mathematics Study* — FIMS (Husén, 1967) — pela utilização de classificações dos professores das *oportunidades de aprendizagem* dos conteúdos relativos a cada item testado. Apesar dos termos *planeado* e *implementado* transportarem a infeliz conotação de que as únicas intenções que contam são as oficiais, e de que os professores não passam de meros executores que implantam no terreno planos de outras pessoas, esta distinção foi útil, na medida em que ajudou a distinguir o planeado do que é a realidade curricular.

Infelizmente, tal como Hans Freudenthal (1975) fez notar há alguns anos, em última análise os professores podem muito bem classificar a dificuldade do item, em vez de avaliar se foi ensinado aos alunos o conteúdo pressuposto nesse item. Além disso, relacionar as “oportunidades de aprendizagem” com o “currículo implementado”, para testar o desempenho, não aborda directamente a questão de saber qual a extensão do currículo que foi aprendida pelos alunos.

Quando damos a todos, num grupo de alunos oriundos de vários sistemas,

um conjunto comum de itens de desempenho, fica assim definido um currículo idealizado que serve de modelo, em contraponto com o desempenho que os alunos têm dentro de cada sistema. As medidas relativas às oportunidades de aprendizagem ou ao currículo implementado tornam-se medidas do modo como o currículo idealizado foi aprendido ou implementado. Independentemente da forma como o grupo de especialistas delineou o currículo idealizado, seleccionando um conjunto de itens de desempenho para o estudo, é impossível, tal como Freudenthal (1975) muito bem assinalou, construir instrumentos internacionalmente *equivalentes*. Freudenthal fez notar que seria possível, embora muito difícil, construir instrumentos que apresentassem relações internacionalmente *comparáveis* com os currículos nacionais. Não explicou como tal poderia ser conseguido, mas o que permanece válido é que, não considerando o currículo como uma variável entre sistemas, os estudos internacionais continuam a comparar maçãs com laranjas ou, nas palavras de Torsten Husén (1983), a “comparar o incomparável”.

As investigações comparativas internacionais têm-se tornado cada vez mais sofisticadas. Em conjunto com os julgamentos dos especialistas sobre o modo como o currículo da Matemática deve ser representado internacionalmente, têm sido feitas análises cuidadosas de documentos oficiais e materiais escritos. Foram efectuadas análises a variáveis como o tempo reservado para vários tópicos em diferentes sistemas, a proporção de sistemas que tratam um dado tópico em cada ano, a forma como varia, nos manuais, o espaço concedido a um tópico, e como difere a organização dos manuais nos diferentes sistemas.

Mesmo assim, o currículo internacional idealizado, definido por um conjunto comum de tarefas organizadas por tópicos de conteúdo, continua a ser a norma para medir o desempenho. Não é concedida nenhuma tolerância pelo facto de existirem objectivos, ques-

tões, histórias e contextos que são diferentes entre os currículos de Matemática dos sistemas em estudo. Ninguém aborda realmente em que medida os alunos de um dado sistema estão a aprender o currículo de Matemática que o seu sistema lhes oferece.

Promessas e perigos das comparações

Tal como as observações anteriores sugerem, os estudos comparativos internacionais que utilizam uma grande quantidade de dados empíricos têm sido objecto de crítica, desde que foram efectuados. Alguns críticos insistem no “jogo de números” e propõem outra organização para a análise de dados, outros modelos matemáticos ou outras representações para os resultados. Críticas mais abrangentes referem-se a problemas fundamentais e que não foram objecto de suficiente atenção nas comparações, tal como “comparar o incomparável”, “muitas visões, muitos objectivos, um teste” e “problemas em comparar currículos de diferentes sistemas educativos” (Bracey, 1997a, 1997b; Husén, 1983; Westbury, 1989). Foi notado com frequência que os estudos testam apenas se e como os currículos avaliados se enquadram ou adaptam ao currículo nacional, ou ao que se toma como tal, e que qualquer conclusão para lá disso não é justificável (Bracey, 1996; ver também o comentário de Baker, 1997). Neste sentido, os estudos comparativos internacionais mais conhecidos dos últimos 30 anos, conduzidos pela *International Association for the Evaluation of Educational Achievement* — IEA (FIMS e SIMS), foram muito criticados por não se adequarem à realidade, ou seja, por não produzirem nem usarem dados “reais” que suscitassem uma compreensão do que se passava, e como, na sala de aula. O TIMSS foi o primeiro dos estudos da IEA que tentou complementar os habituais dados dos testes, os quais tinham sido o foco destes estudos empíricos.



Os testes de desempenho e os questionários foram acompanhados por três outros ambiciosos estudos de investigação que se apoiaram em fontes de dados adicionais:

- o *Videotape Classroom Study*, um estudo das aulas de Matemática no 8º ano de escolaridade nos EUA, Japão e Alemanha, dirigido por James Stigler da Universidade da Califórnia em Los Angeles;
- o *Survey of Mathematics and Science Opportunities Study*, um estudo sobre o ensino da Matemática e Ciências em seis países, dirigido por William Schmidt, da Universidade do Estado de Michigan;
- o *Curriculum Analysis Study*, um estudo dos propósitos curriculares, dirigido por William Schmidt, da Universidade do Estado de Michigan.

Todos estes estudos foram extensos e conduzidos em paralelo com o estudo de desempenho do TIMSS. O primeiro (Stiegler et al., 1996; U.S. Department of Education, 1997) implicou o registo em vídeo e tratamento de 231 aulas do 8º ano de escolaridade, nos EUA, Alemanha e Japão. O segundo (Schmidt et al., 1996) necessitou que se efectuassem mais de 120 observações de aulas de Matemática e Ciências em seis países, numa tentativa de caracterizar "uma aula típica de Matemática ou de Ciências" para alunos de nove e treze anos de idade, nesses países. O último (Schmidt, McKnight, Valverde, Houang e Wiley, 1997) incorporou a tentativa de desenvolver um instrumento e uma metodologia para analisar as orientações curriculares, programas e manuais de cerca de 60 países participantes (ver Schmidt, McKnight e Raizen, 1997, para um relatório que aplica a informação dos dois últimos estudos ao contexto dos EUA). Os três estudos requeriam que os investigadores lidassem com uma quantidade sempre crescente de "dados", que deviam ser recolhidos, processados e interpretados; um número cada vez maior de pessoas

envolvidas nestas actividades; a aplicação de um instrumental técnico, continuamente a aumentar; um nunca acabar de pedidos de financiamento para realizar os estudos. O TIMSS foi, a grande distância, o maior e mais caro dos estudos comparativos alguma vez realizados em educação matemática, e estes três estudos aumentaram o seu alcance e complexidade.

A enorme dimensão do estudo levantou imediatamente alguns problemas de cariz novo: quanto mais dados, pessoas e dinheiro eram necessários, mais importantes se revelavam os problemas de responsabilidade e de racionalidade de tal empresa e mais cuidado era preciso colocar na selecção das pessoas que estavam a dirigir e no controlo do seu trabalho. Estas observações provocam mais considerações fundamentais sobre os estudos comparativos, o seu valor geral e os seus riscos. Vamos colocar três questões sobre "quem" está ligado a estes estudos.

Quem dirige o estudo?

A questão das pessoas que dirigem o estudo liga-se aos conhecimentos específicos, especialização e competência dos que estão envolvidos, assim como à sua capacidade e responsabilidade para assumirem tanto uma perspectiva geral como uma orientação não tendenciosa e para manterem simultaneamente um distanciamento claro e uma proximidade em relação aos dados. Essas pessoas necessitam desenvolver uma compreensão comum, entre todos os que têm responsabilidades no estudo.

Nos estudos da IEA, incluindo o TIMSS, ressalta o facto de quase todos os principais responsáveis pela sua condução terem sido investigadores empíricos em educação, especialistas em psicométrie ou em processamento de dados. O problema dos conteúdos da educação matemática ou científica foi tratado como uma questão meramente técnica. Nestes estudos, a característica mais monumental e abordada com mais rigor foi

o tratamento dos dados, depois de recolhidos. Particularmente notável, foi a maneira como foram tratados problemas de validade metodológica, de fiabilidade e de qualidade, adoptando-se um ponto de vista puramente formal. As questões relacionadas com o conteúdo, em todos os seus aspectos, foram normalmente vistas como secundárias. Esta orientação não surpreende, à luz da especialização que estes investigadores têm e da sua preocupação fundamental com a metodologia. No entanto, a predominância das questões metodológicas levanta com certeza sérios problemas, quando se chega ao aspecto concreto de tomar as tarefas matemáticas como sendo os currículos, os questionários aos professores como representando o ensino e os resultados como a base para comparações dos currículos, em relação aos seus contextos culturais e sociais. As afirmações sobre diferenças e semelhanças revelam-se problemáticas, incluindo os juízos sobre o sucesso e o insucesso ou sobre países fracos ou fortes.

Quem paga o estudo?

A questão dos apoios financeiros tem a ver com o facto dos objectivos do estudo serem meramente determinados politicamente ou orientados pela investigação. Por outras palavras, que papel é desempenhado por grupos de pressão, como departamentos governamentais, forças políticas, organizações sociais ou instituições de investigação e como podem os vários países influenciar a actividade concertada dos diferentes grupos? Esta questão também está relacionada com o problema de saber se os países pobres e ricos participam equitativamente na recolha, interpretação e avaliação dos resultados, referentes ao seu próprio país assim como aos outros.

Na nossa cultura ocidental (euro-americana), as comparações internacionais não têm valor político se não forem aceites como "verdadeiros" estudos de investigação. Isto é, os políticos que financiam um tal estudo confiam fortemente que ele se



justifica, em primeiro lugar, pelo seu carácter internacional e, em segundo lugar, pela exigência dos resultados serem inquestionavelmente válidos. Esta condição explica por que razão os países que pagam mais ficam muito preocupados com a qualidade da investigação e investem não só dinheiro, mas também recursos humanos. A consequência óbvia é que os estudos ficam dominados pelos especialistas dos países que forneceram o apoio financeiro e pelos interesses e conhecimentos profissionais desses especialistas. Não é de estranhar que as pessoas colocadas nas posições de liderança mais importantes do TIMSS fossem oriundas dos países ricos que financiaram o estudo, com todas as consequências que essa influência acarreta. As questões de investigação, o idioma e os estudos complementares foram escolhidos pelas entidades financiadoras, da mesma forma que as finalidades foram determinadas principalmente por aqueles que pagaram.

Por outro lado, os políticos não têm grande interesse por estes estudos. Para eles, o lugar relativo do seu país é o aspecto mais importante dos resultados. Querem ter um bom lugar na competição, querem alardear o brilhar do seu país ou os esforços feitos para sobressair em Matemática e Ciências, ou então querem saber por que razão os países mais bem classificados estão nessas posições. Este interesse está directamente relacionado com a competitividade económica. Algumas críticas recentes dos sistemas escolares exageraram as consequências económicas do desempenho fraco em Matemática e Ciências, uma vez que o estado económico de um país é "mais influenciado pelas políticas comerciais do governo, pelos mercados financeiros mundiais e pela deslocação de companhias para países com baixos salários, do que pelo fraco desempenho escolar" (Stedman, 1997, pp. 4-5). Em todo o caso, a questão fulcral para cada país, em particular para os países ricos que contribuíram com a

maior fatia do dinheiro dos estudos comparativos internacionais e que são altamente competitivos na economia mundial, é a sua posição relativa comparada com a dos que lhe ficaram à frente. Este é o único resultado importante. É raro que os políticos se interessem meramente por perceber melhor ou apenas por explorar as possibilidades de investigação oferecidas pelas comparações internacionais. A sua perspectiva não é a de melhorar a educação matemática e científica, particularmente para os que estão em países pobres e mal classificados, mas, antes, a de serem comparados com os que tiveram mais sucesso. O seu interesse está concentrado nos aspectos internos do seu próprio país e, talvez, em procurar pistas que sugiram razões para o desempenho superior dos melhores países.

Este interesse torna-se especialmente claro quando se olha para os estudos complementares, realizados no contexto do TIMSS. O estudo do TIMSS que utilizou as gravações de vídeo das aulas de Matemática limitou-se aos EUA, Japão e Alemanha. Este estudo, tal como aconteceu no inquérito realizado em seis países sobre o ensino e no estudo que analisou o currículo, foi financiado pelo governo dos EUA. Todos eles foram dirigidos apenas por investigadores norte-americanos, com um pequeno apoio de especialistas pertencentes aos outros países participantes. Os investigadores americanos fizeram um esforço assinalável para trabalhar com os membros das equipas dos países cooperantes, mas estes apenas contribuíram parcialmente para o projecto, fornecendo os dados codificados referentes aos seus países, enquanto o processamento e interpretação, mais importantes, foram efectuados pelos membros da equipa americana. Não pretendemos que estas considerações sejam uma crítica ao procedimento. Contudo, é claro que localizar o financiamento, a gestão e a implantação de um estudo

internacional num só país participante traz problemas. Podem passar despercebidos aspectos fracos dos estudos e não se detectarem questões relacionadas com desequilíbrios ou erros.

Outra questão fundamental e relacionada com esta problemática é: qual é a razão científica para a pressão no sentido de obter resultados que levem a um claro escalonamento dos países? Será que este escalonamento fornece automaticamente melhores "resultados de investigação" ou novas compreensões? Os estudos e os seus resultados fornecem uma mudança de perspectiva? O que significa, para o governo alemão ou o dos EUA, o facto das classificações do Irão ou da África do Sul serem tão decepcionantes? Será que os países ricos tencionam fornecer algum apoio para melhorar esta situação? Os estudos desvendam maneiras de, ou realmente apontam para, melhorar a educação matemática e científica nos diferentes países?

Quem controla a construção e divulgação dos resultados?

Apesar de existir o mito enraizado de que "os números não mentem", particularmente entre pessoas que acreditam firmemente em estudos empíricos "objectivos" — e muitas dessas pessoas integram os grupos que se encarregam de realizar estudos comparativos com base em testes e questionários — há hoje uma aceitação alargada de que os dados podem ser recolhidos, tratados, matematizados e interpretados de modos muito diferentes e de que estes processos não estão totalmente imunes a interesses e intenções especiais. Assim, uma questão deveras importante tem a ver com o modo como os resultados dos estudos são "empacotados", distribuídos e levados ao público. Como é que os investigadores podem tratar, e como é que o fazem realmente, os dados "não desejados" ou as interpretações "não habituais"? Quem



influencia este processo, com que razões e através de que meios? O poder essencial está nas mãos da comunidade de investigação que dirige o estudo ou nas dos financiadores, nomeadamente as instituições políticas? As respostas a estas perguntas podem diferir consideravelmente nos países participantes, por existirem diversas tradições de investigação e estruturas políticas, ou por causa de diferentes "jogos de poder" que podem estar em curso a favor ou contra outros países.

As boas posições que certos países conseguiram, nos testes de desempenho dos estudos comparativos internacionais, levaram alguns comentadores de países em posições mais baixas a exagerar os resultados, criando uma mística sobre a educação dos países com melhor classificação. Nos EUA, por exemplo, os professores foram pressionados a voltar a um currículo baseado nos "conhecimentos básicos" em Matemática, para imitarem os sistemas, presumivelmente mais eficazes, dos países asiáticos (Hirsh, 1996). Na Alemanha, artigos na imprensa sublinhavam quase semanalmente, depois da publicação dos resultados do TIMSS, o contraste entre os alunos alemães, que tiveram níveis de desempenho medíocres nos testes, e os japoneses que — alegadamente por terem trabalhado o tipo de resolução de problemas desafiadores mostrados nas aulas gravadas em vídeo — tiveram bons resultados nos testes. Jogando com as preocupações do público e dos políticos sobre estes artigos, um investigador alemão ligado ao TIMSS conseguiu obter milhões de marcos, não para melhorar a prática nas aulas de Matemática ou a formação de professores de Matemática, mas para desenvolver um instrumento mais sofisticado para medir o desempenho matemático dos alunos. Isto aconteceu numa altura em que os orçamentos para as escolas e universidades baixaram dramaticamente e o financiamento para a investigação, em todas as áreas, foi drasticamente reduzido. Contudo, a

Alemanha não foi o único país onde os investigadores ajudaram a organizar a apresentação dos resultados do TIMSS para os meios de comunicação e, em seguida, recolheram os benefícios, quando os governos responderam alarmados, financiando mais investigação.

Aproximação à realidade da sala de aula

Análise curricular no TIMSS

O estudo do TIMSS de análise curricular (Schmidt, McKnight, Valverde et al., 1997) foi pioneiro na realização da ambiciosa tarefa de analisar as linhas orientadoras dos currículos, programas e manuais, desenvolvendo em simultâneo um poderoso instrumento de comparação para a análise. A descrição da metodologia é impressionante e os complicados processos para a recolha de dados são admiráveis. As várias representações gráficas das conclusões relacionadas com o constructo "oportunidades de aprendizagem e de ensino" e dos correspondentes itens unitários, como "tópicos mais importantes" ou "conteúdo principal" são, contudo, difíceis de perceber. Além disso, a aplicação de outros constructos como "expectativas de desempenho" — conhecer, resolver problemas, raciocinar, para mencionar apenas algumas que podem ser mais facilmente entendidas — e "perspectivas (profissionais)" é simplesmente confusa. De facto, não existe nenhum resultado real deste estudo, além da afirmação de que há "várias perspectivas e muitas finalidades", mas o estudo não está relacionado com nenhum resultado da parte principal do TIMSS, o estudo dos testes de desempenho. Desde o início, foram levantadas críticas gerais sobre a metodologia usada e os pressupostos subjacentes. Queremos agora acrescentar algumas considerações adicionais, que deviam ter sido objecto de reflexão mais substancial.

A ideologia predominante. A ideologia que está por detrás da análise curricular consiste em que através do constructo de um desenho de investigação

de "currículo comum" e da utilização de "uma língua comum inglesa", para as questões que deviam ser respondidas pelos participantes nacionais do estudo e para as descrições sistemáticas que daí resultariam, seria possível encontrar regularidades nacionais e práticas distintas passíveis de serem caracterizadas e avaliadas. Como poderia alguém acreditar que, com este procedimento, fosse possível apreender o sentido dos sistemas educativos participantes?

"Faz o teste!" é o grande imperativo subjacente a investigações comparativas. No entanto, como se pode lidar com culturas diferentes com um teste único, em particular com um teste que tinha a pretensão de apreender objectivos diferentes, entre vários sistemas educativos e dentro de cada um deles?

As ficções associadas. Uma das ficções consistiu em considerar que as linhas orientadoras e os manuais, como medidas dos currículos, influenciam a prática na sala de aula de modo "equitativo" ou universal, representam uma fonte predominante do saber dos professores e dominam o seu conhecimento profissional e as suas acções.

Uma segunda ficção residiu em considerar que a ênfase em aspectos mensuráveis e numeráveis do desempenho dos alunos fornece uma informação importante, suficiente e relevante para ser processada. Esta ênfase fez surgir constructos do tipo "oportunidades para aprender e ensinar Matemática", "expectativas de desempenho" ou "perspectivas profissionais". Estes constructos artificiais podem dar aos investigadores mais espaço para a especulação do que para uma interpretação séria. A linguagem da codificação e o que fica codificado podem ter influenciado os dados de modo tão forte que alguns problemas não podiam sequer ser detectados. Especificamente, no projecto de análise curricular, os dados codificados que serviram de base a todo o processamento e às representações gráficas ou numéricas não podiam ser controlados seriamen-



te para todos os países, nem o próprio procedimento de codificação podia ser verificado pela equipa em todos os casos, quando os dados eram finalmente apresentados. A forte predileção dos investigadores pelos jogos de números — todos os tipos de correlações e agrupamentos, cujas relações não admitiam nenhuma interpretação razoável e, por vezes, pareciam ser completamente desprovidas de sentido e absurdas — não levou, contudo, os investigadores a voltarem aos dados, nomeadamente aos manuais e linhas orientadoras originais, e a examiná-los com cuidado. O estudo dependeu grandemente da codificação nacional, dos especialistas nacionais, da sua fluência no inglês e da sua compreensão das finalidades, instrumentos e procedimentos do estudo. A codificação estava intimamente relacionada com as interpretações dos constructos (frequentemente numa língua estrangeira) e estes constructos, por sua vez, já estavam influenciados pela perspectiva dos que os desenvolveram. Deste modo, a análise curricular perdeu-se no meio dos resultados gerais do estudo principal do TIMSS. Os estudos de contexto que deviam ter complementado e apoiado o estudo empírico do desempenho tornaram-se irrelevantes e não ajudaram à interpretação dos dados dos testes. Os resultados dos estudos de contexto podem mesmo ter contrariado os dados do teste.

Uma terceira ficção do estudo de análise curricular residiu no facto de que *comparação* significa colaboração, e não competição, obrigando a uma igualdade dos participantes. Contudo, existia "naturalmente" uma forte hierarquia entre os países grandes e os pequenos, os "distantes" ou os pobres. Claro que na maioria dos casos a equipa de investigação (de língua inglesa) era definida pelos países maiores, sendo estas pessoas quem sobretudo determinava o que se deveria considerar o currículo comum, assim como o idioma comum.

Ensino na sala de aula no TIMSS

O estudo do TIMSS que utilizou o vídeo das aulas foi outra maneira de

olhar para a aula de Matemática, como um lugar onde as oportunidades de ensino e de aprendizagem são determinadas. A característica singular deste estudo era a tentativa de chegar, tão perto quanto possível, à aprendizagem e ao ensino concretos, e tornar possível que a situação concreta fosse observável posteriormente, por quem quer que visionasse as gravações. Os principais elementos do projecto consistiam em vídeos de 50 aulas de Matemática do Japão, 80 dos EUA e 100 (na realidade foram 123, se considerarmos algumas aulas anteriores e outras posteriores) da Alemanha, no 8º ano de escolaridade, principalmente sobre álgebra e geometria.

As transcrições dos vídeos (traduzidas para inglês!) forneceram oportunidades e problemas específicos, a partir do código especialmente desenvolvido como um procedimento geral antes da análise, e dos constructos da codificação, tais como a forma da interacção, o tempo para trabalho individual ou de grupo, o nível académico do conteúdo e as tarefas matemáticas ou "problemas" usados nas aulas. O processo de codificação, assim como o exame pormenorizado de conteúdos separado de recomendações pedagógicas ou metodológicas, eram abertos a várias interpretações dos psicólogos ou linguistas envolvidos, dos professores participantes, de alguns consultores internacionais ou "especialistas" (num dado país), dos matemáticos e educadores matemáticos e por aí adiante. O facto espantoso é que o critério final para a codificação e interpretação era o *consenso* entre os diversos grupos de especialistas! Como é que os investigadores puderam lidar de forma diferente com as perspectivas, necessariamente enviesadas, destes diversos grupos? Não poderiam ter sido tomadas como matéria de discussão e debate sobre diferentes observações dos episódios? Como poderiam os resultados dos contextos, tal como o fluir pedagógico, que reflectem concepções culturais do conhecimento, de como se aprende e das práticas apropriadas de ensino,

ser utilizados para interpretar os dados?

Um problema que ainda não foi claramente inspeccionado na concepção do estudo foi a questão da representatividade das aulas gravadas em vídeo — a amostragem e a sua relação com circunstâncias importantes e fundamentais da escolarização e com oportunidades de sucesso dentro de um país. A amostra das aulas japonesas foi muito elogiada pelo director do projecto, que afirmou ser um exemplo da aula japonesa típica de resolução de problemas. Será que todos os observadores da educação japonesa (ver George, 1995; Goya, 1993; Ito, 1997) que apresentam descrições diferentes e episódios típicos contraditórios com esta visão da realidade escolar japonesa são mentirosos? Foram respeitadas as diferenças dos contextos sociais, como a aprendizagem extra-escolar e as explicações, entre o Japão e os outros países? Tendo visitado escolas do ensino básico, sem encontrar aulas do tipo das que se mostram nos vídeos, a primeira autora deste artigo duvida que as 50 aulas sejam típicas.

É sabido que os investigadores japoneses desenvolveram protótipos de aulas em vídeo, para a formação de professores, que têm por finalidade a exemplificação de alguns métodos didácticos de ensino altamente valorizados, no contexto da aula. As aulas são semelhantes às do projecto "Secondary Classroom Practice" da Universidade Aberta do Reino Unido, no qual se mostram episódios de abordagens investigativas. Não se pretende que as aulas sejam típicas ou especialmente brilhantes; são apenas exemplos de como conduzir aulas com foco no trabalho de grupo, realizado de modo independente. Muitas das aulas japonesas parecem seguir uma planificação semelhante, embora esse foco não seja tão evidente. Em seu lugar, a ênfase é mais deslocada para um desenvolvimento claramente planeado de certos conceitos ou intuições matemáticas, independentemente dos interesses, sucesso ou desempenho dos alunos, durante o ensino. Não existe a

A Casio possui a linha mais completa e acessível do mercado perfeitamente adaptada ao

procura de motivação ou de feedback da parte do professor, a nível individual, e as várias formas de chegar a uma solução apresentadas no quadro são atribuídas apenas à matemática, em lugar de o serem como possibilidades de diferentes abordagens individuais e formas de as desenvolver. Muitos dos vídeos do estudo do TIMSS podem ser como estas aulas, para serem mostrados em cursos de formação contínua ou inicial de professores, cuidadosamente planificados, preparados e conduzidos como demonstrações úteis da prática de ensino pretendida pelas entidades oficiais. As gravações podem ser úteis para observadores e alunos de cursos de formação de professores, sem representarem necessariamente a realidade da prática.

Como é que os estudos comparativos respeitam os diferentes papéis e tipos de autonomia que os professores dos países participantes têm na tomada de decisões e concepção a nível curricular? Se o professor gozar de grande autonomia, como sucede na Alemanha, as suas decisões curriculares podem diferir de uma turma, ou de grupo de alunos, para outra, podendo assim existir uma maior preocupação com a motivação e envolvimento, por um lado, e com a responsabilização ou a compreensão dos resultados da aprendizagem, por outro lado. E se o professor for o único agente de avaliação no sistema escolar, o conflito típico entre ensino e avaliação torna-se uma característica problemática que tem de ser negociada. No entanto, o ensino e a avaliação estão ligados mais estreitamente do que qualquer teste alguma vez poderá fazer.

As decisões curriculares não são as únicas características do processo de ensino que diferem substancialmente de país para país. As oportunidades de aprendizagem estão também relacionadas com a capacidade do professor em motivar os alunos a seguir as suas escolhas curriculares e em aceitar ou negociar o conteúdo de aprendizagem e o método que será seguido. Portanto, as decisões curriculares dizem respeito não só às

tarefas a conceber e realizar, mas também à discussão, negociação e perspectivas alargadas sobre o "porquê" e o "para quê" do ensino, além do "como". A capacidade e propensão do professor para conduzir estas negociações e propiciar estas perspectivas podem ser muito diferentes de um país para outro. Outro aspecto que ainda não foi tratado nem considerado nos estudos é o das diferenças nacionais entre alunos, e professores, no que diz respeito às suas crenças sobre os testes, às suas atitudes positivas ou negativas face aos testes, uma vez que pode não ser normal que lhes peçam para fazer testes. Mesmo quando o teste é a forma mais comum de avaliação nesse país, a boa vontade dos alunos em se sujeitarem a eles é diferente nos vários países. Por exemplo, os alunos alemães não estão habituados a fazer testes, nem levam a sério um teste se não estiver directamente relacionado com a classificação atribuída pelo professor.

Como é que se pode comparar a visão microscópica e aprofundada fornecida pelos vídeos das aulas com uma impressão global da realidade? Como lidar com o problema dos vários focos da investigação, à semelhança do que acontece com os estudos posteriores, tenderem a tornar-se auto-contidos, independentes e "autónomos"? Poderá um estudo deste tipo continuar sob o controlo da comunidade de investigação ou será fácil, e inevitável, que se torne uma preocupação autónoma de políticos? Será realmente possível obter uma selecção representativa da realidade da aula? Será possível chegar a uma compreensão comum do que constitui um sentido cultural dos sistemas educativos e da aula de Matemática?

Colocando o desempenho em contexto

Os resultados dos alunos em estudos comparativos são tratados como numa olimpíada académica. Apesar de serem pequenas as variações de desempenho representadas na ordenação, não é nada claro o signifi-

cado real das médias divulgadas. Até agora, não foram fornecidas mais do que interpretações superficiais, do tipo "países fracos" ou alunos sem sucesso, mesmo quando estas expressões estão ligadas a números do meio da escala. Será que as médias podem realmente proporcionar uma interpretação sofisticada? A partir de arranjos numéricos, alguns investigadores mostraram que é possível subir ou descer um certo número de lugares, efectuando pequenas alterações na ênfase de conteúdo ou no sistema de ponderação. Outros afirmam que comparações de perfis de respostas mostram mais desigualdades entre países e diferentes "oportunidades de aprendizagem" para os alunos, do que uma média pode indicar. Argumentam que a representatividade das médias nacionais esconde variações subjacentes entre regiões e entre escolas. Nem sempre há garantia de que a amostragem dos alunos tenha sido correcta, ou que as médias não tenham sofrido um abaixamento devido às condições em que o teste foi aplicado em alguns países em vias de desenvolvimento. Além disso, os relatórios de desempenho nos vários anos de escolaridade insistem no "crescimento" do desempenho matemático, apesar de os testes e os alunos serem diferentes nos vários anos considerados. E as diferenças da classificação dos países nos vários níveis de escolaridade poderem dever-se mais a alterações no conjunto de países que participaram no estudo referente a um dado ano de escolaridade, do que ao mérito ou demérito do currículo, do ensino ou dos alunos.

No que respeita à informação cultural e contextual sobre os vários países, sabemos, a partir dos estudos complementares, que os alunos no Japão, Coreia e Singapura têm muito tempo extra para estudarem Matemática, em explicações privadas ou em escolas especiais, fora das escolas oficiais. Em relação aos alunos japoneses, a Matemática e as Ciências são as disciplinas de que gostam



menos, mas, se os compararmos com os alemães, passam mais um terço do horário semanal a rever e preparar as suas aulas de Matemática (Shümer, 1997) e muitos deles frequentam com regularidade aulas depois do almoço no *juku*, onde são preparados, principalmente, para os exames. Em 1996, o governo japonês (Monbusho) publicou um relatório sobre os planos de política educativa. Este relatório fornece dados sobre o tempo de aprendizagem extra que os alunos ocupam, mostrando que ele tem estado a aumentar muito acentuadamente, nos últimos dez anos. Nos primeiros seis anos de escolaridade, 77% dos alunos, em média, têm explicações ou aulas privadas e 24% frequentam as escolas privadas de remediação do tipo *juku*. Nos anos correspondentes ao nosso 3º ciclo, as médias são de 28% em aulas e explicações privadas e 60% no *juku*, havendo alguns alunos que frequentam ambos. No mesmo relatório, são analisadas as opiniões dos pais e professores das escolas oficiais sobre a frequência do *juku*: 60% acham que a frequência do *juku* é um exagero; 58% acreditam que a formação do carácter das crianças é influenciada negativamente pela excessiva competição em exames de admissão; 48% sentem que a saúde e o bem-estar das crianças são afectados negativamente pelo excesso de tempo ocupado no *juku*; 34% acham que a escola é desvalorizada, devido à preocupação com os estudos no *juku*; 11% assinalam que as crianças ficam perturbadas e confusas pelas diferenças existentes entre os tipos de ensino da escola e do *juku*; e 15% acham que as crianças revelam problemas comportamentais, despoletados pela frequência no *juku*, e desenvolvem outros problemas relacionados com a excessiva competição que aí existe. Mais de metade dos pais queixam-se que, por causa da excessiva frequência do *juku*, as crianças têm uma experiência inadequada de vida e de actividades extra-escolares, incluindo lúdicas, comunitárias e interações familiares. Quase 45% pensam que a trajetória profissi-

onal a efectuar pela criança é determinada, em primeiro lugar, de acordo com "classificações normativas", independentemente dos seus desejos ou aptidões, normalmente desprezados. Entre os pais que foram inquiridos, 38% afirmam que o sistema de aulas privadas e do *juku* lhes exige demasiado, sob o ponto de vista financeiro, e apenas 11% não via nenhum problema específico neste sistema (Monbusho, 1996, pp. 33-34). A competição entre este segundo sistema de aulas privadas e o *juku* é considerada pelos professores e pais como um obstáculo importante a qualquer pretensão de reforma (Ito, 1997).

Como é que se pode desenvolver um conceito claro do que é uma prática apropriada de ensino? O ensino e a aprendizagem ineficazes são devidos apenas a problemas metodológicos? A partir da análise dos vídeos, é claro que apenas conta o que pode ser observado e posteriormente interpretado de modo razoável; os métodos analíticos são claramente diferenciáveis, logo preferidos. No entanto, onde podemos alcançar a visão metacognitiva de que as atitudes e crenças, tanto dos alunos como dos professores, assim como vários factores sociais, determinam a falta de eficácia? É fácil descobrir que os professores norte-americanos passam mais tempo na escola, do que os professores japoneses e alemães, e que passam mais tempo nas aulas do que a prepará-las (Schmidt et al., 1996, p. 2). Será então correcto concluir que o método americano é "dizer e mostrar" (McKnight et al., 1987, p.81), ao passo que os japoneses usam o método socrático e actividades práticas apropriadas (Stigler e Stevenson, 1991, p.91)? Como é que os países garantem, do seu ponto de vista, que as prescrições exemplarmente "boas" ou curriculares e pedagogicamente desejáveis são transformadas em realidade?

Conclusão

Dentro de cada sistema educativo, o currículo da Matemática é um organismo que funciona, tanto a nível

retórico, como a nível real. O funcionamento deste todo orgânico não foi examinado, segundo os seus próprios termos, para cada sistema, em nenhum dos grandes estudos recentes. Em vez disso, foi imposto um pseudo consenso (em primeiro lugar pelo mundo de expressão anglo-saxónica) entre todos os sistemas, de modo a tratar o "currículo" como uma constante, e não como uma variável, e a tornar possível examinar as operações efectuadas com outras variáveis. Este modo de tratar o currículo tem tido os seus custos. Chegou a altura de começar a examinar de um modo mais aprofundado e mais sério os currículos de Matemática dos diferentes sistemas educativos, de forma a que as avaliações do que os alunos, no quadro desses sistemas, sabem e são capazes de fazer matematicamente possa ser objecto de interpretações mais válidas e produtivas.

O TIMSS começou com grandes esperanças de poder fornecer um contexto enriquecido, no qual o desempenho dos alunos seria apreendido em pormenor e seria ligado a outras informações sobre a prática na sala de aula e os currículos nacionais. Estas esperanças foram, em parte, atingidas, mas os presumíveis melhoramentos feitos no TIMSS, em relação a estudos anteriores, foram maculados pelo domínio dos EUA no financiamento da maior parte da investigação e na direcção da recolha e análise de dados. Como consequência, trata-se de um estudo impregnado pelas tradições de investigação de um país, que, com demasiada frequência, diz pouco ou nada aos educadores matemáticos de outros países, nomeadamente no que diz respeito ao modo como a educação pode ser melhorada nesses países. Além disso, como o mais abrangente e mais amplamente publicitado estudo comparativo já realizado, o TIMSS ameaça inquinar durante algum tempo as águas da política educativa, enquanto os investigadores e os políticos lutam para tirar dividendos do que o TIMSS alegadamente afirma, sobre o ensino e a aprendizagem da Matemática nos seus países. São



feitas inferências abusivas sobre a ligação entre perspectivas seleccionadas de ensino e indicadores falíveis de desempenho, de modo que educadores insuspeitos, assim como o público em geral, venham a pensar que o ensino descrito produziu o desempenho medido.

Os estudos comparativos internacionais são anunciados aos quatro ventos em revistas de educação e na imprensa, como triunfos da racionalidade. São citados como se os resultados que fornecem não fossem questionáveis. Críticas sérias e expressões de dúvida são rejeitadas, como censuras de provocadores ignorantes ou mal informados. Entretanto, a irracionalidade inerente a estes estudos é varrida para baixo da tapete. Os investigadores que conduziram os estudos investiram demasiado nos resultados, para poderem comprometer-se numa reflexão séria sobre os fundamentos do seu trabalho. E como é que pode haver irracionalidade, quando tantos educadores e cientistas reputados trabalharam tanto para produzir resultados organizados e científicos? O facto de poderem coexistir, nestes estudos, a racionalidade e a irracionalidade, pondo em causa as análises bem fundamentadas e os resultados cuidadosamente apresentados, é algo que não ocorre a muita gente. No entanto, um breve exame dos documentos produzidos nestes estudos mostra que por cada ponto forte há sempre graves lacunas. Os estudos baseiam-se nos fundamentos menos firmes: assumem que a armação científica pode cobrir todas as fraquezas na concepção, os dados incongruentes e os erros de interpretação. Eles não só comparam o incomparável; eles racionalizam o irracional.

Nota

¹ Utilizamos sistemas em lugar de países ou nações porque os países podem estar representados, em investigações comparativas, por estados, províncias, grupos linguísticos ou mesmo cidades (ver Westbury, 1989, p. 19).

Referências

- Baker, D. P. (1997). Good news, bad news, and international comparisons: Comment on Bracey. *Educational Researcher*, 26 (4), 16-17.
- Bracey, G. W. (1996). International comparisons and the condition of American education. *Educational Researcher*, 25 (1), 5-11.
- Bracey, G. W. (1997a). On comparing the incomparable: A response to Baker and Stedman. *Educational Researcher*, 26 (4), 19-26.
- Bracey, G. W. (1997b). Many visions, many aims, one test. *Phi Delta Kappan*, 78 (5), 411-412.
- Freudenthal, H. (1975). Pupils' achievements internationally compared — The IEA. *Educational Studies in Mathematics*, 6, 127-186.
- George, P. (1995). *The Japanese secondary school: A closer look*. Reston, VA: National Association of Secondary School Principals, & Columbus, OH: National Middle Schools Association.
- Goya, S. (1993). The secret of Japanese education. *Phi Delta Kappan*, 75 (2), 126-129.
- Hirsch, E. D., Jr. (1996). *The schools we need and why we don't have them*. New York: Doubleday.
- Husén, T. (1967). *International Study of Achievement in Mathematics: A comparison of twelve countries* (Vols. 1 & 2). New York: Wiley.
- Husén, T. (1983). Are standards in U.S. schools really lagging behind those in other countries? *Phi Delta Kappan*, 64, 455-461.
- Ito, Toshiko (1997). Zwischen "Fassade" und "wirklicher Absicht." Eine Betrachtung über die dritte Erziehungsreform in Japan (Between "façade" and "actual intention": A reflection on the third reform of education in Japan). *Zeitschrift für Pädagogik*, 3, 449-466.
- McKnight, C., Crosswhite, F. J., Dossey, J. A., Kifer, E., Swafford, J. O., Travers, K. J., & Cooney, T. J. (1987). *The Underachieving Curriculum*. Champaign, IL: Stipes Publishing.
- Monbusho (Ministry of Education, Science, Sports and Culture). (1996). *Japanese government policies in education, science, sports and culture*. Tokyo: Author.
- Lapointe, A., Mead, N. A., & Philips, G. W. (1989). *A world of differences: An international assessment of mathematics and science* (Report No. 19-CAEP-01). Princeton, NJ: Educational Testing Service.
- Schmidt, W. H., Jorde, D., Cogan, L. S., Barrier, E., Gonzalo, I., Musor, U., Shimizu, Y., Sawada, T., Valverde, G., McKnight, C., Prawat, R., Wiley, D. E., Raizen, S., Britton, E. D., & Wolfe, R. G. (1996). *Characterizing pedagogical flow*. Dordrecht: Kluwer Academic Publishers.
- Schmidt, W. H., McKnight, C. C., & Raizen, S. A. (1997). *A splintered vision: An investigation of U.S. science and mathematics education*. Dordrecht: Kluwer Academic Publishers.
- Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. I., & Wiley, D. E. (1997). *Many visions, many aims, Vol. 1: A cross-national investigation of curricular intentions in school mathematics*. Dordrecht: Kluwer Academic Publishers.
- Schümer, G. (in press). The TIMSS Videotape Classroom Study. In *Proceedings of the conference on mathematics and elementary science education: German, Japanese and U.S. perspectives* (Dec. 3-5, 1997). Berlin: Japanese-German Center of Berlin.
- Stedman, L. C. (1997). International achievement differences: An assessment of a new perspective. *Educational Researcher*, 26 (4), 4-15.
- Stigler, J. W., & Stevenson, H. W. (1991). How Asian teachers polish each lesson to perfection. *American Educator*, 15 (1), 12-20, 43-47.
- Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (eds.) (1997). *The TIMSS Videotape Classroom Study: Methods and Preliminary Findings* (Draft version), Los Angeles, CA: University of California at Los Angeles, Psychology Department.
- Survey of Science and Mathematics Opportunities (1993). *TIMSS: Concepts, Measurements and Analyses* (Research Report Series No. 56). East Lansing, MI: Michigan State University.
- Travers, K. J., & Westbury, I. (1989). *The IEA Study of Mathematics, I: Analysis of Mathematics Curricula*. Oxford: Pergamon.
- US Department of Education, National Center for Education Statistics (1997). "TIMSS video classroom study" (on-line). Available: <http://nces.ed.gov/timss/viedo/index.html>.
- Westbury, I. (1989) "The problems of comparing curricula across educational systems", in Purves, A. C. (ed.) *International Comparisons and educational Reform*, Alexandria, VA: Association for Supervision and Curriculum Development, 17-34.

Christine Keitel, Universidade Livre de Berlim, Alemanha
Jeremy Kilpatrick, Universidade da Geórgia, EUA

Tradução de Fernando Nunes, Esc. Básica 2,3 Marquesa de Alorna
Revisão de Paulo Abrantes, Universidade de Lisboa