

Fazendo uso dos resultados de testes educacionais

síntese de evidências sobre competências e práticas organizacionais

Outubro, 2022

EVE
Evidência Express

ENAP

Expediente

Presidente

Diogo Costa

Diretora-Executiva

Rebeca Loureiro de Brito

Diretora de Altos Estudos

Flávia de Holanda Schmidt

Diretor de Educação Executiva

Rodrigo Torres

Diretor de Desenvolvimento Profissional

Paulo Marques

Diretora de Inovação

Camila de Castro Barbosa Medeiros

Diretora de Gestão Interna

Alana Regina Biagi Silva Lisboa

Coordenador Geral de Ciência de Dados

Pedro Masson Sesconetto Souza

Capa e Diagramação

Samyra Lima

Equipe Evidência Express

Imagens

Unsplash

Autoria

Géssica C. Souza

Doutora em Economia (UFMG).

Breno Salomon Reis

Mestre em Políticas Públicas (Insper).

O Evidência Express (EvEx) é uma iniciativa da Diretoria de Altos Estudos da Escola Nacional de Administração Pública (Enap) em parceria com a Universidade Federal de Brasília (UnB). A missão do EvEx é melhorar a tomada de decisão do setor público. Para isso a equipe sintetiza, produz e dissemina evidências que possam servir de base para o desenho, monitoramento e avaliação de políticas públicas.

Avaliações completas de políticas públicas são intensivas em tempo e custos. A fim de agilizar esses processos, o EvEx produz relatórios ágeis de evidências para a consolidação do conhecimento disponível e introdução de novos pontos de vista.

Os resultados dos produtos EvEx apoiam tomadores de decisão do setor público federal, subsidiando avaliações Ex Ante, Ex Post ou Análises de Impacto Regulatório. Beneficiam também os gestores públicos subnacionais, pesquisadores, docentes, servidores e demais interessados na sociedade civil.

Os produtos EvEx analisam evidências qualitativas e quantitativas, podendo ser demandados de forma avulsa ou em pacotes, sobre:

- Evolução do problema no Brasil e no mundo;
- Público-alvo de uma política;
- Causas e consequências do problema ou política;
- Soluções existentes para o problema;
- Impactos de intervenções ou políticas públicas.

Para mais informações, consulte nossa página (www.enap.gov.br/pt/servicos/avaliacao-e-organizacao-de-evidencias) ou entre em contato: evidencia.express@enap.gov.br.



Sumário Executivo

- Essa síntese de evidências buscou identificar os recursos e competências (*capabilities*), e práticas organizacionais (*management practices*), necessários para fazer melhor uso dos resultados de avaliações padronizadas educacionais.
- Os sistemas de Monitoramento e Avaliação (M&A) são geralmente construídos para fornecer informações capazes de relatar as realizações de um determinado programa ou política.
- São diversos os obstáculos e fatores que impedem ou afetam a utilização dos resultados das avaliações como: a utilização das avaliações para prestações de contas; a falta de tempo da equipe e a alta carga administrativa; dificuldade de analisar os resultados; falta de habilidade para interpretação e análise de dados (falta de alfabetização de dados); mudanças organizacionais, dentre outros.
- Os fatores que favorecem o uso dos dados podem ser classificados em três tipos: características dos dados, qualidades do usuário e condições organizacionais.
- Destacam-se como aspectos organizacionais que facilitam o uso das avaliações, a existência de uma cultura de uso de dados, a instalação dessa prática na rotina da organização, a existência de espaços e o trabalho colaborativo em torno dos dados (CONTRERAS; TORRES, 2022).
- As habilidades específicas para o domínio de dados são: identificar problemas e formular perguntas a partir do uso dos dados, transformar dados em informações, transformar informações em decisões e avaliar resultados (MANDINACH; GUMMER, 2013).
- Para que o processo de utilização dos dados se torne um instrumento eficiente é necessário que não se comece pelos dados, mas por objetivos claros e mensuráveis (MANDINACH; SCHILDKAMP, 2021).
- Sobre as características dos dados, faz-se necessário garantir a confiabilidade, validade, objetividade e o detalhamento da informação.
- Ainda é menos compreendido como promover facilitadores e remover barreiras para desbloquear, de forma completa, o uso de dados. Assim, são necessárias mais pesquisas que busquem identificar os caminhos necessários para promover a real alfabetização de dados, não somente no âmbito da manipulação, mas também do entendimento e interpretação dos mesmos.

Sumário

1	Introdução	5
2	Metodologia	7
3	O sistema de Monitoramento e Avaliação (M&A)	8
4	O uso das avaliações	12
4.1	Conceito e mapeamento do uso das avaliações	12
4.2	Fatores que afetam o uso das avaliações	14
4.3	Recomendações para o melhor uso das avaliações	16
5	Considerações finais	19
	Referências Bibliográficas	19

1. Introdução

O Monitoramento e Avaliação (M&A) é uma função de gestão contínua para avaliar se houve progresso na obtenção dos resultados esperados, identificar gargalos na implementação e destacar se existem efeitos não intencionais (positivos ou negativos) de um plano de investimento, programa ou projeto e suas atividades (FAO, 2022).

No âmbito educacional, todas as formas de avaliação consistem em coletar informações sistemáticas e fazer algum tipo de julgamento com base nessas informações. Outra expectativa é que essa “informação valiosa” seja usada para decisões sobre o funcionamento cotidiano dos sistemas educacionais ou para decisões mais envolventes sobre a revisão e mudança do sistema. O M&A educacional faz uso de diferentes fontes de dados que se distinguem entre dados baseados na medição de desempenho educacional, dados que estão disponíveis a partir de registros administrativos (incluindo estatísticas educacionais) e dados que se tornam disponíveis a partir de análises de especialistas e métodos de pesquisa educacional (SCHEERENS et al., 2003).

Os modelos de avaliação educacional, que geram os dados de medição de desempenho educacional, assumem como princípio que a missão da escola é a aprendizagem dos alunos. Sendo assim, o foco está no desempenho que os alunos alcançam nos testes cognitivos e no ambiente em que esse desempenho é gerado. Apesar de os resultados dos testes estarem sendo utilizados como um indicativo da aprendizagem, eles sozinhos não são capazes de configurar a qualidade da educação. Além disso, de acordo com Contreras e Torres (2022), tem havido uma subutilização destas informações ou uma discrepância entre as finalidades para as quais foram concebidas e a forma como são utilizadas. Isso vem aumentando o debate sobre o processo e os modelos de avaliação e monitoramento implementados no Brasil, assim como o uso desses resultados pelos diversos atores e esferas educacionais. São exemplos de monitoramento da educação básica os testes padronizados do Sistema de Avaliação da Educação Básica (Saeb) e o Índice de Desenvolvimento da Educação Básica (Ideb) (INEP, 2019).

O Saeb é realizado periodicamente pelo Inep desde os anos 1990. Dentre os vários objetivos do sistema de avaliação, no âmbito da Educação Básica, destaca-se: (i) avaliar a qualidade, a equidade e a eficiência da educação praticada no país em seus diversos níveis governamentais; (ii) subsidiar a elaboração, o monitoramento e o aprimoramento de políticas públicas baseadas em evidências, com vistas ao desenvolvimento social e econômico do Brasil; e (iii) desenvolver competência técnica e científica na área de avaliação educacional, ativando o intercâmbio entre instituições educacionais de ensino e pesquisa (INEP, 2019). No entanto, alguns estudos como o de Neto (2006) indicam que os dados apurados pelo Saeb, não são suficientes, per se, para orientar a tomada de decisões, por diversos motivos que serão apresentados no decorrer desse estudo.

Dessa forma, se torna fundamental a discussão sobre como melhorar os modelos de moni-

toramento e avaliação e como fazer o melhor uso dos resultados dessas avaliações. Os principais motivos para criar ou melhorar as provisões para monitoramento e avaliação na educação são três grandes preocupações: regular formalmente os níveis desejados de qualidade dos resultados e provisões educacionais; responsabilizar os provedores de serviços educacionais e apoiar a melhoria contínua na educação (SCHEERENS et al., 2003).

O uso de dados de avaliações implica uma série de processos, condições e contextos. Envolve processos interpretativos, pois o uso de dados requer que o usuário interprete-os e construa interpretações para os próximos passos. Pré-supõe condições sociais e organizacionais, uma vez que os resultados se desdobram no contexto de um sistema organizacional multinível. E, como os dados podem ser uma fonte de poder, particularmente no ambiente atual de prestação de contas, sua utilização também envolve relações de poder (COBURN; TURNER, 2011).

Portanto, diante da importância do monitoramento e avaliação na educação e diante da relevância e desafios em relação ao uso dos resultados dessas avaliações, o principal objetivo dessa pesquisa, em formato de síntese de evidência, é identificar os recursos e competências (*capabilities*), e práticas organizacionais (*management practices*), que os órgãos públicos e todos os atores de interesse, precisam adquirir e/ou desenvolver para fazer o melhor uso dos resultados de avaliações padronizadas.

A hipótese é que os órgãos públicos, ao desenvolverem certas *capabilities*, ou implementarem certas *management practices*, conseguiriam fazer um melhor uso dos dados resultantes das avaliações para planejar (desenhar/design) e executar políticas educacionais. Assim, devido a essas *capabilities* e *practices*, a educação alcançaria um nível de desempenho superior.

Esse artigo está dividido em cinco seções considerando a introdução. Na próxima seção detalha-se a metodologia adotada. Na seção subsequente o sistema de Monitoramento e Avaliação (M&A) é evidenciado do ponto de vista das práticas que deveriam ser adotadas para garantir o bom uso dos resultados das avaliações. Na quarta seção, traça-se o panorama do uso das avaliações do ponto de vista teórico e prático, evidenciando os fatores que afetam o uso dos resultados e as recomendações. Por fim, constrói-se as considerações finais.

Salientamos que esta pesquisa foi elaborada de maneira independente pelo Evidência Express (EvEx) nos meses de setembro e outubro de 2022, e não representa os resultados e recomendações de políticas públicas elaborados pelo Conselho de Monitoramento e Avaliação de Políticas. Os produtos do EvEx são elaborados como respostas ágeis, oferecendo apenas resultados preliminares e não devem ser tratados como evidências definitivas sobre o fenômeno estudado.

2. Metodologia

Com o objetivo de identificar os recursos e competências (*capabilities*), e práticas organizacionais (*management practices*), necessários para fazer melhor uso dos resultados de avaliações padronizadas, foram realizadas buscas de artigos, livros, capítulos de livro, trabalhos apresentados em conferências e revisões de literatura nos portais indexadores Scopus, a partir da combinação de palavras chaves apresentadas na Tabela 2.1 e a partir de busca manual via Google Acadêmico.

Os principais achados, que contribuíram efetivamente com o objetivo da pesquisa, foram encontrados por meio das referências e citações do Livro *Educational evaluation, assessment, and monitoring: A systemic approach* (SCHEERENS et al., 2003). A partir desse livro identificou-se os principais teóricos do "evaluation use" e da "avaliação focada na utilização", são eles: Alkin, Daillak e White (1979), Patton (1978) e Huberman (1987). A partir das citações cruzadas com esses autores *mainstream* e de consultas direcionadas à revista *Studies in Educational Evaluation*, chegou-se no total de 15 artigos científicos capazes de contribuir diretamente com o problema de pesquisa. Vale destacar que para chegar nesses artigos finais, foi necessária a realização de um filtro a partir de uma vasta leitura de títulos e resumos.

Tabela 2.1: Palavras-chaves utilizadas para executar as buscas

Tema	Termos
Gestão Pública	<i>public policy design, capability, competencies, monitoring and evaluation systems, evaluation capacity, evidence based public policy, research utilization, goal setting, effective management, gestão por resultados, data-based, data-based decision making (DBDM)</i>
Educação	<i>basic education, economia da educação, standardized test, avaliação de larga escala, Large-scale Assessments in Education, evaluation of basic education, use of administrative data, student monitoring systems</i>

Fonte: Elaborada pelos autores.

3. O sistema de Monitoramento e Avaliação (M&A)

Para conseguir extrair o melhor uso de qualquer resultado de avaliação se faz necessário entender todo o sistema de monitoramento que envolve, ou deveria envolver, a construção dessas avaliações. Os sistemas de Monitoramento e Avaliação (M&A) são geralmente construídos para fornecer informações capazes de relatar as realizações de um determinado programa ou política. Um dos objetivos é cumprir as responsabilidades de prestação de contas, o que fez com que o M&A fosse amplamente associado a uma função de controle e prestação de contas. Porém, cada vez mais, há o reconhecimento de que os sistemas de M&A também podem contribuir para a gestão estratégica, para o aprendizado de lições, e para alimentar experiências em processos políticos (HILHORST; GUIJT, 2006).

O desenvolvimento de M&A deve se concentrar em incutir um processo que produza dados de resultados regulares que possam ser usados pelos órgãos governamentais e por funcionários de nível superior para prestação de contas e gestão, ajudando assim no melhoramento das prestações de serviços aos cidadãos e públicos ao qual foi direcionado a política (LOPEZ-ACEVEDO; KRAUSE, 2012). São necessários vários passos básicos para conceber e implementar o processo de M&A que vão desde a delimitação do comitê de funcionários-chave até a delimitação clara dos objetivos do M&A. Todos os passos são muito bem detalhados em Lopez-Acevedo e Krause (2012)¹, mas vale destacar dois passos importantes para o bom uso dos resultados de monitoramento:

- É necessário enfatizar a importância das informações de resultados aos órgãos governamentais e atores interessados, principalmente quando as informações de M&A serão utilizadas na tomada de decisão. Eles também precisam usar as informações coletadas sobre os resultados para identificar as melhorias necessárias no programa e as necessidades de treinamento ou assistência técnica, e chamar a atenção para programas bem-sucedidos. O uso de dados de M&A em todos os níveis governamentais, como planejamento, orçamento e revisão do programa, destaca a importância do sistema de M&A.
- Desde o início, enfatize que alcançar bons resultados será recompensado. Enfatize que um dos principais objetivos das informações de M&A é ajudar os funcionários a identificar o que funciona e o que não funciona. Use as informações de M&A para estabelecer as bases para incentivos positivos para melhorar a eficácia dos serviços públicos.

De acordo com Lopez-Acevedo e Krause (2012), várias experiências mostraram que gastar enormes quantias de dinheiro para desenvolver sistemas tecnológicos sofisticados de M&A não garante necessariamente que a base de indicadores atende aos padrões técnicos de qualidade e garanta a utilização dos dados de desempenho. Uma abordagem personalizada e de bom senso pode ser muito

¹Páginas 64 a 70.

mais importante. Essa abordagem consideraria os arranjos institucionais necessários, responderia às necessidades específicas de medição de desempenho dos usuários-chave e garantiria que as fontes de informação estejam disponíveis, sejam confiáveis e tenham uma medição adequada.

Como apontado por Lopez-Acevedo e Krause (2012), o modelo de gestão da informação por trás da base de indicadores de desempenho é a verdadeira espinha dorsal de um sistema de M&A. É necessário garantir que as fontes de informação e os fluxos de dados que emergem dos indicadores possam ter credibilidade suficiente para garantir a utilização por diferentes usuários. O acesso público à base de indicadores e fontes de informação, além de auditorias externas periódicas, têm se mostrado eficaz no controle da qualidade dos indicadores e sua base de informações (LOPEZ-ACEVEDO; KRAUSE, 2012).

Outro fator de sucesso está na diferenciação entre os sistemas de indicadores de desempenho e regimes de estabelecimento de metas. Tanto os indicadores quanto as metas são elementos-chave de um sistema de M&A ou de um modelo de gestão de desempenho. No entanto, a definição de indicadores de desempenho não deve ser interpretada como tendo metas associadas a cada um deles. Para serem significativas, as metas dependem de uma medida subjacente, mas os indicadores de desempenho não precisam necessariamente ter metas associadas para serem úteis. O processo de definição de metas pelas quais um governo será responsável, e contra as quais os controles gerenciais serão exercidos, requer um certo nível de institucionalização de M&A e boas práticas de gestão de desempenho (LOPEZ-ACEVEDO; KRAUSE, 2012).

Uma das lições aprendidas pelo Canadá sobre a construção e a capacidade de utilização do M&A é a importância de considerar não apenas o treinamento técnico, mas também o treinamento e orientação de M&A para funcionários não técnicos (ou seja, os usuários de informações de M&A). Além disso, a capacitação precisa abordar uma área frequentemente ignorada – desenvolvimento de dados e criação de bancos de dados confiáveis. Os órgãos nacionais de estatísticas devem desempenhar um papel central no desenvolvimento, armazenamento, supervisão e controle de qualidade associados à coleta e levantamento de dados (LAHEY, 2010).

De acordo com a FAO (2022), os componentes de M&A e ferramentas associadas podem ser implementados mecanicamente sem oferecer muita utilidade para a tomada de decisões. No entanto, existem algumas sugestões práticas para fortalecer os sistemas de M&A:

- Garantir que a administração e as partes interessadas se envolvam e assumam a propriedade do trabalho de M&A. A gestão do projeto e as responsabilidades da equipe devem ser internalizadas, para evitar a percepção de que M&A é uma tarefa de relatório independente da política/programa avaliado.
- Fortalecer a capacidade de planejamento e M&A da política/programa e criar uma cultura de aprendizagem. Os programas geralmente são experimentais e precisam maximizar o aprendizado de lições para uma possível ampliação e consolidação.
- Use triangulação e métodos mistos e participativos: o uso complementar de ferramentas de

M&A e métodos quantitativos e qualitativos é importante porque nenhuma ferramenta fornecerá todas as informações. Às vezes, visões e conhecimentos externos são essenciais para novos *insights* críticos e análises especializadas.

- Construir redes e compartilhar aprendizado entre projetos e programas. A colaboração é fundamental para disseminar e acessar rapidamente o conhecimento prático sobre M&A.

Diante dessas sugestões práticas, vale destacar o conceito e contribuições do Monitoramento e Avaliação Participativa (*Participatory Monitoring and Evaluation* - PM&E). No PM&E, as partes interessadas estão envolvidas e influenciam a tomada de decisão, alocação de recursos, implementação e controle de iniciativas de desenvolvimento dos programas e políticas de avaliação. As partes interessadas podem ser a nível de governança subnacional mais baixo, onde o governo local eleito e os provedores de serviços da “linha de frente” se envolvem com os cidadãos, organizações, programas e políticas (HILHORST; GUIJT, 2006).

O PM&E pode agregar valor de duas maneiras: garantindo que informações e experiências relevantes sejam coletadas daqueles que são imediatamente usuários do projeto, programa ou política, e aumentando a responsabilidade desses participantes que têm interesse direto no sucesso da implementação. O processo de participação aumenta ainda mais a propriedade das atividades e a probabilidade de replicação e sustentabilidade. Esforços especiais precisam ser feitos para incorporar as partes interessadas em todos os níveis para garantir que elas contribuam e se beneficiem do compartilhamento de conhecimento (HILHORST; GUIJT, 2006).

No contexto educacional, todas as formas de avaliação consistem em coletar informações sistemáticas e fazer algum tipo de julgamento com base nessas informações. Outra expectativa é que essa “informação valiosa” seja usada para decisões sobre o funcionamento cotidiano dos sistemas educacionais ou para decisões sobre a revisão e mudança do sistema educacional (SCHEERENS et al., 2003).

De um ponto de vista mais pragmático e orientado para a ação, o contexto institucional, político e organizacional em que o M&A ocorre pode ser analisado em termos de restrições. Compreender essas restrições é um primeiro passo para tentar superá-las e criar melhores pré-condições para que as avaliações se desenrolem de maneira tecnicamente adequada. A implementação de dispositivos de M&A educacional a nível nacional, particularmente no contexto dos países em desenvolvimento, deve ser vista como um processo de inovação, que muitas vezes requer desenvolvimento organizacional. Isso pode envolver o recrutamento e a coordenação do potencial de recursos humanos técnicos necessários e a facilitação do uso da informação para os tomadores de decisão e outras partes interessadas (SCHEERENS et al., 2003).

Existem diferentes tipos de M&A educacional, mas, como observado por Scheerens et al. (2003), a escolha deve levar em consideração aspectos políticos, institucionais e organizacionais do contexto local. O uso e a implementação bem-sucedidos dessas opções técnicas dependem da criação de condições de apoio nessas áreas. O sucesso da melhoria do M&A como função de apoio ao dia a

dia e ao planejamento estratégico dos sistemas educacionais, naturalmente, depende do grau em que o país já tem um histórico no emprego dessa função. Trata-se de arranjos estruturais e formais, como um sistema de exames, mas também de algo que poderia ser indicado como uma “cultura de avaliação” (SCHEERENS et al., 2003).

4. O uso das avaliações

4.1 Conceito e mapeamento do uso das avaliações

Nesta subseção serão apresentados diferentes possibilidades de uso dos resultados de avaliações presentes em sistemas de M&A para que possa se ter um vislumbre de quais capacidades órgãos públicos precisam adquirir para fazer melhor uso dos resultados das avaliações educacionais. Existe uma literatura interessada no uso dos resultados das pesquisas de avaliação pelos formuladores de políticas ou “*evaluation use*”.

A suposição ao aplicar a pesquisa de avaliação é que os resultados desempenharão um papel importante na tomada de decisão política. Alguns teóricos da avaliação proclamaram um tipo de avaliação, indicada como "avaliação focada na utilização"(ALKIN; DAILLAK; WHITE, 1979; PATTON, 1978), que se distingue por propagar um diálogo entre pesquisadores e praticantes ou formuladores de políticas, buscando o comprometimento dos tomadores de decisão com a avaliação e tentando apresentar o resultado da avaliação de uma forma “amigável” (SCHEERENS et al., 2003).

No caminho da "avaliação focada na utilização", Huberman (1987) desenvolveu uma estrutura em que a utilização dos resultados são descritas como dependentes das condições estruturais de fundo organizacional dos atores-chave e em aspectos comunicativos, tanto no fornecimento quanto no uso das pesquisas. No seu modelo organizacional são destacadas algumas características relevantes dos usuário, como experiência anterior com pesquisa e avaliação, *know-how* em avaliação e clima positivo para o uso dos resultados da pesquisa. Esse estrutura desenvolvida por Huberman é bastante rica em explicitar a complexidade dos fatores organizacionais, motivacionais e “políticos” que estão em jogo na implementação, interpretação e uso das avaliações. O conhecimento desses fatores no planejamento e na implementação de avaliações escolares pode ajudar no uso adequado dos resultados.

De acordo com Alkin e King (2017), as primeiras discussões sobre o “*evaluation use*”, ou uso da avaliação, procuravam identificar os componentes que seriam considerados na utilização das avaliações. Braskamp (1982) *apud* Alkin e King (2017), por exemplo, diz que o uso de uma avaliação ocorre quando qualquer uma das seguintes condições estiver presente: “alocação, uso direto” (instrumental), “iluminação conceitual” (conceitual), “impacto nas políticas e práticas organizacionais como resultado da ameaça de uma avaliação” (simbólica) e “contribuição para a gestão da organização” (aprendizagem organizacional ou capacitação).

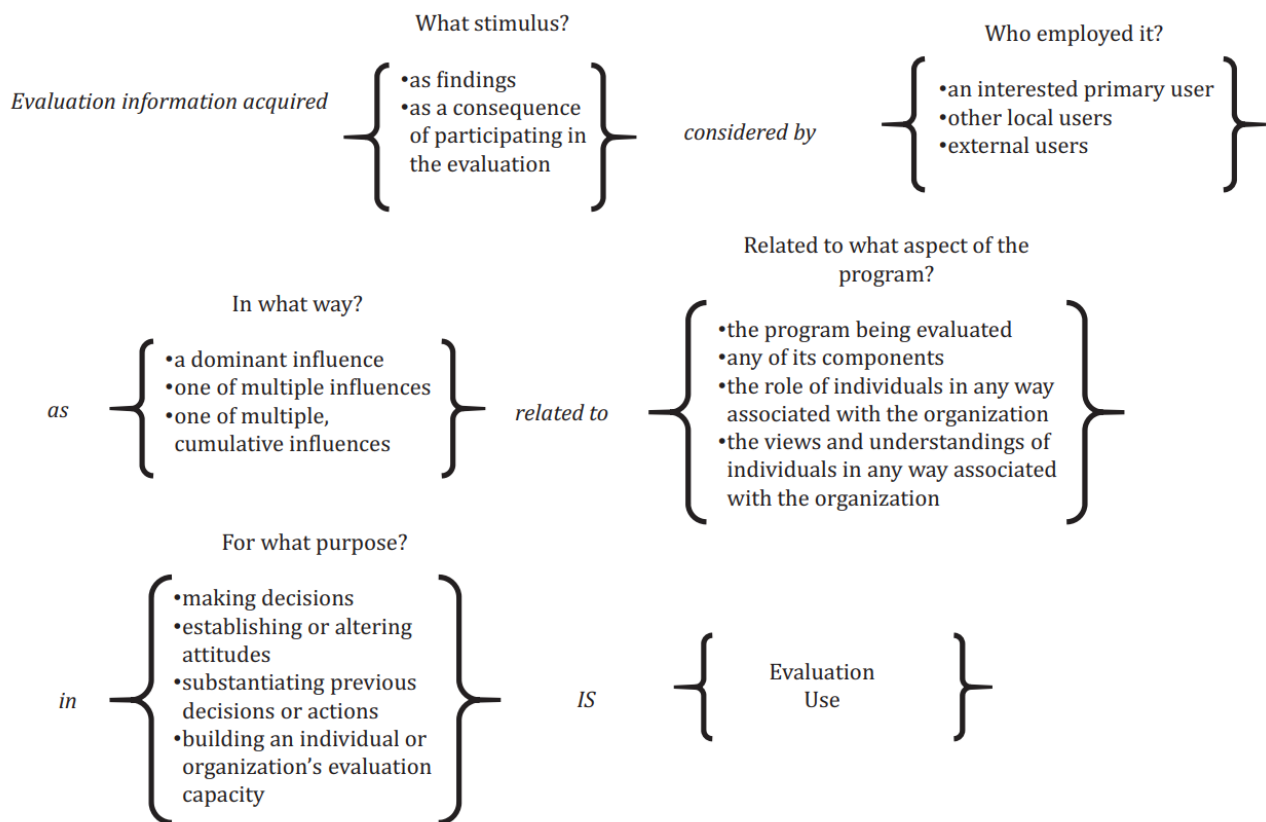
Diante dessas condições de uso, Alkin e King (2017) procurou mapear o uso das avaliações considerando também a sentença de mapeamento de Guttman¹, um formato útil como representação teórica que definem instâncias de uso. O mapeamento é dividido em cinco processos: estímulo,

¹As sentenças de mapeamento de Guttman fornecem aos pesquisadores e avaliadores um método para organizar sistematicamente uma série de características relevantes sob uma ou varias variáveis.

usuário, tipo de influência, aspectos e proposta. Os processos seguem na Figura 4.1 e serão brevemente detalhados ao longo do texto, dada a capacidade de estruturar uma sequência racional lógica para o bom uso dos resultados de avaliação.

A etapa de estímulo, como o próprio nome diz, procura estimular o usuário em potencial. Os resultados da avaliação irão estimular o uso e nesse primeiro momento a disponibilização dos resultados da avaliação deve ser feita de forma resumida, por meio de algum tipo de relatório (por exemplo, *policy brief* e exposição oral). Alternativamente, o estímulo pode consistir nas informações ou entendimentos derivados como consequência da participação na avaliação (ALKIN; KING, 2017).

Figura 4.1: Mapeamento do uso das avaliações por meio de uma escala de Guttman adaptada



Fonte: Alkin e King (2017)

A etapa do usuário indica quem responde aos estímulos. Uma variedade de potenciais interessados podem examinar os resultados ou descobertas das avaliações e processar as informações. Os principais usuários são os primários, as partes interessadas mais estreitamente associadas ao programa que está sendo avaliado, que estão potencialmente em posição de tomar decisões com base nas informações. Além dos usuários primários, outros usuários locais do programa podem estar em posição de influenciar decisões, ações ou entendimento. Greene (1988) *apud* Alkin e King (2017) destaca a importância do envolvimento das partes interessadas, distinguindo entre a “pessoa muito envolvida”, “a pessoa um pouco ou às vezes envolvida” e a “pessoa marginalmente envolvida”. A extensão e a maneira como essas pessoas podem influenciar o uso potencial dos resultados, determinam até que

ponto elas são incluídas como parte desta definição.

Em relação aos tipos de influência da avaliação, os autor sugere três maneiras pelas quais o uso dos resultados de uma avaliação podem afetar um programa: como uma influência dominante, como uma das múltiplas influências em um determinado momento, ou como uma das múltiplas influências ao longo do tempo. Os tomadores de decisão podem se fazer valer dos resultados apresentados no relatório de avaliação, considerando que os resultados fornecidos pelos testes concordem com os apresentados no relatório de avaliação, ou os tomadores de decisão podem empregar informações de avaliação juntamente com outros insumos não avaliados, como custos do programa ou recursos distritais. Uma terceira possibilidade é que a avaliação possa ser uma das múltiplas e cumulativas influências. Por exemplo, a avaliação pode fundamentar as conclusões de um relatório anterior e, portanto, ser suficiente para estimular as partes interessadas a agir (ALKIN; KING, 2017).

Por fim, os aspectos e a proposta do programa estão associados aos usuários finais, seja diretamente (por exemplo, um membro da equipe que adapta as atividades do programa) ou indiretamente (por exemplo, um membro influente da comunidade em posição de influenciar o programa) que tenha a capacidade de afetar o desempenho do programa avaliado. Assim, as pessoas podem usar as informações de avaliação de diferentes maneiras: (1) tomando decisões, (2) estabelecendo ou alterando atitudes, (3) fundamentando decisões ou ações anteriores, ou (4) construindo a capacidade de avaliação de um indivíduo ou de uma organização. Em todos os casos, é útil determinar a motivação e a intenção dos usuários em potencial (ALKIN; KING, 2017).

Em relação as avaliações da educação, mais especificamente as avaliações padronizadas, estas devem ser utilizadas não apenas por quem desenha e avalia as políticas públicas, mas também pelos níveis decisórios intermediários e locais, incluindo equipes de gestão e professores (CONTRERAS; TORRES, 2022). Esse tipo de avaliação visa medir os desempenhos cognitivos dos alunos e constituem evidências sobre quais decisões são tomadas. No entanto, de acordo com Contreras e Torres (2022), tem havido uma subutilização destas informações ou uma discrepância entre as finalidades para as quais foram concebidas e a forma como são utilizadas.

4.2 Fatores que afetam o uso das avaliações

O estudo dos usos das avaliações educacionais é praticamente inseparável dos fatores que as afetam, no entanto, de acordo com Contreras e Torres (2022), esse aspecto tem recebido pouca atenção na América Latina. Além disso, as evidências mostram que, apesar do crescente volume de dados disponíveis, as equipes pedagógicas continuam tomando decisões com base na intuição, em vez de resultados de pesquisas e avaliações.

São diversos os obstáculos e fatores que impedem ou afetam a utilização dos resultados das avaliações. A utilização das avaliações para prestações de contas é mencionado por muitos trabalhos como um dos principais obstáculos uma vez que existe uma relação psicológica entre os propósitos de melhoria da escola e a responsabilização do baixo desempenho. Com frequência, as equipes

pedagógicas atribuem os resultados insatisfatórios dos alunos às suas características pessoais ou socio-demográficas, dificultando uma abordagem de treinamento (LASATER; BENGTON; ALBILADI, 2021).

Por meio de entrevistas direcionadas a atores educacionais, Contreras e Torres (2022) constata que uma grande barreira para o uso eficaz dos dados, a nível organizacional, seria a falta de tempo da equipe e a alta carga administrativa. Nesses casos, reconhece-se a importância e o desejo de analisar as informações, no entanto, os espaços de encontro são limitados e a possibilidade de os dirigentes escolares prestarem apoio pedagógico fica relegada às tarefas administrativas.

O desenho dos relatórios de resultados e acompanhamento dos usuários do programa também atuam como obstáculos de sua efetiva utilização. Existiria uma dificuldade de analisar os resultados dado a grande quantidade de informações, por vezes difíceis de compreender. De forma concomitante haveria a falta de habilidade para interpretação e análise de dados em alguns profissionais, bem como uma cultura de profissionais pouco acostumados a conceber ações concretas a partir dos resultados das avaliações (CONTRERAS; TORRES, 2022).

Com o objetivo de analisar de que forma a Secretaria de Educação do Distrito Federal utiliza as informações produzidas pelo Saeb, para gerir o ensino fundamental da rede pública, Neto (2006) conclui que a Secretaria não utiliza as informações geradas pelos resultados do Saeb no sentido de nortear o desenvolvimento de programas e ações para a solução dos problemas apontados pela avaliação. Os testes estariam sendo utilizados não como um instrumento para avaliar os sistemas educacionais e auxiliar no processo de aprendizagem dos alunos, mas sim como um instrumento de regulação do trabalho realizado pela escola e por seus profissionais.

O autor identificou como fatores que limitam o uso dos resultados a dificuldade de lidar com os conceitos definidos pela avaliação, provavelmente devido à dificuldade de compreensão da metodologia utilizada e pelas constantes mudanças que ocorreram na forma de interpretar os resultados ao longo dos ciclos de avaliação. Além disso, existiria uma ausência de um planejamento estratégico, apontando objetivos e definindo metas e instrumentos de acompanhamento de resultados (NETO, 2006).

Outro fator limitador seria a dificuldade de obter recursos orçamentários suficientes e o número insuficiente de pessoal especializado para executar eficientemente um grande conjunto de ações, fazendo com que muitas ações sejam iniciadas e poucas concluídas (NETO, 2006).

Segundo Coburn e Turner (2011), a maneira que os indivíduos irão utilizar os resultados das avaliações depende de como eles percebem, interpretam e constroem uma compreensão das implicações dos dados para a ação. Esses processos interpretativos são influenciados por contextos complexos e multifacetados, incluindo rotinas de uso de dados, normas de interação, relações de poder e autoridade, entre outros.

A falta de conexão entre intervenção, contexto, processo e resultados não beneficia melhorias no sistema educacional. Na ausência de informações sobre o processo de uso de dados, é possível saber

que algo leva a um determinado resultado, mas não se sabe como ou por quê. Sem informações sobre o contexto, não se pode explicar por que as mesmas ferramentas ou iniciativas promovem resultados positivos em alguns ambientes e em outros não. Ao mesmo tempo, informações sobre contexto e processo na ausência de informações sobre resultados pouco ajudam os formuladores de políticas e líderes de escolas a tomar decisões informadas sobre se determinadas intervenções de uso de dados valem o investimento de seus esforços e recursos escasso (COBURN; TURNER, 2011).

Outro fator que influencia o uso dos dados das avaliações são as mudanças organizacionais que podem ser bastante frequentes. A mudança na política e no direcionamento estratégico, ou nas formas como o trabalho e os papéis de trabalho são organizados, têm o potencial de alterar amplamente a prática individual do administrador, gestor e professor. Sendo assim, a atenção aos resultados relacionados à mudança organizacional é uma abordagem potencialmente frutífera para entender as consequências do uso de dados em uma escala maior e se estendendo além do alcance de uma única intervenção (COBURN; TURNER, 2011).

Schildkamp et al. (2017) procuram verificar quais fatores influenciam o uso de dados e quais efeitos percebidos pelas equipes escolares da Suécia. Os resultados qualitativo mostram que o trabalho das equipes de dados é influenciado pelos mesmos fatores que na Holanda: características dos dados (relevância dos dados), características da equipe (heterogeneidade da equipe) e características organizacionais da escola (por exemplo, apoio do líder da escola).

Os autores observaram que, várias características de usuários e equipes também podem atuar como facilitadores ou barreiras no desenvolvimento do trabalho, que será viabilizado se algum de seus membros já possuem algum conhecimento e habilidades sobre como usar os dados (ou seja, alfabetização de dados), ter conhecimento pedagógico do conteúdo e ter uma atitude positiva. Outro fator relevante pontuado é a importância dos membros da equipe trabalharem diretamente com o problema, além de terem metas compartilhadas.

4.3 Recomendações para o melhor uso das avaliações

De acordo com Schildkamp et al. (2017), os fatores que favorecem o uso dos dados podem ser classificados em três tipos: características dos dados, qualidades do usuário e condições organizacionais. Os primeiros incluem aspectos como a pontualidade da entrega dos resultados, sua acessibilidade, usabilidade, confiabilidade ou precisão. Em relação às qualidades do usuário, destacam-se as habilidades de compreensão e interpretação de dados, com ênfase para a alfabetização estatística como fator-chave e aspecto a ser promovido na formação inicial de professores, gestores e atores educacionais.

Dentro das condições organizacionais, destacam-se a liderança dos diretores e o apoio ao uso dos dados, a visão compartilhada e as normas institucionais, o tempo e os espaços disponíveis para o trabalho em torno dos dados, além da colaboração em equipe que tem se mostrado facilitadora no uso de dados (SCHILDKAMP et al., 2017).

Na mesma direção, Contreras e Torres (2022) também destacam como aspectos organizacionais

que facilitam o uso das avaliações, a existência de uma cultura de uso de dados, a instalação dessa prática na rotina da organização, a existência de espaços e o trabalho colaborativo em torno dos dados. Além disso, apoio por meio de reuniões e palestras facilitaria a compreensão e o uso dos dados. Também em relação às características organizacionais, Schildkamp et al. (2017) ponderam que elas podem tanto possibilitar quanto dificultar o trabalho das equipes de dados e o gestor líder possui papel importante nesse processo. Ele pode habilitar o trabalho das equipes de dados, fornecendo tempo aos membros, incentivando e motivando os membros da equipe, deixando claro o objetivo e a importância da equipe e distribuindo as tarefas de liderança e manipulação de dados.

De acordo com Neto (2006), alguns gestores sentiriam falta de um planejamento efetivo para orientar a utilização das avaliações, indicando que planejamento poderia ser útil para apontar caminhos e acompanhar os resultados. Além disso, se faz necessário que os relatórios apontem não somente os resultados numéricos, mas que sejam disponibilizadas informações complementares que ofereçam uma visão mais clara sobre os problemas educacionais apontados pela avaliação. Neto (2006) também considera que os dados apurados pelo Saeb, não são suficientes, per se, para orientar a tomada de decisões. É necessário que cada sistema de ensino desenvolva pesquisas adicionais, contando com o suporte técnico do Inep, para estudar os aspectos mais problemáticos apontados pelo Saeb.

Em relação à qualidade do usuário, fica clara a necessidade de promover a alfabetização de dados, tanto no processo de manipulação quando de interpretação, aos atores educacionais. De acordo com Mandinach e Gummer (2013), o domínio de uso de dados é composto por cinco componentes sob os quais associa-se conhecimentos e habilidades específicas. Os domínios incluem: identificar problemas e formular perguntas a partir do uso dos dados, transformar dados em informações, transformar informações em decisões e avaliar resultados. A maioria das decisões dos processos não são lineares ou mesmo finitos.

Schildkamp (2019), reitera que embora muitas intervenções de desenvolvimento profissional de uso de dados estejam disponíveis em todo o mundo, muitas dessas intervenções não foram estudadas sistematicamente. Os autores identificaram e analisaram 11 intervenções de uso de dados que foram estudadas cientificamente. Com base nessas intervenções foi possível identificar vários recursos de desenvolvimento profissional efetivo em serviço no uso de dados. Por exemplo, criar estruturas e protocolos para desenvolver o uso de dados, fornecer desenvolvimento profissional por um período de tempo mais longo, e fazer o link entre dados e instruções explícitas. Os autores acreditam que esses recursos devem ser levados em consideração ao treinar os atores educacionais no uso de dados de avaliação.

Neste sentido, especialmente no primeiro ano, é importante que o treinador de dados dirija o processo e apoie a equipe na coleta e análise de dados. No segundo ano, os líderes podem reduzir lentamente esse papel mais ativo para estimular a independência e a sustentabilidade. No entanto, são necessárias mais pesquisas sobre seu papel no ensino e análise dos dados. Por exemplo, sobre quais funções são cruciais para o progresso de uma equipe (SCHILDKAMP et al., 2017).

Outra questão que precisa ser analisada é em relação à composição mais eficaz da equipe de análise de dados. As diferenças entre os membros da equipe podem ser estimulantes se a equipe tiver discussões que levam a conflitos cognitivos, isto é, diferenças de opinião, premissas e ideias, por exemplo em relação às ações a serem tomadas com base em dados específicos, pode ser positiva para a equipe (SCHILDKAMP et al., 2017).

Mandinach e Schildkamp (2021) fazem algumas recomendações para que o processo de utilização dos dados se torne um instrumento eficiente. Eles recomendam que não se comece com os dados, mas com objetivos claros e mensuráveis. Além de utilizar diferentes fontes de dados, para capturar as necessidades educacionais, é preciso envolver-se coletivamente em um processo de criação de sentido, por exemplo, por meio de uma equipe de dados. Também é necessário conectar o julgamento profissional e o uso de dados para aumentar a qualidade de tomada de decisão.

5. Considerações finais

Essa pesquisa buscou identificar os recursos, competências e práticas organizacionais, necessários para fazer o melhor uso dos resultados de avaliações padronizadas, por parte dos órgãos públicos e atores de interesse.

Foram identificados como principais competências e práticas organizacionais a cultura de uso de dados, a instalação dessa prática na rotina da organização, a existência de espaço e tempo, o trabalho colaborativo entre os atores e a estabilidade em relação à mudança organizacional. Do ponto de vista da qualidade do usuário, destaca-se o desenvolvimento de habilidades estatísticas e capacidade de transformar os dados em tomada de decisão. Os principais domínios incluem: identificar problemas e formular perguntas, usar dados, transformar dados em informações, transformar informações em decisões e avaliar resultados. Sobre as características dos dados, faz-se necessário garantir a confiabilidade, validade, objetividade e o detalhamento da informação.

De acordo com Neto (2006), cabe ao Inep disseminar entre os gestores públicos de todos os níveis os conhecimentos a percepção dos significados que envolvem os resultados de uma avaliação. Esta seria a condição para que se possa compreender a realidade das escolas e buscar soluções para melhorar o seu desempenho. A Plataforma Foco Aprendizagem (PFA)¹, adotada por diversos estados brasileiros, seria um exemplo de formação integrada e continuada às diferentes instâncias e aos profissionais da rede pública educacional. A PFA é um exemplo de política de gestão que visa não somente fornecer dados interpretáveis, mas também instruir os gestores em como utilizar os dados de avaliação.

De acordo com as conclusões de Schildkamp (2019), quando se trata de usar dados para melhorar a qualidade do ensino e da aprendizagem, é evidente que alguns dos facilitadores e barreiras importantes incluem alfabetização de dados e liderança. No entanto, o que é menos compreendido é como podemos promover os facilitadores e remover as barreiras para desbloquear, de forma mais completa, o potencial de uso de dados. Só então a utilização dos dados pode levar a uma melhoria escolar sustentável. Portanto, são necessárias mais pesquisas que busquem identificar os caminhos necessários para promover a real alfabetização de dados, não somente no âmbito da manipulação, mas também do entendimento e interpretação dos mesmos.

¹A plataforma pode ser acessada pelo link: <<https://www.focoaprendizagem.com.br/>>

Referências Bibliográficas

ALKIN, M. C.; DAILLAK, R.; WHITE, P. Using evaluations. *Beverly Hills: Sage*, 1979. Citado 2 vezes nas páginas 7 e 12.

ALKIN, M. C.; KING, J. A. Definitions of evaluation use and misuse, evaluation influence, and factors affecting use. *American Journal of Evaluation*, SAGE Publications Sage CA: Los Angeles, CA, v. 38, n. 3, p. 434–450, 2017. Citado 3 vezes nas páginas 12, 13 e 14.

BRASKAMP, L. A. A definition of use. *Studies in Educational Evaluation*, ERIC, v. 8, n. 2, p. 169–74, 1982. Citado na página 12.

COBURN, C. E.; TURNER, E. O. Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research & Perspective*, Taylor & Francis, v. 9, n. 4, p. 173–206, 2011. Citado 3 vezes nas páginas 6, 15 e 16.

CONTRERAS, J.; TORRES, Á. Perspectivas de directivos educacionales sobre el uso de pruebas estandarizadas. *Revista Mexicana de*, p. 511, 2022. Citado 5 vezes nas páginas 3, 5, 14, 15 e 16.

FAO. *Monitoring and Evaluation for learning and performance improvement*. 2022. <<https://www.fao.org/investment-learning-platform/themes-and-tasks/monitoring-and-evaluation/en/#c325490>>. Acessado em: 2022-10-25. Citado 2 vezes nas páginas 5 e 9.

GREENE, J. G. Stakeholder participation and utilization in program evaluation. *Evaluation review*, Sage Publications Sage CA: Thousand Oaks, CA, v. 12, n. 2, p. 91–116, 1988. Citado na página 13.

HILHORST, T.; GUIJT, I. Participatory monitoring and evaluation a process to support governance and empowerment at the local level. *Guidance Paper*. Amsterdam: KIT, 2006. Citado 2 vezes nas páginas 8 e 10.

HUBERMAN, M. Steps toward an integrated model of research utilization. *Knowledge*, Sage Publications, v. 8, n. 4, p. 586–611, 1987. Citado 2 vezes nas páginas 7 e 12.

INEP. Sistema de avaliação da educação básica - documentos de referência - versão preliminar. Ministério da Educação, 2019. Citado na página 5.

LAHEY, R. *The Canadian M&E System*. [S.l.]: Washington, DC: World Bank, 2010. Citado na página 9.

LASATER, K.; BENGTON, E.; ALBILADI, W. S. Data use for equity?: How data practices incite deficit thinking in schools. *Studies in Educational Evaluation*, Elsevier, v. 69, p. 100845, 2021. Citado na página 15.

LOPEZ-ACEVEDO, G.; KRAUSE, P. *Building better policies: The nuts and bolts of monitoring and evaluation systems*. [S.l.]: World Bank Publications, 2012. Citado 2 vezes nas páginas 8 e 9.

MANDINACH, E. B.; GUMMER, E. S. A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, SAGE Publications Sage CA: Los Angeles, CA, v. 42, n. 1, p. 30–37, 2013. Citado 2 vezes nas páginas 3 e 17.

MANDINACH, E. B.; SCHILDKAMP, K. Misconceptions about data-based decision making in education: An exploration of the literature. *Studies in Educational Evaluation*, Elsevier, v. 69, p. 100842, 2021. Citado 2 vezes nas páginas 3 e 18.

NETO, J. L. H. Avaliação externa: a utilização dos resultados do saeb 2003 na gestão do sistema público de ensino fundamental no distrito federal. 2006. Citado 4 vezes nas páginas 5, 15, 17 e 19.

PATTON, M. Q. *Utilization-focused evaluation*. [S.l.]: Sage publications, 1978. Citado 2 vezes nas páginas 7 e 12.

SCHEERENS, J. et al. *Educational evaluation, assessment, and monitoring: A systemic approach*. [S.l.]: Taylor & Francis, 2003. v. 13. Citado 6 vezes nas páginas 5, 6, 7, 10, 11 e 12.

SCHILDKAMP, K. Data-based decision-making for school improvement: Research insights and gaps. *Educational research*, Taylor & Francis, v. 61, n. 3, p. 257–273, 2019. Citado 2 vezes nas páginas 17 e 19.

SCHILDKAMP, K. et al. Factors promoting and hindering data-based decision making in schools. *School effectiveness and school improvement*, Taylor & Francis, v. 28, n. 2, p. 242–258, 2017. Citado 3 vezes nas páginas 16, 17 e 18.

Evidência Express

EVEX **ENAP**