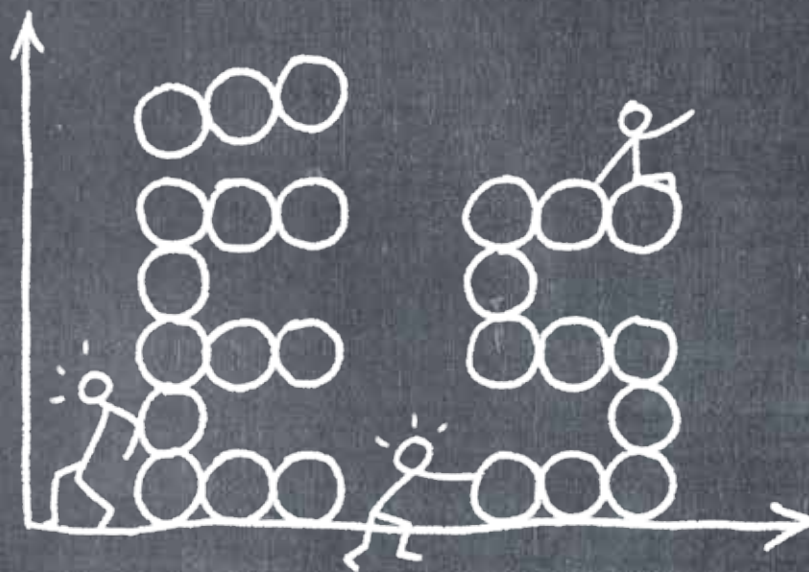


# LES ÉPREUVES STANDARDISÉES :

Comment sont-elles perçues par les acteurs concernés ?



PRINCIPAUX CONSTATS ET CLARIFICATIONS



UNIVERSITY OF LUXEMBOURG  
Luxembourg Centre for  
Educational Testing (LUCET)

**Titre :**

Les Épreuves Standardisées : comment sont-elles perçues par les acteurs concernés ?

**Sous-titre :**

Principaux constats et clarifications

**Directeurs de la publication :**

Sonja Ugen, Romain Martin, Antoine Fischbach

**Lieu :**

Esch/Alzette

**Institution :**

Université du Luxembourg, Luxembourg Centre for Educational Testing (LUCET)

**Auteurs :**

Christophe Dierendonck (ECCS/LLLG), Pascale Esch (LUCET), Antoine Fischbach (LUCET), Romain Martin (LUCET), Marianne Milmeister (LUCET), Paul Milmeister (ECCS/LLLG), Sonja Ugen (LUCET), Christiane Weis (LUCET)

**Layout et illustrations :**

Sons Design, Bonn

**Copyright :**

Université du Luxembourg

**Année :**

2017

**URL :**

lucet.uni.lu  
www.epstan.lu

**ISBN :**

ISBN 978-99959-0-352-7

# LES ÉPREUVES STANDARDISÉES :

Comment sont-elles perçues par  
les acteurs concernés ?

PRINCIPAUX CONSTATS ET CLARIFICATIONS



UNIVERSITY OF LUXEMBOURG  
Luxembourg Centre for  
Educational Testing (LUCET)

Auteurs :

Sonja Ugen

Pascale Esch

Antoine Fischbach

## Préface

**La mission principale, octroyée par le gouvernement, et la raison d'être du Luxembourg Centre for Educational Testing (LUCET) est l'évaluation externe du système scolaire.**

Depuis les années 2000, les constats issus d'études telles que PISA (*Programme for International Student Assessment*) ont renforcé la prise de conscience des défis auxquels le système scolaire luxembourgeois se voit confronté, à savoir une population scolaire de plus en plus hétérogène et un curriculum exigeant en terme d'apprentissages à réaliser. Face à ces observations, des changements conceptuels majeurs ont été opérés sur le plan du développement et de l'assurance de la qualité scolaire, à savoir la transition vers une approche fondée sur les acquis scolaires ainsi qu'une volonté d'implémenter une évaluation systématique du système scolaire fournissant un appui nécessaire à toute politique éducative ambitieuse.

Pas d'évaluation sans auto-évaluation. C'est dans cette optique que le LUCET s'engage à être régulièrement évalué à différents niveaux.

La structure du LUCET inclut un *scientific advisory board* constitué de spécialistes nationaux et internationaux dans le domaine du monitoring scolaire, externes au LUCET. Sur base annuelle, le LUCET leur présente ses méthodes de travail, comme par exemple le développement d'outils de mesure ou encore les méthodes statistiques utilisées et reçoit ensuite un retour détaillé avec des conseils d'amélioration. Par ailleurs, le LUCET participe aux

évaluations externes de l'Université du Luxembourg, qui sont commanditées par le gouvernement et qui vérifient régulièrement la qualité de l'output scientifique et de l'enseignement. L'output scientifique des chercheurs mêmes du LUCET est aussi évalué par des pairs à chaque fois qu'ils soumettent leurs travaux pour être publiés dans des journaux scientifiques.

Au niveau interne, le LUCET se constitue de différents labs dont les membres se réunissent régulièrement pour échanger et se donner un retour sur leurs travaux. Par exemple, les tâches qui font partie des Épreuves Standardisées (ÉpStan) sont discutées lors des réunions du *test development lab* qui réunit tous les chercheurs responsables des groupes développant ces tâches, afin de garantir la meilleure qualité de ces épreuves.

Dans l'idée de s'améliorer et de s'adapter continuellement, le LUCET s'est intéressé à la perception des Épreuves Standardisées par les acteurs concernés. La transmission des résultats se fait actuellement aux niveaux système (scolaire), école, classe et élève. En admettant que le processus d'évaluation ne soit pas une fin en soi, ces retours aux différents niveaux ont pour ambition d'augmenter l'utilité voire la plus-value du dispositif et d'en faire un outil de travail pouvant s'intégrer au quotidien scolaire. ●

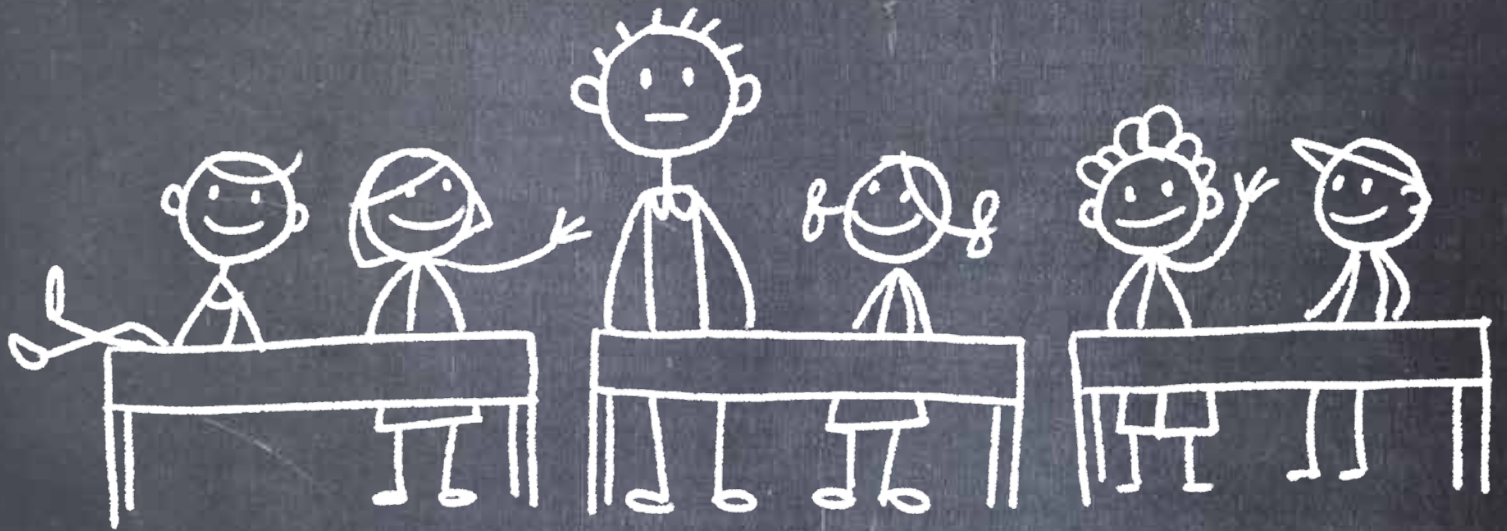


# SOMMAIRE

02	Préface
<b>05</b>	<b>1   VERS UNE DÉMARCHE D'ASSURANCE-QUALITÉ POUR LES ÉPREUVES STANDARDISÉES</b>
06	1.1 Contexte et aperçu
07	1.2 Méthodologie
<b>09</b>	<b>2   DANS QUELLE MESURE LE PRINCIPE D'UNE ÉVALUATION EXTERNE DES ACQUIS SCOLAIRES DES ÉLÈVES – ET EN PARTICULIER LE DISPOSITIF ÉPSTAN – EST-IL SOUTENU PAR LES ACTEURS SCOLAIRES ?</b>
10	2.1 Acceptation des évaluations externes en général et des ÉpStan en particulier
11	2.2 Compréhension du dispositif de monitoring du système scolaire
12	2.3 Compréhension du sens et de l'utilité des ÉpStan
13	2.4 Avis des acteurs concernés par rapport au timing des ÉpStan
14	2.5 Avis des acteurs concernés par rapport à l'idée d'élever les enjeux des ÉpStan
15	2.6 Avis des acteurs concernés par rapport au contenu et au format des tests
16	2.7 Avis des acteurs concernés par rapport au questionnaire des élèves
17	2.8 Avis des acteurs concernés par rapport aux processus d'information et de communication autour du dispositif ÉpStan
<b>19</b>	<b>3   DANS QUELLE MESURE LES ÉLÈVES FONT-ILS LES TESTS ÉPSTAN SÉRIEUSEMENT ?</b>
20	3.1 Sérieux des élèves pendant les ÉpStan
25	3.2 Facteurs qui influencent les attitudes des élèves et des enseignants à l'égard des ÉpStan
<b>27</b>	<b>4   DANS QUELLE MESURE LES RAPPORTS ENVOYÉS AUX ACTEURS DANS LE CADRE DES ÉPSTAN SONT-ILS COMPRIS ET UTILISÉS ?</b>
28	4.1 Aspects des rapports qui suscitent le plus d'intérêt auprès des concernés
28	4.2 Améliorations relatives aux rapports qui pourraient intéresser les concernés
30	4.3 Compréhension des rapports « école », « classe » et « élève » par les acteurs concernés
31	4.4 Utilisation des rapports
33	4.5 Impact des résultats ÉpStan
34	4.6 Difficultés liées à l'accès aux différents rapports
35	4.7 Questions et revendications des concernés à propos des résultats
36	4.8 Risques et appréhensions des concernés concernant les rapports
<b>37</b>	<b>LISTE DES RÉFÉRENCES</b>
<b>38</b>	<b>LISTE DES ABRÉVIATIONS</b>

## LISTE DES ENCADRÉS EXPLICATIFS

10	<b>Encadré 1 :</b> Utilité des évaluations externes
11	<b>Encadré 2 :</b> Cadre des évaluations externes
12	<b>Encadré 3 :</b> Objectif des ÉpStan
13	<b>Encadré 4 :</b> Timing des ÉpStan
14	<b>Encadré 5 :</b> Enjeux des ÉpStan
15	<b>Encadré 6 :</b> Élaboration des ÉpStan
16	<b>Encadré 7 :</b> Prise en compte de la composition de la population scolaire dans les ÉpStan
17	<b>Encadré 8 :</b> Communication autour des ÉpStan
24	<b>Encadré 9 :</b> Implication des élèves lors des ÉpStan
25	<b>Encadré 10 :</b> Attitude des élèves (du secondaire) à l'égard des ÉpStan
29	<b>Encadré 11 :</b> Apports et limites des rapports ÉpStan
30	<b>Encadré 12 :</b> Présentation des rapports ÉpStan
32	<b>Encadré 13 :</b> Utilisations potentielles des rapports ÉpStan
33	<b>Encadré 14 :</b> Impact potentiel des ÉpStan (classe/école)
34	<b>Encadré 15 :</b> Accès aux rapports ÉpStan
35	<b>Encadré 16 :</b> Impact des résultats des ÉpStan (niveau système)
36	<b>Encadré 17 :</b> Appréhensions face aux résultats des ÉpStan



*Auteurs :*

*Christophe Dierendonck*

*Marianne Milmeister*

*Paul Milmeister*

*Christiane Weis*

*Sonja Ugen*

*Antoine Fischbach*

*Romain Martin*

# 1 Vers une démarche d'assurance-qualité pour les Épreuves Standardisées

05

Depuis la première édition des Épreuves Standardisées (ÉpStan) durant l'année scolaire 2008-2009, le dispositif a considérablement évolué sur les plans conceptuel, organisationnel et technologique, mais également en termes de rapports adressés aux différents acteurs concernés par ces évaluations externes des acquis scolaires. Ces transformations successives ont été opérées en fonction des contraintes et des opportunités rencontrées au fil du temps, mais sans référence explicite à une démarche d'assurance-qualité globale. ●

## 1.1 CONTEXTE ET APERÇU

Le dispositif étant stabilisé depuis 2010-2011, le moment semblait idéal pour prendre du recul et initier une démarche d'assurance-qualité. En 2013, il a donc été décidé de collecter des données auprès des acteurs concernés par les ÉpStan, tant en externe au niveau des acteurs du terrain scolaire (directions, coordinateurs ÉpStan dans les lycées, enseignants du fondamental et du secondaire, inspecteurs, présidents des comités d'école, élèves, parents et syndicats actifs dans le domaine de l'éducation) qu'en interne au niveau de l'ancienne Agence pour le développement de la qualité scolaire (ADQS)<sup>1</sup> du Service de Coordination de la Recherche et de l'Innovation pédagogiques et technologiques (SCRIPT) du Ministère de l'Éducation nationale, de l'Enfance et de la Jeunesse (MENJE) et du *Luxembourg Centre for Educational Testing* (LUCET) de l'Université du Luxembourg (UL). Le but de cette collecte de données était de permettre la conduite d'une démarche d'assurance-qualité en fondant les réflexions et les éventuelles décisions sur des constats tirés de données empiriques.

L'étude s'est articulée autour de trois questions principales qui ont progressivement émergé lors de réunions de travail de l'équipe ÉpStan, lors de plusieurs échanges informels avec des enseignants et suite à certaines observations menées dans les classes durant les tests :

- (1) Dans quelle mesure le principe d'une évaluation externe des acquis scolaires des élèves est-il soutenu par les acteurs scolaires ?
- (2) Dans quelle mesure les élèves font-ils les tests ÉpStan sérieusement ?
- (3) Dans quelle mesure les rapports envoyés aux acteurs dans le cadre des ÉpStan sont-ils compris et utilisés ?

Pour répondre à ces questions, quatre types de données ont été récoltées. Les opinions de tous les types d'acteurs concernés par le dispositif ÉpStan ont été sondées par des *focus groups* (a). Les enseignants du cycle 3.1 et de V<sup>e</sup>/9<sup>e</sup> ont en outre été invités à compléter un questionnaire à large échelle (b). Au niveau des élèves de V<sup>e</sup>/9<sup>e</sup>, deux autres types de données ont été recueillies : des données auto-rapportées de motivation et d'effort ont été collectées au moyen de questions posées au début et à la fin de chacun des trois tests ÉpStan (c) et des données comportementales ont été enregistrées par la plate-forme de testing lors de la passation des tests par les élèves (d).

Le présent texte a pour but de synthétiser les résultats de recherche et de tirer un bilan intermédiaire de la démarche d'assurance-qualité entreprise. Il est structuré en fonction des trois questions de recherche présentées ci-dessus. Dans chaque partie sont repris deux types de propos : d'une part, des constats formulés par les auteurs de l'étude à partir des données récoltées et, d'autre part, des clarifications sur l'outil des ÉpStan présentées sous la forme d'encadrés explicatifs rédigés par l'équipe ÉpStan.

Avant de poursuivre, que les représentants du terrain soient remerciés pour leur précieuse collaboration. Sans leur participation au questionnaire et l'échange franc lors des *focus groups*, cette étude n'aurait pas été possible. ●

<sup>1</sup> Suite à la restructuration du SCRIPT en 2017, le service comprend actuellement six divisions. La division « du traitement de données sur la qualité de l'encadrement et de l'offre scolaire et éducative » reprend notamment la partie monitoring du portfolio de l'ancienne « Agence pour le Développement de la Qualité Scolaire » (ADQS).



## 1.2 MÉTHODOLOGIE

Les méthodologies utilisées pour les différentes collectes de données, de type quantitatif et qualitatif, ont été décrites de façon précise dans le chapitre 7 du rapport national 2011-2013 des ÉpStan (Dierendonck et al., 2015, p.99-109).

Seules quelques informations essentielles sont reprises ici pour permettre au lecteur de comprendre la nature des données récoltées :

**a)** En 2014, douze **focus groups** réunissant au total 77 participants ont été organisés. Ces discussions en groupe ont rassemblé différents participants : des directeurs de lycée, des coordinateurs ÉpStan des lycées, des enseignants du cycle 3.1 et de V<sup>e</sup>/9<sup>e</sup>, des inspecteurs, des élèves de V<sup>e</sup>/9<sup>e</sup>, des représentants de syndicats actifs dans le domaine de l'éducation, des membres de l'ancienne ADQS (SCRIPT) et des membres du LUCET. Faute de volontaires, deux *focus groups* prévus initialement n'ont pas pu avoir lieu : ceux avec les parents d'élèves du cycle 3.1 et avec les présidents des comités d'école.

**b)** Les enseignants du cycle 3.1 et les professeurs d'allemand, de français et de mathématiques de V<sup>e</sup>/9<sup>e</sup> ont été sollicités en 2014 pour participer à l'étude par **questionnaire**.

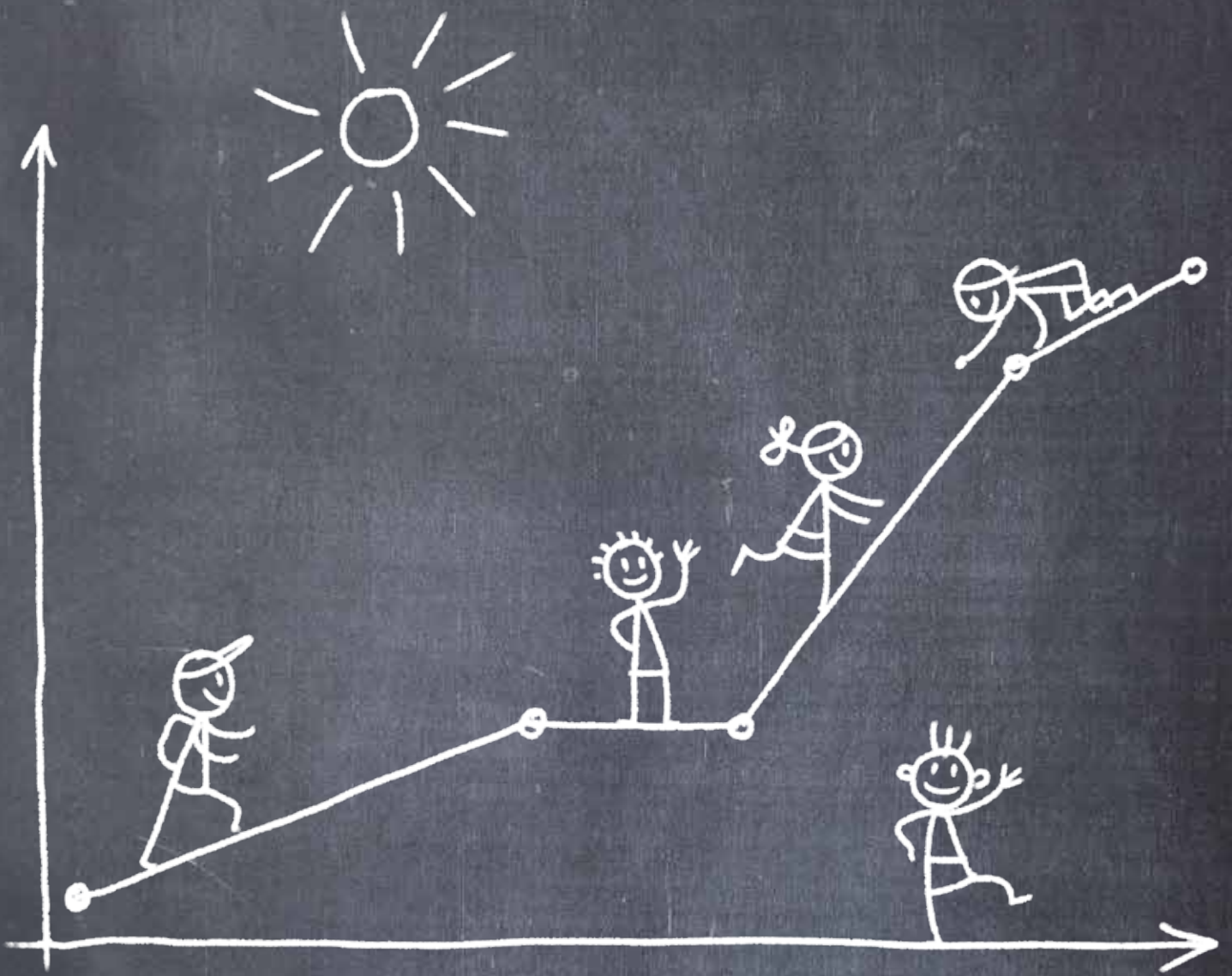
Sur l'ensemble des sollicitations, 152 enseignants du cycle 3.1 et 131 enseignants de V<sup>e</sup>/9<sup>e</sup> se sont connectés à la plate-forme de sondage. Les participants avec un taux d'omissions aux questions de plus de 25% ont été retirés des échantillons. Au final, l'échantillon pour le cycle 3.1 est constitué de 150 enseignants (taux de participation d'environ 41%) et celui pour la V<sup>e</sup>/9<sup>e</sup> se compose de 121 enseignants (taux de participation d'environ 13%) dont 43 professeurs de mathématiques, 34 professeurs d'allemand et 44 professeurs de français. Il est difficile de spécifier le nombre exact d'enseignants sollicités, même en se basant sur les listes des classes ayant participé aux ÉpStan : au secondaire, certains enseignants exercent dans plusieurs classes de V<sup>e</sup>/9<sup>e</sup> et au fondamental, des enseignants travaillent en *team teaching* ou à temps partiel. Le nombre d'enseignants sollicités a été estimé autour de 360 pour

le fondamental et autour de 1010 pour le secondaire. Les taux de participation mentionnés ci-dessus sont par conséquent aussi à considérer comme une estimation de la participation des enseignants.

**c)** Les **données auto-rapportées de motivation et d'effort** ont été recueillies lors des ÉpStan de V<sup>e</sup>/9<sup>e</sup>. Au début de chacun des trois tests, les élèves ont été invités à estimer, sur une échelle à quatre modalités, leur degré de motivation pour le test. En fin de test, ils ont dû estimer l'effort consenti pour le test et l'effort qu'ils auraient consenti si le test avait compté comme note scolaire. L'instrument utilisé pour mesurer l'effort des élèves (composé d'échelles de 1 à 10) est une adaptation du « thermomètre d'effort » développé dans le cadre de PISA (*Programme for International Student Assessment*) de l'OCDE (Dierendonck et al., 2013).

**d)** Afin de disposer d'un autre indicateur du sérieux des élèves lors des tests ÉpStan, il a semblé intéressant d'exploiter les potentialités du testing par ordinateur en enregistrant une série de **données de type comportemental**. Actuellement, il est possible d'exploiter les informations rendant compte du temps durant lequel chaque stimulus<sup>2</sup> a été affiché à l'écran. En totalisant les temps d'affichage de tous les stimuli, on dispose d'une variable rendant compte du temps total *a priori* consacré au traitement des stimuli par les élèves. Il faut nuancer qu'il est impossible d'assimiler avec certitude le temps d'affichage des tâches au temps investi effectivement par les élèves pour les traiter : on peut en effet envisager la possibilité que lors de l'affichage des questions, les élèves regardent (ou non) l'écran, sans véritablement traiter les tâches et les questions proposées. Cette variable est, à l'heure actuelle, le meilleur indicateur à disposition pour tenter d'estimer objectivement le degré de sérieux avec lequel les élèves ont pris part au test. ●

<sup>2</sup> Les épreuves sont constituées de plusieurs stimuli (par exemple, différents textes dans le test d'allemand) et plusieurs questions sont associées à chaque stimulus.



*Auteurs :*

*Paul Milmeister  
Christiane Weis  
Christophe Dierendonck  
Marianne Milmeister*

## 2 Dans quelle mesure le principe d'une évaluation externe des acquis scolaires des élèves – et en particulier le dispositif ÉpStan – est-il soutenu par les acteurs scolaires ?

09

La présente étude a voulu recueillir l'avis des acteurs de l'enseignement par rapport à l'idée générale d'une évaluation externe des acquis scolaires et par rapport aux ÉpStan en particulier. Cette partie traitera de sujets comme l'acceptation générale des évaluations externes, la compréhension du dispositif de monitoring du système scolaire, la compréhension du sens et de l'utilité des ÉpStan et d'autres sujets qui ont émergé des entretiens. Il s'agit notamment de questions en rapport avec les caractéristiques du dispositif ÉpStan (timing, enjeux, contenu des tests, questionnaire des élèves) ainsi que d'observations relatives aux processus d'information et de communication mis en place actuellement. ●

## 2.1 ACCEPTATION DES ÉVALUATIONS EXTERNES EN GÉNÉRAL ET DES ÉPSTAN EN PARTICULIER

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS* ET DU QUESTIONNAIRE

Les enseignants interrogés par questionnaire, aussi bien ceux du fondamental que ceux du secondaire, semblent être favorables au principe d'une évaluation externe des acquis scolaires des élèves. Ainsi, 9 enseignants sur 10 considèrent l'organisation de telles épreuves comme légitime et semblent d'avis que ces épreuves externes pourraient constituer un outil complémentaire à l'évaluation interne qu'ils pratiquent quotidiennement. Parmi les participants des *focus groups*, l'idée d'une évaluation externe semble être généralement acceptée par les interlocuteurs qui voient comme principal point positif, l'objectivité des évaluations externes. Cela dit, d'après les données du questionnaire, 4 enseignants sur 10 au cycle 3.1 et 6 enseignants sur 10 au secondaire rejettent l'idée que les évaluations externes soient plus objectives que leurs propres évaluations. Plus largement, 8 enseignants sur 10 jugent qu'ils restent les mieux placés pour construire des évaluations valides et fiables.

Dans les *focus groups*, les enseignants soulignent qu'il est important que ces tests externes soient utiles à leur pratique de terrain. Les enseignants ayant participé à l'enquête par ques-

tionnaire apprécient généralement l'idée d'une évaluation externe qui permet, au travers d'une comparaison objective, de situer les performances des élèves de leur école ou de leur classe par rapport à la moyenne nationale. Par rapport à une appréciation individuelle de l'élève, ils sont respectivement 8 enseignants sur 10 au cycle 3.1, 6 enseignants sur 10 en V<sup>e</sup> et 7 enseignants sur 10 en 9<sup>e</sup> à considérer que les évaluations externes confirment leurs propres évaluations.

Par ailleurs, même si les interlocuteurs des *focus groups* adhèrent au principe général d'une évaluation externe, ils rendent toutefois compte d'une certaine méfiance à l'égard d'une évaluation qui les touche directement, notamment parce qu'ils craignent qu'elle soit utilisée pour évaluer leur travail. Certains enseignants ne reconnaissent pas la plus-value des ÉpStan et les considèrent comme une chose parmi tant d'autres qui se rajoute et parfois même dérange l'activité principale d'enseignement du programme scolaire. D'après quelques interlocuteurs, les ÉpStan sont placées au rang d'activités secondaires, noyées par les affaires courantes de la vie scolaire. ●

#### Les évaluations externes des acquis scolaires des élèves sont-elles utilisées pour évaluer les enseignants ?

Non. Les évaluations externes des acquis scolaires comme PISA ou les ÉpStan n'ont pas pour but primaire d'analyser la performance moyenne des classes ou la compétence individuelle des élèves. Dans les rapports internationaux de PISA qui sont rédigés par l'OCDE, les informations récoltées sont agrégées pour comparer les moyennes de chaque pays. Dans les rapports nationaux de PISA, rédigés par l'Université du Luxembourg et le MENJE, on procède à des comparaisons entre les types d'enseignement et entre les établissements qui accueillent des publics scolaires similaires. Ces comparaisons sont cependant toujours anonymes : il est impossible d'identifier des établissements en particulier mais cela permet de montrer que certains établissements sont plus efficaces que d'autres.

La finalité principale des ÉpStan se situe au niveau du système éducatif. Il s'agit avant tout d'opérer un suivi longitudinal de l'efficacité du système au travers d'un monitoring annuel, à des fins de pilotage politique (cf. encadrés 2 et 3). Le dispositif permet aussi de mener des comparaisons entre les types d'enseignement et d'adresser des rapports au niveau de l'établissement scolaire, au niveau de la classe et au niveau de l'élève. Les rapports envoyés au niveau de l'établissement scolaire ne reprennent pas les résultats de chacune des classes mais présentent les résultats pour la population entière de l'école. L'accès au « rapport classe » est limité aux seuls enseignants concernés ; ni l'inspecteur, ni le directeur ne peut savoir comment une classe particulière a travaillé (cf. encadré 14) et ne peut donc pas utiliser ces résultats pour évaluer le travail d'un enseignant.

#### En quoi les évaluations externes sont-elles complémentaires aux évaluations réalisées par les enseignants ?

Les évaluations externes et les évaluations des enseignants présentent des avantages et des désavantages spécifiques. Un avantage des évaluations externes est de proposer un dispositif standardisé tant au niveau du contenu des épreuves d'évaluation que des conditions de passation et des modalités de correction des tests. Il s'agit donc ici de limiter au maximum les biais potentiels de mesure liés aux conditions d'administration et de correction des épreuves ainsi qu'au degré d'exigence variable des enseignants. De nombreuses études empiriques ont en effet montré qu'à performances identiques lors de devoirs en classe, les notes attribuées aux élèves pouvaient diverger fortement selon les enseignants.

Un autre avantage des évaluations externes, en particulier des ÉpStan ou de PISA, est de permettre une comparaison dans le temps et, potentiellement, un suivi longitudinal pour mesurer les progrès réalisés par le système.

À l'inverse des évaluations conduites par les enseignants, les évaluations externes ne portent que sur une partie du curriculum enseigné, ceci essentiellement pour des raisons de temps d'évaluation disponible, mais aussi parce que certaines compétences scolaires sont très difficiles à évaluer de manière standardisée (p. ex. production écrite pour les ÉpStan). Les évaluations externes comme les ÉpStan procèdent à l'évaluation d'un échantillon des compétences à maîtriser, et ce au départ de questions fermées ou à réponses brèves permettant une correction objective. →

→ Les évaluations des enseignants s'inscrivent quant à elles dans l'exhaustivité et le long terme, en proposant des questions qui demandent parfois des réponses très élaborées dont l'évaluation est évidemment plus riche, mais plus difficile aussi.

La complémentarité entre les évaluations externes et les évaluations des enseignants réside donc essentiellement

dans l'opportunité de pouvoir disposer avec les évaluations externes d'une mesure des compétences des élèves qui soit comparable d'une classe ou d'une école à l'autre, permettant ainsi aux enseignants de situer le niveau de leurs élèves par rapport aux autres élèves testés avec le même instrument et dans des conditions comparables.

← Encadré 1 : Utilité des évaluations externes

## 2.2 COMPRÉHENSION DU DISPOSITIF DE MONITORING DU SYSTÈME SCOLAIRE

### CONSTATS ISSUS DES DONNÉES DES FOCUS GROUPS

Le dispositif de monitoring du système scolaire semble peu clair. Les interlocuteurs disent ne pas comprendre les différences entre les évaluations externes existantes et leur insertion dans un dispositif ou une politique plus large. Les inspecteurs mentionnent la brochure du Ministère sur le monitoring et le pilotage du système scolaire<sup>3</sup>, publiée en 2007, mais regrettent qu'il n'y ait pas eu de mise à jour de ce document de référence. Les interlocuteurs déplorent l'absence d'un concept global de monitoring et de pilotage qui intégrerait toutes les évaluations externes internationales et nationales, qui clarifierait la manière dont elles s'articulent, quels objectifs particuliers elles poursuivent

et comment elles se complètent. Selon les interlocuteurs, cette mission de définition ou de clarification revient aussi bien au MENJE qu'à l'Université. Ils soulignent également l'importance d'une stratégie globale pour la politique d'éducation (par exemple clarifier et communiquer si l'on se situe ou non dans l'optique d'un pilotage du système par les résultats) qui fournirait le cadre dans lequel les ÉpStan et les autres évaluations externes pourraient devenir des outils importants. D'après certains inspecteurs, la définition d'une telle stratégie relève d'une décision et d'une communication politiques. ●

11

#### Les évaluations externes actuelles s'intègrent-elles dans un cadre conceptuel clair ?

Le document de référence de 2007 « Die Steuerung des Luxemburger Schulwesens » définit les grandes lignes du pilotage du système éducatif luxembourgeois. Ce document annonce notamment le passage d'un système dont la gestion se focalise sur les ressources et les conditions d'apprentissage (*inputs*) vers un système qui se concentre davantage sur les résultats obtenus (*outputs*). Cette évolution est à l'image de ce qui s'est fait dans d'autres systèmes éducatifs. Une première étape était de définir l'ensemble des connaissances théoriques et pratiques (socles de compétences) à maîtriser aux différents moments de la scolarité. Les écoles et les enseignants sont relativement autonomes concernant le choix des moyens pour arriver aux résultats attendus. En contrepartie, afin d'assurer la qualité de l'enseignement, des évaluations externes des acquis scolaires des élèves sont prévues, l'objectif étant d'un côté, de contrôler si les socles de compétences ont été atteints et de l'autre côté, de voir si les socles ont été fixés à un niveau approprié.

Les évaluations externes internationales et nationales au Luxembourg s'inscrivent donc dans une approche d'assurance de la qualité de l'enseignement. La loi de 2009 portant restructuration du SCRIPT (Service de Coordination de la Recherche et de l'Innovation pédagogiques et techno-

logiques), dont une des missions est de promouvoir et de mettre en œuvre l'assurance de la qualité de l'enseignement, spécifie qu'il y a une évaluation interne et externe du système éducatif qui porte, d'un côté, sur la qualité de l'enseignement dans les écoles et les lycées et, de l'autre côté, sur les compétences atteintes par les élèves à différents niveaux de leur scolarité. Toutes les écoles fondamentales et secondaires sont censées participer aux évaluations externes de la qualité de l'enseignement mises en œuvre par le SCRIPT.

Les ÉpStan (élaborées par le LUCET au sein de l'Université du Luxembourg) sont un outil externe parmi d'autres pour évaluer le système scolaire. Elles ont pour finalité principale de fournir au Ministère une vue d'ensemble des acquis scolaires à certaines étapes de la scolarité obligatoire (grades 1-3-5-7-9) et de permettre le pilotage ainsi que le suivi de l'impact des réformes. Quant aux Épreuves Communes élaborées par le MENJE en interne, elles sont de deux types : les Épreuves Communes de cycle 4 servent à orienter les élèves en fin d'école fondamentale tandis que les Épreuves Communes (ÉpCom) de V<sup>e</sup> ES/9<sup>e</sup> EST (grade 9), qui comptent comme un devoir en classe, sont plutôt utilisées pour réguler le degré d'exigence des enseignants en proposant des critères de correction communs.

← Encadré 2 : Cadre des évaluations externes

<sup>3</sup> Il s'agit d'un numéro spécial du Courrier de l'Éducation nationale, intitulé « Die Steuerung des Luxemburger Schulwesens », qui explique le cadre théorique et décrit les objectifs que le Ministère s'est donnés en matière de monitoring et de pilotage du système éducatif.

## 2.3 COMPRÉHENSION DU SENS ET DE L'UTILITÉ DES ÉPSTAN

### CONSTATS ISSUS DES DONNÉES DES FOCUS GROUPS

Une partie considérable des interlocuteurs dit ne pas bien comprendre la finalité, le sens et l'utilité du dispositif ÉpStan. Ils se posent beaucoup de questions à cet égard, et des hypothèses sont parfois formulées : certains pensent par exemple que les ÉpStan servent à préparer les élèves à PISA, à faire un ranking des lycées ou encore à évaluer le travail des enseignants. Par ailleurs, certains coordinateurs ÉpStan dans les lycées, personnes *a priori* bien informées, se disent incapables d'expliquer à leurs collègues la finalité de ces épreuves externes. Il semble évident, et les interlocuteurs le confirment d'ailleurs, que cette compréhension lacunaire ou erronée de la finalité du dispositif ÉpStan ait un impact négatif sur son acceptation.

Les interlocuteurs disent avoir du mal à voir la plus-value des ÉpStan et semblent chercher l'utilité du dispositif principalement à leur niveau d'action. Par exemple, quelques enseignants du cycle 3.1 apprécieraient de pouvoir utiliser les ÉpStan comme une aide à leur propre évaluation des élèves. D'autres interlocuteurs considèrent par contre que les ÉpStan sont plus utiles à certains niveaux qu'à d'autres : d'après quelques professeurs, l'utilité des ÉpStan se situe surtout au niveau des directions, mais pas à leur niveau, et encore moins au niveau des élèves. Des directeurs estiment que l'utilité au niveau national prévaut sur celle

au niveau interne de l'établissement scolaire. D'après certains coordinateurs, l'utilité se trouve plutôt au niveau de la comparaison de cohortes ou bien de la comparaison des écoles avec une moyenne nationale, alors que des inspecteurs voient une utilité potentielle surtout au niveau d'un suivi longitudinal des élèves qui permettrait de voir leur évolution individuelle. Des interlocuteurs pensent que les ÉpStan pourraient être utiles au niveau national pour évaluer et adapter des documents de référence comme le Plan d'études du fondamental ou les programmes scolaires du secondaire. ●

#### Quels sont le but et l'utilité des ÉpStan ?

Le but des ÉpStan est de dresser un rapport fiable sur l'enseignement par une évaluation nationale de type standardisé des compétences des élèves. Ce monitoring du système éducatif doit permettre le pilotage et le développement du système éducatif au départ de données objectives (*evidence-based policy*). De plus, l'outil permet de comparer les performances globales des élèves d'une année à l'autre et de dégager des tendances. Il est également possible d'envisager un suivi longitudinal de cohortes tout au long de la scolarité obligatoire. Cette opportunité distingue les ÉpStan d'autres épreuves.

À la différence d'autres dispositifs de monitoring à l'étranger, les ÉpStan concernent l'ensemble des élèves d'une cohorte (et non un échantillon). Cette particularité s'explique par la taille du pays : pour sélectionner un échantillon d'élèves représentatif sur le plan statistique, il faudrait en effet inclure quasiment tous les élèves de la cohorte concernée. La décision de tester tous les élèves d'une cohorte a donc été prise. Elle implique que les résultats sont disponibles aussi au niveau individuel pour chaque élève. Au fil des années et des échanges avec les acteurs de terrain, il a paru intéressant, en parallèle à l'objectif principal de monitoring du système, de partager ces résultats avec les enseignants, les élèves et les parents qui peuvent y trouver aussi une utilité (cf. encadré 13).

#### Quelle est la plus-value des ÉpStan pour les acteurs du terrain ?

À côté des avantages du dispositif en termes de monitoring et de pilotage du système, il existe également une plus-value des ÉpStan pour les acteurs du terrain. Ces épreuves aboutissent en effet à la rédaction de rapports (école, classe, élève) qui permettent une comparaison objective puisque les conditions de passation sont standardisées et que le format des questions du test ÉpStan limite au maximum les biais de correction.

Les indicateurs de compétence des élèves obtenus avec les ÉpStan peuvent donc être considérés comme complémentaires à l'évaluation opérée par l'enseignant (cf. encadré 1).

La comparaison d'une classe ou d'une école avec les scores moyens nationaux permet de les situer par rapport à une population de référence et de relativiser les performances des élèves. En outre, l'intervalle de performance attendue (*Erwartungsbereich*), qui est renseigné au niveau de l'établissement scolaire et de la classe, permet une comparaison équitable. Cet intervalle de performance attendue indique l'intervalle dans lequel on retrouve la majorité des classes ou des écoles comparables en termes de population scolarisée (cf. encadré 7).

## 2.4 AVIS DES ACTEURS CONCERNÉS PAR RAPPORT AU TIMING DES ÉPSTAN

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS* ET DU QUESTIONNAIRE

La problématique du timing a été traitée par les enseignants dans l'enquête par questionnaire dans le cadre d'une prise de position par rapport à des pistes de modification éventuelles du dispositif ÉpStan. La proposition de déplacer le moment de l'évaluation en fin d'année scolaire reçoit un avis favorable auprès de 7 enseignants sur 10 au cycle 3.1 et 6 enseignants sur 10 en V<sup>e</sup>/9<sup>e</sup>. La formulation de la proposition, « Les Épreuves Standardisées devraient être menées en fin d'année scolaire et pas en début », ne permet cependant pas de dire avec certitude si l'évaluation devrait avoir lieu en fin de l'année précédente (cycle 2.2 ou VI<sup>e</sup>/8<sup>e</sup>) ou en fin de l'année du testing actuel.

Dans les *focus groups*, des instituteurs questionnent la pertinence d'administrer les ÉpStan en début de cycle 3.1. Comme c'est la maîtrise du programme du cycle précédent qui est évaluée, les résultats ne seraient pas très utiles aux

enseignants du cycle 3.1 et leur causeraient parfois même des problèmes, par exemple lorsque des élèves qui sont au cycle 3.1 n'atteignent pas, selon les ÉpStan, les socles du cycle précédent. C'est pourquoi certains jugent préférable de faire les ÉpStan à un autre moment : à la fin du cycle 2.2 ou au milieu du cycle 3.1. Ceci permettrait à l'enseignant du cycle 2.2 de voir le niveau de l'élève à la fin de l'année ou à l'enseignant du cycle 3.1 de tirer des conclusions et d'ajuster le tir au cours du cycle.

Le délai entre l'administration des épreuves et la mise à disposition des rapports est également critiqué dans la mesure où ces rapports arriveraient trop tard pour être réellement utiles aux enseignants et pertinents pour les élèves et leurs parents. Les instituteurs aimeraient les avoir plus tôt dans l'année, idéalement pour pouvoir accompagner le premier bilan. ●

#### Pourquoi les ÉpStan ont-elles lieu au début de l'année scolaire ?

Les ÉpStan évaluent la maîtrise d'une sélection de compétences du cycle précédent et la décision de conduire les épreuves au début de l'année scolaire s'appuie sur les volontés suivantes :

- (1) S'assurer que l'évaluation externe porte bien sur des contenus dont l'enseignement a été finalisé dans toutes les classes.
- (2) Donner aux enseignants qui viennent de prendre en charge une classe une vue complémentaire, de mettre à leur disposition des informations sur leurs élèves et d'offrir par conséquent aux enseignants une opportunité pour ajuster éventuellement leurs pratiques d'enseignement.
- (3) Maintenir les ÉpStan au rang d'épreuves à faibles enjeux (c'est-à-dire qui ne comptent pas pour l'évaluation de l'année). Les enseignants ne sont pas censés les utiliser pour déterminer les appréciations des bilans ou calculer les notes des bulletins. Si les ÉpStan étaient conduites en fin de l'année, elles entreraient en concurrence avec les évaluations de fin d'année des enseignants et elles pourraient influencer la décision des enseignants de faire redoubler ou de faire réussir les élèves.
- (4) Éviter un *teaching to the test* en classe.
- (5) Permettre aux élèves de démontrer leurs performances sans trop de stress.
- (6) Ne pas permettre l'utilisation des ÉpStan pour évaluer le travail des enseignants du cycle précédent. Ce risque existe évidemment, mais il est moins important que si les épreuves étaient organisées en fin d'année scolaire.

#### Pourquoi les rapports ÉpStan viennent-ils « si tard » ?

Les rapports issus des ÉpStan sont mis à la disposition des acteurs fin janvier-début février.

Pour être en mesure de produire les rapports plus tôt, il faudrait que les épreuves aient lieu plus tôt, tel que c'était le cas au début des ÉpStan. En effet, les premières années, les épreuves se déroulaient en octobre, tout de suite après la rentrée scolaire. Ceci causait un certain nombre de problèmes logistiques, notamment parce que les écoles et les lycées étaient en pleine organisation scolaire et que les bases de données officielles, indispensables pour connaître le nombre d'élèves et la composition des classes, n'étaient pas encore actualisées à ce stade. Voilà pourquoi il a été décidé, notamment à la demande des coordinateurs ÉpStan dans les lycées, de repousser les épreuves au mois de novembre. À l'heure actuelle, la collecte des données est clôturée fin novembre-début décembre et les analyses statistiques débutent tout de suite après une série de contrôles de qualité.

En tout, ce sont environ 24 000 élèves qui participent chaque année aux ÉpStan. Dans ce dispositif, il faut produire environ 60 000 rapports individualisés. Compte tenu du nombre d'élèves testés, on peut donc dire que les rapports ÉpStan sont disponibles rapidement. En guise de comparaison, les rapports PISA, qui ne sont pas personnalisés, ne sont publiés qu'un an après les épreuves.

## 2.5 AVIS DES ACTEURS CONCERNÉS PAR RAPPORT À L'IDÉE D'ÉLEVER LES ENJEUX DES ÉPSTAN

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS* ET DU QUESTIONNAIRE

Dans l'enquête par questionnaire, l'idée d'élever les enjeux des ÉpStan pour les élèves a été soumise aux enseignants comme une piste de modification éventuelle du dispositif ÉpStan. Cette proposition, « Les Épreuves Standardisées devraient compter pour des points », relativement vague quant à son implémentation et à l'importance à accorder aux ÉpStan dans l'évaluation périodique ou annuelle, reçoit un avis défavorable auprès de 56% des enseignants au cycle 3.1 et de 62% des enseignants en V<sup>e</sup>/9<sup>e</sup>.

On retrouve cette attitude plutôt négative à l'idée de faire compter les ÉpStan chez les élèves du secondaire interviewés dans les *focus groups*. Ils préfèrent la situation actuelle puisqu'il y a moins de stress et de nervosité.

Pour les enseignants du fondamental et du secondaire, les coordinateurs ÉpStan et les membres de direction, les opinions sont plus mitigées. Certains défendent l'idée que les élèves et les enseignants seraient plus motivés à faire les ÉpStan si les épreuves comptaient, d'une manière ou d'une autre, pour l'évaluation de l'année. D'autres interlocuteurs adoptent une vision plus négative en énumérant les désavantages que cette nouvelle donnerait : cela augmenterait la pression sur les élèves et inciterait les enseignants au *teaching to the test*. Certains professeurs estiment également qu'élever les enjeux des ÉpStan ne pourrait s'envisager qu'à certaines conditions : il faudrait pouvoir retracer ce qui a été demandé exactement lors des tests et rendre leur correction plus transparente. ●

#### Pourquoi les ÉpStan ne comptent-elles pas ?

Comme développé précédemment (cf. encadré 3), les ÉpStan visent en premier lieu le niveau du système et il n'y a pas de volonté d'élever les enjeux de ces épreuves.

Contrairement aux devoirs en classe qui contrôlent si la matière enseignée pendant quelques semaines a été bien assimilée, les ÉpStan vérifient si certains socles de compétences définis pour un niveau d'études donné ont été atteints, voire dépassés. Voilà pourquoi les tests sont construits de façon à mesurer l'atteinte des socles de compétence, mais également certaines compétences en-dessous et au-delà des socles dans l'optique de situer les compétences actuelles des élèves. Les analyses statis-

tiques utilisées pour calculer les résultats sont complexes et prennent en compte tous les facteurs possibles (ordre de passation du test, ordre des questions,...) qui pourraient influencer les résultats pour refléter au mieux les compétences des élèves. Il s'avérerait difficile de traduire ce résultat en note scolaire classique.

Par ailleurs, les ÉpStan sont à faibles enjeux comme la plupart des évaluations externes (entre autres PISA) pour éviter que les élèves ne ressentent trop de stress et que les enseignants ne soient incités à préparer leurs élèves à ces tests (effet *teaching to the test*).



## 2.6 AVIS DES ACTEURS CONCERNÉS PAR RAPPORT AU CONTENU ET AU FORMAT DES TESTS

### CONSTATS ISSUS DES DONNÉES DES FOCUS GROUPS ET DU QUESTIONNAIRE

Par rapport au contenu et au format des tests, les enseignants du secondaire ont des perceptions différentes que ceux du fondamental. Ces divergences s'expliquent par le fait que les enseignants du cycle 3.1 ont eu, contrairement aux enseignants de V<sup>e</sup>/9<sup>e</sup>, l'occasion de voir l'ensemble des questions (items) composant les tests en les administrant dans leur classe.

D'après l'enquête par questionnaire, l'avis des enseignants du cycle 3.1 est généralement positif quant aux tests de compréhension en allemand et quant à l'épreuve de mathématiques : la majorité des répondants jugent les items des trois tests comme étant d'un niveau de difficulté adapté, similaires à ce qui est proposé par les enseignants lors de leurs activités d'enseignement ou lors de leurs évaluations et en cohérence avec les documents de référence. Dans les *focus groups*, les principales critiques avancées par les instituteurs tournent autour de la question de la couverture du programme scolaire par les ÉpStan : d'un côté, ils critiquent que certains items se réfèrent à des compétences qui sont au programme du cycle suivant et de l'autre côté,

ils regrettent que quelques-unes des compétences les plus importantes, en particulier l'expression écrite, fassent défaut. Concernant le format des épreuves, une partie des instituteurs se réjouit de leur belle apparence, tandis que d'autres pensent que le format (test QCM où il faut cocher des cases) est assez inhabituel et donc potentiellement problématique pour les enfants de cet âge.

Le nombre restreint d'enseignants de V<sup>e</sup>/9<sup>e</sup>, qui ont pu voir les items et qui ont répondu au questionnaire, sont sensiblement moins positifs que ceux du cycle 3.1, notamment en termes de niveau de difficulté adapté et de similarité entre les items du test et les tâches utilisées par les enseignants lors de leurs activités d'enseignement ou lors de leurs évaluations. Dans les *focus groups*, les critiques formulées par les enseignants de V<sup>e</sup>/9<sup>e</sup> sont moins concrètes. Ils disent ne pas comprendre comment les tests sont construits et comment ils peuvent être adaptés aux différents niveaux des élèves. Certains estiment que les tests ne sont pas vraiment adaptés au niveau des élèves, surtout concernant les plus faibles (au préparatoire en particulier). ●

#### Comment les tests des ÉpStan sont-ils élaborés ?

Le développement des items commence en général une année avant le test. Afin de garantir la qualité des tests, les items pour chaque niveau scolaire et chaque matière sont élaborés par des enseignants qui travaillent ou ont travaillé dans le niveau et la branche scolaires concernés. Ils sont encadrés sur le plan scientifique et méthodologique par les chercheurs du LUCET. Les groupes de travail se réunissent à intervalles réguliers et discutent des nouveaux items proposés (compétence testée, niveau de difficulté, complexité de la tâche, adaptation au groupe cible, qualité psychométrique des items). Les items sont modifiés jusqu'à un accord unanime sur leur contenu et leurs caractéristiques. Les items finalisés sont pré-testés dans quelques classes l'année scolaire précédant le test. Sur base de ces résultats, les items sont soit modifiés soit jugés inadaptés et alors éliminés.

Le développement des épreuves se fait en fonction d'un cadre conceptuel d'élaboration (cadre développé sur la base des socles de compétence à atteindre à la fin des niveaux scolaires ciblés). Ce cadre spécifie le nombre d'items nécessaires par type de compétence et par niveau de difficulté. La compilation du test tient également compte d'autres critères comme l'ordre des items, l'alternance des formats de réponse ou la place des items d'ancrage longitudinal (pour la comparaison d'une année à l'autre) et d'ancrage transversal (pour la comparaison entre types d'enseignement).

Pour garantir que les tests aient la meilleure qualité possible, les chercheurs responsables pour le développement d'items se réunissent aussi régulièrement en interne pour discuter du développement des ÉpStan en général.

#### Pourquoi les tests ne peuvent-ils pas évaluer toutes les compétences du programme scolaire ?

Les ÉpStan se focalisent sur les compétences pour lesquelles un testing standardisé est possible. Voilà pourquoi la plupart des items sont à format fermé (questions à choix multiples, vrai-faux, mettre les items dans le bon ordre, etc.) ou ne requièrent que des réponses courtes. L'expression écrite ne fait pas partie des compétences testées étant donné qu'un tel test standardisé serait très laborieux et coûteux à mettre en œuvre, aussi bien du point de vue de la définition de critères de correction que de l'évaluation elle-même (formation de correcteurs spécialisés).

#### Comment est-il possible d'élaborer un test unique qui soit adapté aux élèves de niveaux aussi différents ?

Au niveau des épreuves du secondaire, trois versions sont élaborées pour chaque test de compétence, une par type d'enseignement (versions es, est et est-pr). Ces tests ont des degrés de difficulté adaptés aux différentes filières. Afin de garantir la comparabilité des résultats entre les filières, les versions de chaque test comprennent des items spécifiques et des items communs pour assurer un ancrage transversal et la comparabilité entre types d'enseignement.

Les enseignants du fondamental et du secondaire intéressés à participer au développement des tests peuvent nous contacter : [info@epstan.lu](mailto:info@epstan.lu)

## 2.7 AVIS DES ACTEURS CONCERNÉS PAR RAPPORT AU QUESTIONNAIRE DES ÉLÈVES

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS*

Le questionnaire soumis aux élèves<sup>4</sup> suscite intérêt et réactions positives quand il s'agit de récolter des informations sur la motivation scolaire et le climat de classe. Par contre, la partie du questionnaire sur les données socio-économiques est vue d'un œil critique car touchant à la sphère privée. Les interlocuteurs se posent des questions par rap-

port au but et à la pertinence de telles questions. Certains voient le risque d'une interprétation déterministe des résultats basés sur les données socio-économiques et ils soulèvent que l'analyse en fonction du statut socio-économique n'apporte, de toute façon, aucune piste de solution pour dépasser ces constats connus depuis longtemps. ●

#### Pourquoi a-t-on besoin de récolter des données sur les caractéristiques familiales des élèves ?

Des études nationales, comme le projet « Matière grise perdue (MAGRIP) » par exemple, et internationales, comme PISA par exemple, ont montré qu'il existe des disparités importantes en termes de compétences entre différents sous-groupes de la population scolaire. La situation est d'autant plus alarmante que ces différences se retrouvent, voire s'accroissent, tout au long du parcours scolaire. Afin de pouvoir mesurer les effets du système éducatif et de comprendre les dynamiques sous-jacentes, il s'avère indispensable de prendre en compte, dans le monitoring scolaire, les caractéristiques familiales des élèves sur les plans social, culturel, économique et linguistique.

Dans le cas des ÉpStan, les données récoltées via les questionnaires des élèves et des parents sont utilisées pour mener des analyses et présenter les résultats en fonction du statut socio-économique, du sexe, de l'origine migratoire et des langues parlées à la maison. Elles servent également à définir, de manière probabiliste, un intervalle de performance attendue (*Erwartungsbereich*), c'est-à-dire un intervalle statistique dans lequel on pourrait s'attendre que la performance d'une classe ou d'une école se situe

compte tenu de la composition de sa population scolaire. Pour le calcul de l'intervalle de performance attendue, plusieurs données personnelles des élèves sont retenues : le sexe, la ou les langues parlées à la maison, l'origine migratoire, le statut socio-économique (estimé au départ d'un indice de richesse, du nombre de livres dans le foyer et de l'occupation professionnelle des parents), l'année de naissance et le parcours scolaire antérieur (redoublement, scolarité effectuée ou non au Luxembourg). Par ailleurs, la composition de la classe est aussi considérée (p.ex. nombre de filles/garçons).

L'objectif principal est de tenir compte du contexte et de la population d'une école ou d'une classe pour interpréter les résultats et permettre une comparaison équitable des performances.

← Encadré 7 : Prise en compte de la composition de la population scolaire dans les ÉpStan

<sup>4</sup> Au secondaire, les élèves sont priés de remplir un questionnaire portant sur les données socio-économiques et socio-culturelles de la famille, sur le parcours scolaire antérieur, sur l'intérêt scolaire et le climat scolaire. Au fondamental, ces informations sont fournies en partie par les élèves et en partie par leurs parents.

## 2.8 AVIS DES ACTEURS CONCERNÉS PAR RAPPORT AUX PROCESSUS D'INFORMATION ET DE COMMUNICATION AUTOUR DU DISPOSITIF ÉPSTAN

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS*

En ce qui concerne l'information donnée en amont quant au déroulement des épreuves, les concernés jugent être assez bien informés, grâce notamment à une documentation très complète. Plusieurs problèmes d'information et de communication existent cependant à d'autres niveaux et expliqueraient en partie le manque de compréhension et d'acceptation du dispositif ÉpStan.

Les interlocuteurs esquissent des propositions visant une meilleure acceptation des ÉpStan et une plus grande motivation des concernés à travers une communication adéquate : miser sur les échanges entre les acteurs-clés du terrain (directions-enseignants, inspecteurs-enseignants,

enseignants-élèves) ; investir dans des formes de communication plus vivantes (films ou vidéos), interactives et personnalisées (réunions, conférences ou présentations) qui privilégient le contact personnel et/ou font en sorte que toutes les personnes concernées reçoivent des informations et explications de première main. Des interlocuteurs demandent par ailleurs une promotion plus ouverte du dispositif ÉpStan et un investissement dans le suivi des ÉpStan. Par « suivi », ils entendent le soutien de l'acteur du terrain ainsi que des explications quant à l'interprétation et l'utilisation des données. Ils s'attendent surtout à une prise de position face aux questions essentielles du terrain quant aux finalités poursuivies par les ÉpStan. ●

#### Quelles informations sont d'ores et déjà disponibles à propos des ÉpStan ?

La communication et la transparence sont sans doute des éléments importants pour la compréhension et l'acceptation du dispositif ÉpStan par les acteurs concernés.

Beaucoup d'efforts de communication ont été fournis ces dernières années. En guise d'exemple on peut citer le matériel qui est disponible sur le site Internet des ÉpStan ([www.epstan.lu](http://www.epstan.lu)) : manuels pour les enseignants, les coordinateurs et les responsables de test ; manuels en format vidéo ; vidéos d'information sur les ÉpStan ; exemples de questions du test ; explications sur les rapports école, classe et élève ; exemples de rapports ; rapport présentant les résultats au niveau national (« Épreuves Standardisées : Bildungsmonitoring für Luxemburg. Nationaler Bericht 2011 bis 2013 ») ; rapport technique (*ÉpStan Technical Report*) ; articles explicatifs adressés aux enseignants.

Des réunions annuelles s'adressant aux enseignants, plusieurs réunions d'information organisées avec la Fédération des associations de parents d'élèves du Luxembourg (FAPEL), des participations à des journées pédagogiques sur demande ainsi que des présentations ou conférences offrant la possibilité de poser des questions aux responsables des ÉpStan ont été organisées pour mieux informer les concernés.

Pendant toute l'année, une *helpline* (Tél. 46644-9777) et une adresse e-mail ([info@epstan.lu](mailto:info@epstan.lu)) permettent de se renseigner auprès de l'équipe ÉpStan.



*Auteurs :*

*Christiane Weis*

*Paul Milmeister*

*Marianne Milmeister*

*Christophe Dierendonck*

# 3 Dans quelle mesure les élèves font-ils les tests ÉpStan sérieusement ?

19

Les ÉpStan ont uniquement une fonction informative puisque les performances à ces tests ne sont pas prises en compte pour influencer ou déterminer le parcours scolaire des élèves. Face à ce genre d'évaluations externes « à faibles enjeux » pour les élèves, on peut interroger leur niveau d'effort fourni lors de la réalisation de ces tests.

Les résultats de recherche à ce sujet sont cependant contradictoires : certaines études ont mis en évidence que les élèves sont raisonnablement motivés à donner le meilleur d'eux-mêmes quand le test est à faibles enjeux (Baumert & Demmrich, 2001) tandis que d'autres ont conclu que les enjeux du test exerçaient effectivement une influence significative positive sur la motivation et la performance (Wolf & Smith, 1995). Les élèves placés dans une situation de testing à faibles enjeux seraient donc plus ou moins motivés et pourraient dès lors s'impliquer différemment, avec le risque que les performances des élèves moins impliqués lors du testing ne reflètent pas exactement leurs compétences réelles et que les indicateurs moyens calculés au niveau de la population soient sous-estimés.

Dans les lignes qui suivent, nous présentons les constats issus des données dont nous disposons actuellement au sujet du sérieux des élèves lors des ÉpStan et quelques facteurs ayant une influence sur les attitudes des enseignants et des élèves à l'égard des épreuves. ●

## 3.1 SÉRIEUX DES ÉLÈVES PENDANT LES ÉPSTAN

### CONSTATS ISSUS DES DONNÉES DU QUESTIONNAIRE ET DES *FOCUS GROUPS*

Pour les enseignants du fondamental, la très grande majorité des élèves du cycle 3.1 (d'après les données du questionnaire, 8 à 9 élèves sur 10 en moyenne) seraient très impliqués pendant les tests. Les enseignants du cycle 3.1, interviewés dans les *focus groups*, confirment cette tendance en décrivant le sérieux, la motivation, voire le plaisir que les élèves ont à faire les tests ÉpStan.

Au secondaire, la proportion des élèves jugés sérieux lors des ÉpStan chute quelque peu (6 à 7 élèves sur 10 en moyenne d'après les données du questionnaire). Dans les deux cas, les indicateurs moyens cacheraient cependant une grande variabilité entre les classes. Selon plusieurs coordinateurs et professeurs, interviewés dans les *focus groups*, une partie des élèves jugés peu sérieux répondraient trop rapidement afin de terminer au plus vite. Cela dit, ce type de comportements se retrouverait dans une

certaine mesure aussi lors des devoirs en classe.

Les élèves rencontrés dans les *focus groups* semblent accepter les ÉpStan comme une normalité à laquelle ils n'accordent cependant pas une importance considérable. La plupart des élèves interviewés disent avoir travaillé un peu moins sérieusement lors des tests ÉpStan que pendant les devoirs en classe. Ils estiment que leur effort pour les ÉpStan est comparable à celui qu'ils font pour les exercices qui ne comptent pas pour l'année. Certains apprécient que ce ne soit pas le même stress que pour les devoirs en classe et prennent un certain plaisir à faire les ÉpStan car cela signifie de passer une matinée sans cours, tout en apprenant éventuellement quelque chose de nouveau. Selon les enseignants interrogés par questionnaire, les élèves seraient en moyenne 1 sur 10 à être davantage stressés que lors des devoirs en classe. La proportion monte à 3 sur 10 au cycle 3.1. ●

### CONSTATS ISSUS DES DONNÉES AUTO-RAPPORTÉES PAR LES ÉLÈVES (V<sup>e</sup>/9<sup>e</sup>)

Les élèves de V<sup>e</sup>/9<sup>e</sup> ont été invités à estimer leur degré de motivation pour le test à l'entame de chacun des trois tests (« motivation initiale »). En fin de test, à partir d'une version adaptée du « thermomètre d'effort » (utilisé pour PISA), ils ont dû estimer leur effort pour le test (« effort consenti ») et l'effort qu'ils auraient fourni si le test avait compté comme note scolaire (« effort si noté »). La description statistique de ces données auto-rapportées est synthétisée dans les tableaux suivants en distinguant quatre groupes d'élèves selon la filière d'enseignement et l'orientation sui-

vies : (1) Enseignement Secondaire (ES), (2) Enseignement Secondaire Technique (EST) – Orientation théorique ou polyvalente, (3) Enseignement Secondaire Technique (EST) – Orientation pratique et (4) Enseignement Secondaire Technique – Régime Préparatoire (EST-PRE). Trois versions (es, est, est-pr) de chaque test ÉpStan sont développées pour s'ajuster au mieux à la compétence scolaire moyenne de chaque groupe d'élèves. Les groupes 1 et 2 reçoivent respectivement la version 'es' et la version 'est' tandis que les groupes 3 et 4 reçoivent la version 'est-pr'.

Groupes d'élèves		Test								
		allemand			français			mathématiques		
Filière	Version du test	N	Moyenne de la motivation initiale	Ecart-type	N	Moyenne de la motivation initiale	Ecart-type	N	Moyenne de la motivation initiale	Ecart-type
(1) ES	es	1743	2.8	0.8	1722	2.7	0.8	1746	2.7	0.8
(2) EST	est	2917	2.8	0.8	3183	2.7	0.9	3195	2.7	0.8
(3) EST	est-pr	671	2.8	0.8	671	2.7	0.9	684	2.7	0.8
(4) EST-PRE	est-pr	661	2.9	0.8	791	2.8	0.9	797	2.9	0.8
<b>Total</b>		<b>5992</b>	<b>2.8</b>	<b>0.8</b>	<b>6367</b>	<b>2.7</b>	<b>0.9</b>	<b>6422</b>	<b>2.7</b>	<b>0.8</b>

Modalités de réponse : 1 – pas du tout motivé(e), 2 – peu motivé(e), 3 – motivé(e), 4 – très motivé(e)

Tableau 1 : Moyenne de la « motivation initiale » selon les groupes d'élèves et les matières testées (sur une échelle de 1 à 4)

En termes de « motivation initiale », les valeurs moyennes au niveau de la population totale se situent autour de 2.7, soit entre les modalités « peu motivé(e) » et « motivé(e) »

pour chacun des tests (voir tableau 1). Le degré moyen de motivation à l'entame des tests semble être indépendant du groupe d'élèves considéré et de la matière testée.

Groupes d'élèves		Test								
		allemand			français			mathématiques		
Filière	Version du test	N	Moyenne de l'effort consenti	Ecart-type	N	Moyenne de l'effort consenti	Ecart-type	N	Moyenne de l'effort consenti	Ecart-type
(1) ES	es	1677	6.4	2.1	1701	6.7	2.1	1422	6.9	2.2
(2) EST	est	2856	6.1	2.3	3108	6.2	2.4	3020	6.3	2.4
(3) EST	est-pr	663	5.7	2.4	677	5.8	2.4	673	5.8	2.4
(4) EST-PRE	est-pr	625	5.5	2.6	738	5.9	2.6	724	5.9	2.6
<b>Total</b>		<b>5821</b>	<b>6.0</b>	<b>2.3</b>	<b>6224</b>	<b>6.3</b>	<b>2.4</b>	<b>5839</b>	<b>6.3</b>	<b>2.4</b>

Tableau 2 : Moyenne de l'« effort consenti » selon les groupes d'élèves et les matières testées (sur une échelle de 1 à 10)

Au niveau de l'« effort consenti » déclaré par les élèves, les valeurs moyennes calculées au niveau de la population totale sont relativement proches entre les matières testées : 6.0 pour le test d'allemand et 6.3 pour les tests de français et de mathématiques (voir tableau 2). On constate des

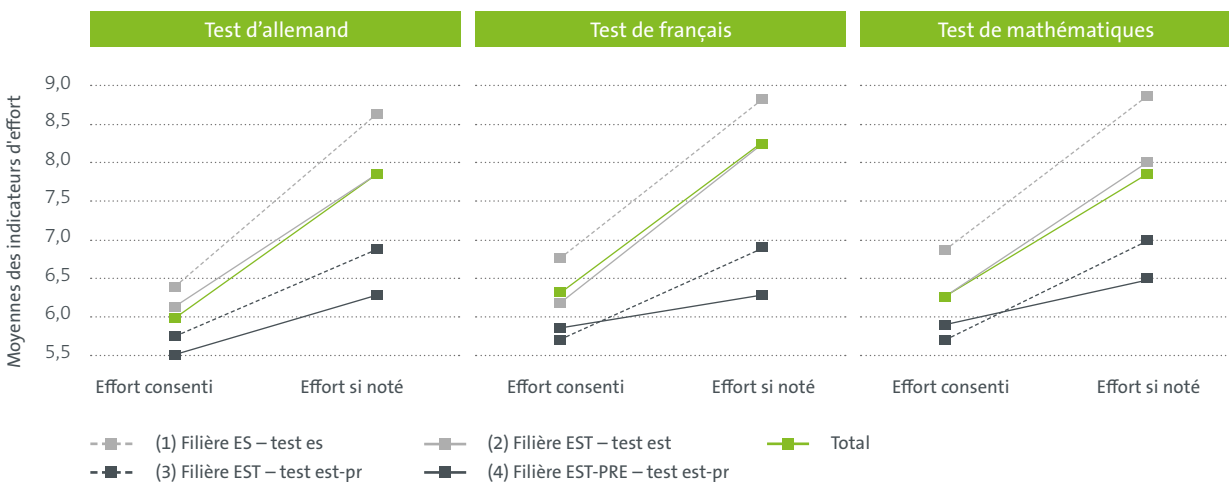
différences légèrement plus marquées entre les groupes considérés. Par exemple, l'écart entre la filière ES et la filière EST-PRE est d'environ un point quelle que soit la matière testée.

Groupes d'élèves		Test								
		allemand			français			mathématiques		
Filière	Version du test	N	Moyenne de l'effort si noté	Ecart-type	N	Moyenne de l'effort si noté	Ecart-type	N	Moyenne de l'effort si noté	Ecart-type
(1) ES	es	1623	8.6	1.8	1648	8.7	1.8	1368	8.9	1.7
(2) EST	est	2792	7.8	2.4	3042	7.8	2.4	2955	8.0	2.3
(3) EST	est-pr	633	6.9	2.7	658	6.9	2.6	652	7.0	2.6
(4) EST-PRE	est-pr	605	6.3	2.8	726	6.3	2.8	709	6.5	2.8
<b>Total</b>		<b>5653</b>	<b>7.8</b>	<b>2.4</b>	<b>6074</b>	<b>7.8</b>	<b>2.4</b>	<b>5684</b>	<b>7.9</b>	<b>2.4</b>

Tableau 3 : Moyenne de l'« effort si noté » selon les groupes d'élèves et les matières testées (sur une échelle de 1 à 10)

Au niveau de l'effort consenti si les ÉpStan avaient compté comme note scolaire (« effort si noté »), les valeurs moyennes calculées pour la population totale se situent toutes autour de 7.8 (voir tableau 3). On observe cette fois

des différences très marquées entre les groupes d'élèves considérés, l'amplitude de l'écart entre la filière ES et la filière EST-PRE tournant autour de 2.4 points.



Graphique 1 : Moyennes des indicateurs d'« effort consenti » et d'« effort si noté » selon les groupes d'élèves considérés et les matières testées

Le graphique 1 permet de comparer l'« effort consenti » et l'« effort si noté » et met en lumière une différence entre les deux indicateurs moyens d'effort puisque l'« effort si noté » est systématiquement supérieur à l'« effort consenti » pour le test. Cette différence est cependant plus ou moins marquée selon le groupe d'élèves considéré : elle

tourne autour de deux points pour la filière ES mais seulement autour d'un demi-point pour la filière EST-PRE.

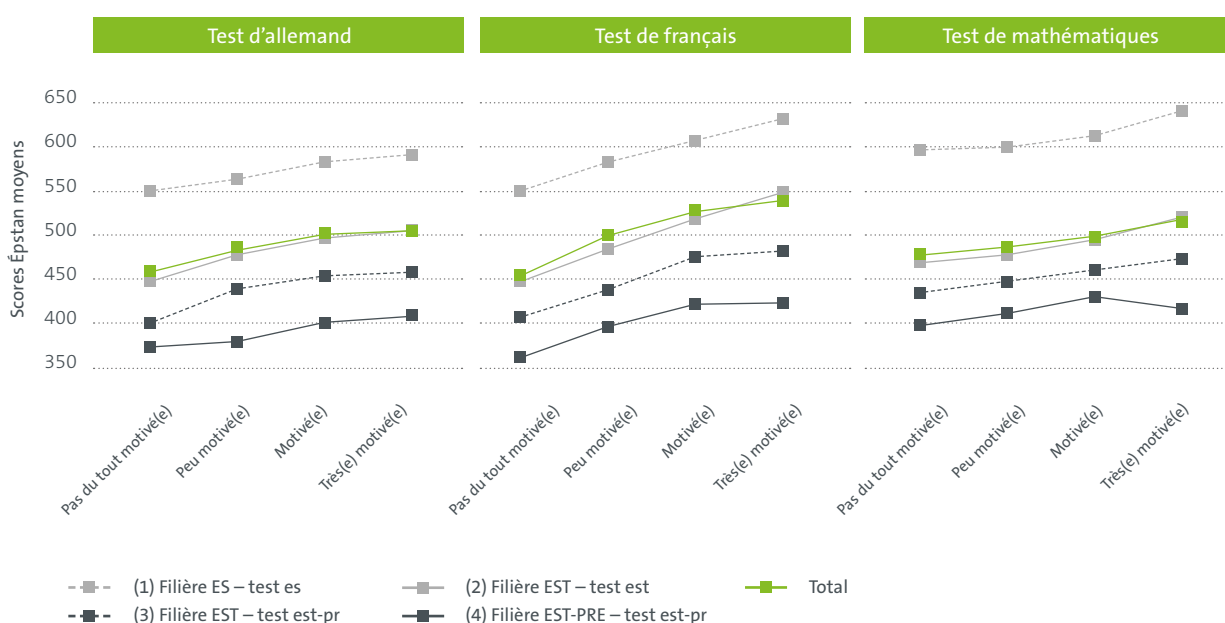
Les liens entre ces données auto-rapportées et les scores ÉpStan obtenus par les élèves sont détaillés dans le tableau 4 et les graphiques 2 et 3.

Motivation initiale	Test								
	allemand			français			mathématiques		
	N	Score moyen	Ecart-type	N	Score moyen	Ecart-type	N	Score moyen	Ecart-type
Pas du tout motivé(e)	396	465	104.1	645	454	105.8	584	476	84.4
Peu motivé(e)	1359	489	97.9	1685	496	105.3	1584	484	77.7
Motivé(e)	3123	510	95.6	2939	527	107.2	3175	500	84.0
Tres motivé(e)	947	508	99.8	936	536	122.3	926	508	95.3
<b>Total</b>	<b>5825</b>	<b>502</b>	<b>98.3</b>	<b>6205</b>	<b>512</b>	<b>111.7</b>	<b>6269</b>	<b>495</b>	<b>84.9</b>

Tableau 4 : Scores ÉpStan moyens pour la population totale selon les matières testées et la « motivation initiale »

Le tableau 4 met en évidence l'existence d'un lien positif entre la « motivation initiale » rapportée par les élèves à l'entame de chaque test et la performance observée : plus

les élèves se déclarent motivés en début de test, plus ils obtiennent, en moyenne, un score élevé aux ÉpStan.



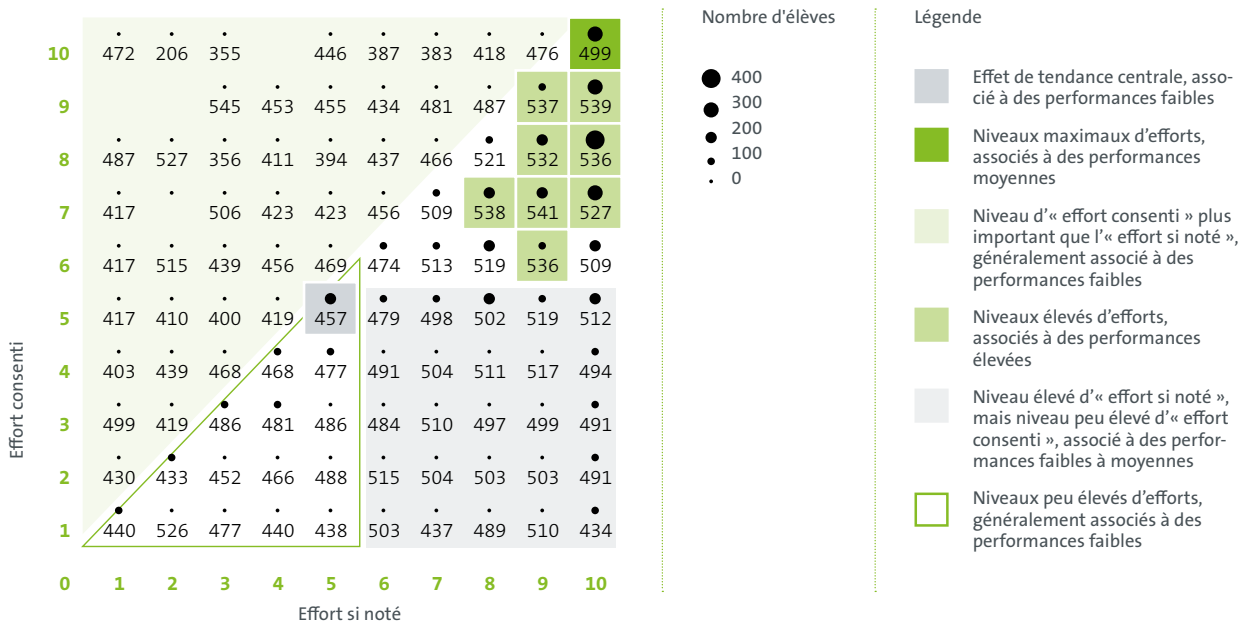
Graphique 2 : Scores ÉpStan moyens selon les groupes d'élèves, les matières testées et la « motivation initiale »

Le graphique 2 montre que les mêmes tendances existent pour les trois tests et pour les quatre groupes d'élèves considérés. Les différences de scores moyens entre les élèves très motivés et les élèves pas du tout motivés tournent, en moyenne, autour de 50 points mais semblent être les plus marquées pour le test de français.

Au niveau des variables d'effort, on peut synthétiser la situation par le graphique 3 qui permet de situer les élèves

en fonction de leurs réponses aux deux questions : « effort consenti » en ordonnée et « effort si noté » en abscisse. La taille des bulles représente le nombre d'élèves concernés. Le score moyen au test d'allemand obtenu par ces élèves est indiqué en-dessous de chaque bulle. Les scores moyens sont cependant à considérer avec précaution lorsque les effectifs sont faibles. Les constats dressés ici ne se réfèrent qu'au test d'allemand mais les constats sont similaires pour les deux autres tests (français et mathématiques).





Graphique 3 : Scores ÉpStan moyens en allemand selon les différentes modalités des variables d'« effort consenti » et d'« effort si noté »

On remarque tout d'abord que la grande majorité des élèves se situent dans la partie supérieure droite du graphique 3, là où les indicateurs d'effort sont les plus élevés. Les performances moyennes les plus élevées (en vert) sont généralement associées à des niveaux d'effort élevés, même si cette tendance ne se confirme pas dans le cas particulier des élèves qui ont répondu avoir consenti un effort maximal sur les deux échelles (en vert foncé). Leur performance moyenne est de 499. Une piste explicative pourrait être que l'on est ici face à des élèves qui présentent une tendance à l'exagération ou à la surestimation. On peut ensuite relever l'existence d'élèves ayant une façon particulière de répondre au « thermomètre d'effort », à savoir ceux qui déclarent avoir consacré davantage d'effort au test d'allemand ÉpStan qu'ils ne l'auraient fait si le test avait compté comme note scolaire (en vert clair dans le graphique). Ces élèves sont généralement crédités de performances faibles. On peut aussi remarquer l'existence d'un biais de tendance centrale de l'instrument utilisé puisqu'un nombre significatif d'élèves ont répondu au milieu des deux échelles (en gris foncé). Ces élèves sont

crédités de performances faibles. Enfin, le graphique permet d'identifier deux groupes<sup>5</sup> d'élèves qui sont particulièrement intéressants lorsqu'on traite de la problématique du sérieux des élèves lors des tests à faibles enjeux. D'un côté, on distingue un groupe d'élèves ayant déclaré des niveaux d'efforts peu élevés sur les deux échelles et présentant en général des performances faibles (triangle vert foncé). Ce sont des élèves qui sont probablement, de manière générale, non motivés pour l'école. De l'autre côté, on retrouve un groupe d'élèves ayant déclaré un niveau élevé d'« effort si noté », mais un niveau peu élevé d'« effort consenti » pour le test (carré grisé). Ce groupe obtient des performances faibles à moyennes aux ÉpStan. C'est ce dernier groupe d'élèves qui, en fonction de sa taille plus ou moins grande, pourrait éventuellement menacer la validité des constats dressés par les ÉpStan puisqu'il est envisageable que ces élèves auraient pu s'investir davantage et obtenir de meilleurs résultats. Il convient cependant de rester prudent car nous n'avons aucune possibilité de vérifier cette hypothèse sans connaître le niveau scolaire de ces élèves lors des devoirs en classe. ●

<sup>5</sup> Les seuils pour définir les deux groupes ont été placés au milieu des deux échelles à titre d'illustration, mais d'autres seuils auraient pu être envisagés, certes avec des conséquences différentes.

## CONSTATS ISSUS DES DONNÉES COMPORTEMENTALES (V<sup>e</sup>/9<sup>e</sup>)

L'analyse exploratoire des données comportementales à notre disposition, à savoir les temps d'affichage à l'écran de chaque stimulus, devrait permettre d'identifier, de manière plus objective, les élèves de V<sup>e</sup>/9<sup>e</sup> peu appliqués lors des ÉpStan. En totalisant les temps d'affichage de tous les stimuli, on dispose d'une variable rendant compte, théoriquement, du temps total consacré au traitement du test par les élèves.

En guise d'exemple, les constats dressés à propos du test d'allemand sont présentés dans le tableau 5. Les constats sont similaires pour les tests de français et de mathématiques. Agrégés au niveau du test, les temps d'affichage des stimuli indiquent que, sur les 50 minutes à disposition pour les différentes versions du test d'allemand, entre 25 et 35 minutes ont en moyenne été utilisées.

Groupes d'élèves		N	Temps total d'affichage moyen (en min)
Filière	Version du test		
(1) ES	es	1501	35
(2) EST	est	2591	31
(3) EST	est-pr	647	26
(4) EST-PRE	est-pr	560	25

Tableau 5 : Temps total d'affichage moyen des stimuli du test d'allemand selon les groupes d'élèves

On peut se demander dans quelle mesure le temps utilisé a un impact sur la performance au test. Il convient donc

d'examiner le lien qui existe entre cet indicateur et le score observé au test d'allemand en calculant les corrélations.

Groupes d'élèves		Corrélation
Filière	Version du test	
(1) ES	es	.11
(2) EST	est	.40
(3) EST	est-pr	.40
(4) EST-PRE	est-pr	.50
Total		.46

Tableau 6 : Corrélation entre le temps total d'affichage et le score ÉpStan obtenu en allemand selon les groupes d'élèves

La corrélation calculée entre le temps total d'affichage des stimuli et le score obtenu au test d'allemand est positive et significative (voir tableau 6), mais d'ampleur modérée (.46). Dans une certaine mesure, plus les élèves consacrent

du temps au test, plus ils obtiennent de bons résultats. Cette corrélation varie cependant selon le groupe d'élèves considéré : elle est modérée pour les groupes 2, 3 et 4, et faible pour le groupe 1. ●

### Le manque d'implication apparent de certains élèves lors des tests ÉpStan est-il un problème ?

Les ÉpStan visent *in fine* à calculer des indicateurs moyens de la compétence d'une cohorte et de sous-groupes d'élèves dans différentes disciplines scolaires à des fins de monitoring et de pilotage du système éducatif. Tout comme pour d'autres évaluations externes à faibles enjeux, le manque d'implication de certains élèves lors des épreuves peut s'avérer problématique au fur et à mesure qu'augmente la proportion d'une certaine catégorie d'élèves : ceux qui s'impliquent moins lors des épreuves à faibles enjeux que lors des épreuves à enjeux élevés. En effet, la présence de ces élèves sous-performants conduit à une sous-estimation de la compétence moyenne, cette sous-estimation étant plus ou moins prononcée selon le nombre d'élèves concernés par cette attitude particulière.

D'après les constats dressés dans la présente section au départ de données auto-rapportées et de données comportementales recueillies avant, pendant et après les tests, on reconnaîtra qu'il est effectivement possible d'observer ce genre de comportements dans le cadre des ÉpStan, mais il est impossible d'estimer, avec précision, le nombre d'élèves concernés par cette attitude, et donc l'ampleur du biais éventuel.

Les données ont par ailleurs permis de relever un autre groupe d'élèves : ceux qui ont déclaré des niveaux d'effort peu élevés aussi bien pour le test ÉpStan que pour des situations à enjeux élevés. Ce sont sans doute des élèves qui sont démotivés pour l'école en général. Cette problématique dépasse le cadre des ÉpStan, mais on peut se poser la question de savoir pourquoi il y a, dans le système éducatif, des élèves qui ne veulent plus performer et comment on peut y remédier.

## 3.2 FACTEURS QUI INFLUENCENT LES ATTITUDES DES ÉLÈVES ET DES ENSEIGNANTS À L'ÉGARD DES ÉPSTAN

### CONSTATS ISSUS DES DONNÉES DES FOCUS GROUPS

Les interlocuteurs mentionnent certains facteurs qui, selon eux, influencent les comportements et les attitudes des élèves et des enseignants à l'égard des ÉpStan. Un des plus importants est sans doute qu'il subsiste encore des questions par rapport à l'outil et que les informations quant au but de ces tests ne sont pas claires ou, en tout cas, pas clairement transmises. La plus grande revendication des interlocuteurs est donc peut-être d'avoir une explication sur les objectifs des ÉpStan.

Un autre facteur influant sur l'attitude des enseignants est leur implication lors des tests. Les instituteurs semblent être plus favorables aux ÉpStan que les professeurs qui se plaignent de ne pas en connaître assez du contenu des épreuves. D'après certains enseignants (du fondamental et du secondaire), le test ne correspond pas au programme scolaire. Cette non-correspondance est également citée pour expliquer le comportement des élèves lors des tests. Certains élèves, se sentant incapables de répondre aux questions, se démotiveraient rapidement et n'auraient

par conséquent plus envie de faire le test sérieusement. D'autres arguments, comme le fait que les ÉpStan ne comptent pas pour l'année, les effets de fatigue induits par la passation des trois tests en affilée ou encore la lecture moins concentrée sur PC expliqueraient les problèmes de motivation d'un certain nombre d'élèves.

À côté de tous ces éléments directement liés au dispositif ÉpStan, le milieu dans lequel les élèves et les enseignants évoluent jouerait également un rôle important. Ainsi, l'attitude et la manière de communiquer d'une direction seraient importantes pour les professeurs tout comme celles des professeurs le seraient pour leurs élèves. Par leur attitude, leur présence lors des tests ou leur manière de préparer les élèves en soulignant l'importance des tests, les professeurs pourraient influencer le comportement des élèves lors de la passation. La motivation devrait en quelque sorte se transmettre tout au long de la chaîne, en partant des directions pour arriver, via les enseignants, auprès des élèves. ●

#### Comment peut-on améliorer l'attitude de certains élèves du secondaire à l'égard des ÉpStan ?

Les interlocuteurs ont donné un certain nombre de pistes qui méritent d'être creusées davantage. La plus importante est sans doute d'expliquer clairement les enjeux des ÉpStan aux élèves, de leur faire comprendre que même si cela ne compte pas pour l'évaluation de l'année, leurs enseignants et leurs parents vont voir les résultats, que leurs

résultats auront une implication sur le résultat de leur lycée et que les ÉpStan leur donnent la possibilité de montrer ce dont ils sont capables.

Par ailleurs, il faudrait tenir compte de la « chaîne de la motivation », c'est-à-dire le fait que la motivation et l'information sont relayées d'une personne à l'autre, en essayant de préparer au maximum et le mieux possible tous les acteurs concernés.

← Encadré 10 : Attitude des élèves (du secondaire) à l'égard des ÉpStan



*Auteurs :*

*Marianne Milmeister  
Christophe Dierendonck  
Christiane Weis  
Paul Milmeister*

# 4 Dans quelle mesure les rapports envoyés aux acteurs dans le cadre des ÉpStan sont-ils compris et utilisés ?

27

Depuis 2008, les rapports ÉpStan mis à la disposition des différents acteurs suite aux épreuves ont connu des développements marqués tant au niveau de la forme que des contenus produits. Ainsi, la perspective initiale quasi-exclusivement psychométrique ne permettant pas d'identifier les élèves (jusqu'en 2009-2010) s'est transformée progressivement en une perspective mixte (depuis 2010-2011). Cette dernière recherche un équilibre entre comparaisons normatives et informations faisant davantage référence au curriculum scolaire. Elle autorise en outre une identification des élèves. Dans les lignes qui suivent, une série de réactions concernant plusieurs aspects des rapports sont synthétisées. ●

## 4.1 ASPECTS DES RAPPORTS QUI SUSCITENT LE PLUS D'INTÉRÊT AUPRÈS DES CONCERNÉS

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS* ET DU QUESTIONNAIRE

Certains interlocuteurs rencontrés dans les *focus groups* jugent que l'intérêt pour les rapports est assez limité. Pourtant, lors de l'enquête par questionnaire, les participants témoignent d'un certain intérêt pour les rapports : ils sont 95% des enseignants du cycle 3.1 et 88% des enseignants de V<sup>e</sup>/9<sup>e</sup> ayant participé à l'enquête à déclarer avoir lu le « rapport classe ». La plupart des enseignants (cycle 3.1 et V<sup>e</sup>/9<sup>e</sup>) trouvent intéressant de pouvoir situer les résultats de leurs élèves par rapport à la moyenne nationale et apprécient la comparaison qui tient compte du niveau socio-économique. De même, certains interlocuteurs rencontrés dans les *focus groups* considèrent l'intervalle de performance attendue (*Erwartungsbereich*) comme important pour mettre les résultats obtenus, tant les bons que les mauvais, dans un contexte. Pour la plupart des enseignants du cycle 3.1 et la moitié des enseignants en V<sup>e</sup>/9<sup>e</sup> enquêtés par questionnaire, le « rapport classe » apporte des informations que l'on peut mettre en relation avec les compétences définies dans les documents de référence. À peu près la moitié des enseignants (cycle 3.1 et V<sup>e</sup>/9<sup>e</sup>) estiment que les graphiques illustrant la distribution des résultats donnent une information intéressante et utile.

Les enseignants du cycle 3.1, enquêtés par questionnaire, sont davantage intéressés par l'interprétation critériée des résultats que par l'interprétation normative/comparative<sup>6</sup>

dans les « rapports élève ». En revanche, en V<sup>e</sup>/9<sup>e</sup>, la plupart des enseignants trouvent intéressantes les comparaisons par rapport à la moyenne de la classe et par rapport à la moyenne observée dans les différents types d'enseignement. Certains interlocuteurs des *focus groups* voient de manière positive la comparaison des différents scores avec des valeurs moyennes. Ces interlocuteurs estiment que ces comparaisons leur permettent de s'orienter, de se positionner face à une « norme ». En même temps, ils y voient aussi des inconvénients, surtout au niveau de l'élève : si cette comparaison est stimulante pour un bon élève, elle peut être démotivante pour un élève faible. La comparaison peut selon eux créer des situations de concurrence malsaine entre les élèves et entre les différentes filières. À peu près la moitié des enseignants enquêtés par questionnaire pensent que le « rapport élève » apporte des informations que l'on peut mettre en relation avec les compétences définies dans les documents de référence.

Les informations sur la motivation scolaire des élèves et le climat de classe dans les rapports sont généralement bien reçues par les enseignants interviewés lors des *focus groups* et suscitent leur intérêt. ●

## 4.2 AMÉLIORATIONS RELATIVES AUX RAPPORTS QUI POURRAIENT INTÉRESSER LES CONCERNÉS

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS* ET DU QUESTIONNAIRE

Des interlocuteurs rencontrés dans les *focus groups* – surtout les enseignants du secondaire – regrettent que le dispositif ÉpStan dans sa forme actuelle n'apporte pas davantage d'informations d'ordre diagnostique. Ces enseignants voudraient avoir plus de détails dans le « rapport élève » ou dans le « rapport classe » pour comprendre quelle est la cause d'un mauvais résultat : ils disent ne pas savoir si l'élève ne sait pas calculer, n'a pas appris les formules, n'a pas compris l'énoncé, n'avait plus envie de continuer le test... Ils trouvent que le rapport n'est pas assez concret pour pouvoir interpréter les résultats. De même, dans l'enquête par questionnaire, parmi les pistes d'amélioration proposées, celles visant à développer le caractère diagnostique des rapports issus des ÉpStan ont été le plus souvent plébiscitées (deux tiers d'accords ou plus).

De même, certains enseignants du secondaire, ayant participé aux *focus groups*, jugent l'utilité des ÉpStan limitée puisqu'il y a juste un « nombre » (score) par élève. Ils sont demandeurs, comme leurs collègues du fondamental, d'informations plus précises sur les faiblesses ou les lacunes des élèves. Certains professeurs disent n'avoir aucune idée où se

situent les faiblesses vu qu'ils n'ont pas accès aux questions du test ÉpStan. Les enseignants enquêtés par questionnaire semblent également largement en faveur d'une mise à disposition des questions utilisées dans les épreuves.

Dans le questionnaire, la suggestion « Les rapports devraient comprendre des pistes didactiques pouvant être exploitées par les enseignants afin de combler les lacunes des élèves » a été parmi les plus plébiscitées. En général, la proposition des pistes didactiques est plutôt bien accueillie par les enseignants ayant participé aux *focus groups*. Certains professeurs se disent ouverts à des idées novatrices et être prêts pour essayer de nouvelles choses dans leurs classes. Ces pistes didactiques pourraient être une aide pour l'enseignant pour travailler de façon ciblée avec les élèves, surtout si les classes sont trop grandes (dans les lycées). Néanmoins, selon des professeurs, elles seraient seulement utiles si elles étaient concrètes : par exemple des exercices ou des exemples de questions. Des instituteurs trouvent intéressante l'idée des pistes didactiques, mais à la seule condition que ces dernières soient des suggestions et non des prescriptions à suivre. ●

<sup>6</sup> Une mesure peut être interprétée de deux manières : par rapport aux autres individus testés (approche normative/comparative) ou par rapport à un critère prédéterminé ou un ensemble de compétences ou de tâches à réaliser (approche critériée).

### Est-il possible de fournir plus d'informations d'ordre diagnostique ou des pistes didactiques ?

Comme l'objectif des ÉpStan est en premier lieu le monitoring du système éducatif, l'investissement dans la mise à disposition de plus d'informations aux niveaux de la classe et de l'élève sera limité (cf. encadré 3). Les ÉpStan se focalisent sur les performances globales atteintes par les élèves par rapport aux différentes compétences telles qu'elles sont définies dans les plans d'études respectivement les programmes scolaires (cf. encadré 5). Les résultats permettent donc de situer la performance de manière générale et de voir dans quels domaines l'élève a des forces ou des faiblesses. Cette information peut donner suite à un diagnostic plus détaillé dans un ou plusieurs domaines.

Il serait sans doute intéressant et utile – d'autant plus que la demande semble exister – de développer l'aspect diagnostique pour les types d'évaluations spécifiquement/prioritairement conçues pour viser le niveau individuel comme p. ex. les Épreuves Communes (ÉpCom). Dans une optique de complémentarité entre les différentes évaluations, il serait tout à fait cohérent que des aspects favorisant les développements en classe soient davantage pris en compte par les ÉpCom.

### Pourquoi les rapports « élève » et « classe » se présentent-ils sous la forme actuelle ?

Bien que les ÉpStan aient été développées pour permettre une analyse au niveau du système éducatif, les résultats sont disponibles pour toute la population scolaire (cf. encadré 3) et ce serait dommage de ne pas partager ces informations avec les enseignants, les élèves et les parents sous forme de résultats individuels.

Le principe de la présentation des résultats selon la métrique ÉpStan (score standardisé avec une moyenne de 500 et un écart-type de 100), développée pour les besoins de monitoring au niveau du système, reste identique pour tous les rapports. Ceci permet une comparabilité maximale entre tous les rapports produits aux différents niveaux (national, établissement scolaire, classe et élève) et au fil des années. Rompre avec ce principe (en simplifiant par exemple les rapports avec des pourcentages de réussite) reviendrait à abandonner cette possibilité de comparaison maximale.

### Est-il envisageable de montrer aux enseignants du secondaire les questions des tests ?

On ne peut ignorer le constat de l'étude selon lequel les enseignants éprouvent beaucoup de difficultés à interpréter les résultats ÉpStan et à travailler avec ceux-ci lorsqu'ils ne connaissent pas le contenu testé. Le problème se pose surtout pour les enseignants du secondaire parce qu'ils ne corrigent pas – comme leurs collègues du fondamental – les copies, le test étant administré par ordinateur. La faisabilité logistique et technique d'accorder aux enseignants du secondaire l'accès aux questions du test par une consultation unique lors de la publication des résultats est étudiée pour le moment.

Un accès non-restreint aux items des tests n'est cependant pas possible car le développement des épreuves représente un grand investissement en ressources humaines et en temps. Ceci est d'autant plus vrai pour les items d'ancrage qui autorisent la comparabilité des résultats entre filières et la comparabilité des résultats entre cohortes d'élèves au fil des années. Une publication inconditionnelle des items empêcherait la réutilisation de ceux-ci dans des épreuves ultérieures.

## 4.3 COMPRÉHENSION DES RAPPORTS « ÉCOLE », « CLASSE » ET « ÉLÈVE » PAR LES ACTEURS CONCERNÉS

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS* ET DU QUESTIONNAIRE

Les directeurs et inspecteurs ayant participé aux *focus groups* disent, en général, comprendre les « rapports école » sans difficulté. Des inspecteurs émettent des doutes sur la compréhension des « rapports école » par les instituteurs, soit par manque de formation à cet égard, soit par manque d'intérêt.

De nombreux interlocuteurs doutent de la compréhension des « rapports élève » (en tout ou en partie) par les élèves et les parents, notamment chez les élèves fréquentant les classes de 9<sup>e</sup> polyvalente, pratique et préparatoire. Les élèves ayant participé aux *focus groups* disent comprendre le rapport s'ils se penchent dessus ou si l'enseignant le leur explique. De l'enquête par questionnaire, il ressort qu'il y a compréhension totale ou quasi-totale du « rapport élève » par la grande majorité des enseignants (cycle 3.1 et V<sup>e</sup>/9<sup>e</sup>). Par contre, toujours selon les enseignants, la compréhension du « rapport élève » ne serait totale ou quasi-totale que pour 4 élèves sur 10 du cycle 3.1 et pour une majorité (7 sur 10) de parents d'élèves du cycle 3.1, de parents d'élèves de V<sup>e</sup>/9<sup>e</sup> et d'élèves de V<sup>e</sup>/9<sup>e</sup>.

Les données du questionnaire montrent que sur le plan de l'intelligibilité du « rapport classe », la grande majorité des enseignants (cycle 3.1 et V<sup>e</sup>/9<sup>e</sup>) déclarent avoir compris

totalemment ou quasi-totalemment le document, le plus souvent après avoir lu les explications qui accompagnaient le rapport. Or, selon les propos recueillis dans les *focus groups*, les « rapports classe » et/ou « rapports élève » poseraient des problèmes de compréhension au moins à une partie des professeurs – plutôt aux professeurs de langues que de mathématiques. Ces enseignants disent ne pas comprendre la métrique ÉpStan et le principe de la moyenne de 500 points. Le graphique pose également problème à certains instituteurs : la représentation choisie consistant à mettre la moyenne nationale au-dessus de la moyenne de la classe les aurait induit à penser que leur classe était systématiquement en-dessous de la moyenne. En même temps, certains admettent avoir survolé seulement les documents et les textes explicatifs.

Par ailleurs, les données recueillies lors de l'enquête par questionnaire montrent que lorsqu'on tente de vérifier la compréhension effective des graphiques des rapports par les enseignants à l'aide de questions plus poussées, on s'aperçoit que ceux-ci peuvent interpréter les graphiques et le score ÉpStan dans des directions très contrastées et parfois erronées. ●

#### Pourquoi les rapports sont-ils présentés sous la forme actuelle ?

À partir des constats de l'étude, on peut retenir qu'il y a des différences dans la compréhension des rapports selon les groupes concernés (élèves, parents, enseignants, directeurs, inspecteurs, etc.), selon leur intérêt et selon la consultation des explications fournies (lecture des brochures d'information, informations relayées par l'enseignant à l'élève, etc.). Il ne semble donc pas indiqué de changer fondamentalement les rapports, mais plutôt de cibler les différents publics pour promouvoir la compréhension de ces derniers. D'autant plus que des modifica-

tions sur l'un ou l'autre rapport, en optant p. ex. pour une simplification de la présentation, entraîneraient la perte de la comparabilité maximale de tous les rapports telle qu'elle existe à présent (cf. encadré 11).

Il faudra, dans les prochaines années, de manière systématique et dans le dialogue avec les concernés, chercher à s'assurer d'une meilleure compréhension des rapports, en passant p. ex. par les formations initiales et continues des enseignants et en développant plus généralement les connaissances des enseignants dans les domaines de l'évaluation et de l'*assessment*.



## 4.4 UTILISATION DES RAPPORTS

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS* ET DU QUESTIONNAIRE

Les interlocuteurs ayant participé aux *focus groups* ont décrit des exemples d'utilisation des rapports ÉpStan qui peuvent être regroupés selon différents types et niveaux d'utilisation. En guise d'exemples, les **types d'utilisation des rapports** évoqués par les interlocuteurs sont entre autres des remises en question personnelles, des échanges entre concernés ou encore des utilisations jugées délicates par certains.

Certains interlocuteurs ont abordé la remise en question personnelle suite aux résultats ÉpStan : les élèves montrent une certaine facilité à se remettre en question et certains professeurs pensent qu'ils pourraient bien reconsidérer leurs routines. Dans le groupe des instituteurs, on estime que ce processus de remise en question n'est pas automatique. Selon eux, tirer des conclusions constructives des résultats ÉpStan est une question d'attitude. Sinon, il suffit simplement de conclure que ce sont les élèves qui sont faibles ou difficiles.

Il ressort des données du questionnaire qu'une minorité d'enseignants (cycle 3.1 et V<sup>e</sup>/9<sup>e</sup>) déclarent que le rapport leur a permis de prendre conscience que certaines facettes de leur enseignement devaient être modifiées. Seule une minorité d'enseignants (cycle 3.1 et V<sup>e</sup>/9<sup>e</sup>) disent comprendre, suite au « rapport classe », comment améliorer certaines choses dans leur enseignement.

Les données recueillies dans le cadre de l'enquête par questionnaire montrent que les « rapports classe » ont fait l'objet d'échanges entre collègues dans la majorité des cas. Les données issues des *focus groups* donnent une image plus mitigée. Certains interlocuteurs rapportent qu'il n'y a, selon eux, que peu d'échanges sur les rapports dans leur établissement scolaire. D'autres évoquent cependant des exemples d'échanges : pour le fondamental, des interlocuteurs rapportent des échanges entre inspecteurs et

écoles/comités d'écoles, entre enseignants du cycle 3.1, entre enseignants des cycles 3.1 et 2.2, entre enseignants du cycle 3, entre enseignants et parents lors des entretiens de bilan. Pour le secondaire, des échanges concernant les rapports ont eu lieu dans les directions, dans les Cellules de développement scolaire (CDS), dans les départements, entre enseignants et élèves, entre enseignants ainsi qu'entre élèves.

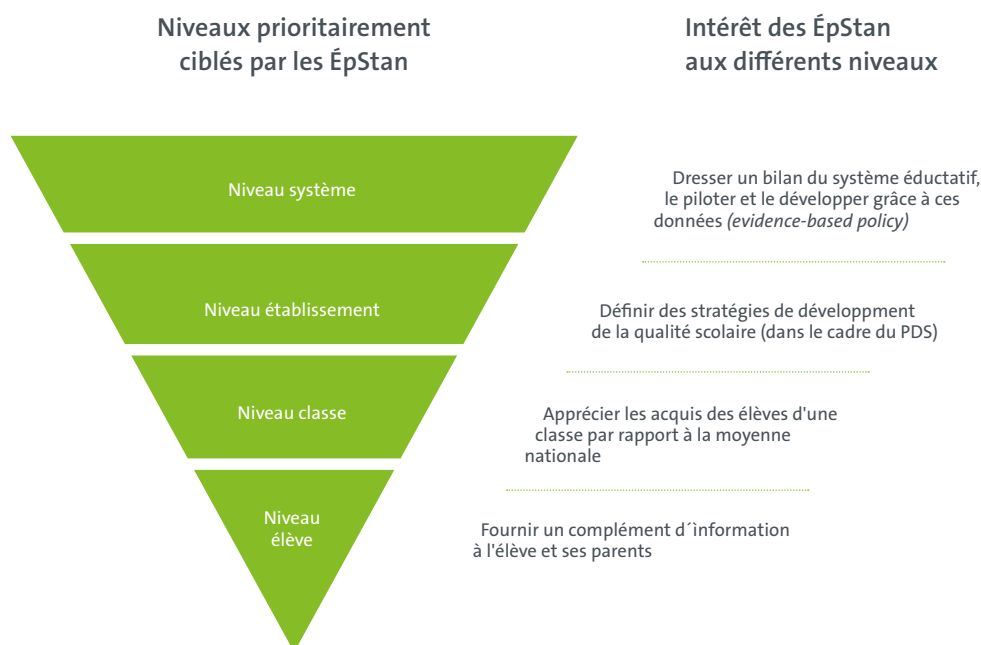
Certaines utilisations sont jugées délicates comme par exemple au fondamental, la difficulté de faire respecter l'anonymat des « rapports classe » et des « rapports élève » dans les petites écoles. Au secondaire, des enseignants décrivent par exemple des cas où les directions de lycée ont convoqué les enseignants concernés pour demander des explications sur les résultats des tests. Ce genre d'utilisation a été mal vécu par le personnel enseignant.

Les exemples d'utilisations décrits par les participants des *focus groups* peuvent être regroupés aussi selon différents **niveaux d'utilisation** (p. ex. « utilisation légère » vs « utilisation plus poussée »). Les utilisations caractérisées comme « légères » relèvent plutôt de l'ordre d'un intérêt pour les rapports, comme par exemple dans le cas de directions de lycées qui sont confrontées à des résultats contre attente, à savoir que leurs classes EST s'en sont tirées mieux que leurs classes ES. Les directions concluent qu'il faut réfléchir aux causes de cette sorte de résultats. De même, certains instituteurs estiment que la comparaison de la moyenne de leur classe avec la moyenne nationale est intéressante pour situer leurs élèves. Les exemples d'utilisations plus « poussées » se caractérisent par une volonté des concernés d'aller au-delà d'un échange, d'une réflexion ou d'un certain intérêt. Dans l'enseignement fondamental, par exemple, l'utilisation des données ÉpStan dans le cadre du Plan de réussite scolaire (PRS) et des comités d'école est évoquée par plusieurs interlocuteurs. ●

## Quelles sont les utilisations potentielles des rapports aux différents niveaux (niveau système, établissement scolaire, classe et élève) ?

Suite à l'évolution du dispositif des ÉpStan au fil du temps, des rapports ont progressivement été mis à la disposition

des concernés de manière à ce que différentes utilisations soient désormais possibles. Bien que les utilisations aux niveaux de l'établissement scolaire, de la classe ou de l'élève restent secondaires par rapport au but principal du monitoring (cf. encadré 3), les informations mises à disposition à ces niveaux peuvent être utiles et intéressantes



Au **niveau du système**, les informations recueillies visent à dresser un rapport fiable sur le système éducatif et à contribuer au pilotage et au développement du système au départ de données (*evidence-based policy*). Il s'agit de vérifier, d'un côté, si les socles de compétences ont été atteints et de l'autre côté, de voir si les socles ont été fixés à un niveau approprié (cf. encadré 2).

Au **niveau de l'établissement scolaire**, les données ÉpStan peuvent servir à alimenter le processus de développement de la qualité scolaire. La comparaison de l'école ou du lycée avec les scores moyens nationaux (*Landesmittelwert* pour le fondamental et *Mittelwerte der Schulformen* pour le secondaire) permet de situer l'école par rapport à une population de référence et de mieux juger les performances des élèves. L'intervalle de performance attendue (*Erwartungsbereich*), qui est renseigné au niveau de l'établissement scolaire, permet une comparaison équitable : les performances des élèves sont relativisées par la prise en compte des caractéristiques socio-économiques des publics scolaires testés (cf. encadrés 3 et 7). Ces informations pourront contribuer

à identifier des forces et faiblesses de leur population scolaire et à définir des stratégies de développement scolaire, notamment dans le cadre des Plans de développement de l'établissement scolaire (PDS).

Au **niveau de la classe**, l'enseignant reçoit des informations complémentaires permettant de vérifier l'atteinte des niveaux de compétence par ses élèves et de juger des acquis de ses élèves en comparaison à la moyenne nationale. Ici encore une comparaison équitable est possible grâce à l'intervalle de performance attendue indiqué pour la classe. La comparaison des résultats ÉpStan avec sa propre évaluation lui permet éventuellement de réfléchir à son système d'évaluation.

Au **niveau de l'élève**, le résultat ÉpStan constitue un complément d'information à l'évaluation de l'enseignant pour l'élève et ses parents.

## 4.5 IMPACT DES RÉSULTATS ÉPSTAN

### CONSTATS ISSUS DES DONNÉES DES FOCUS GROUPS ET DU QUESTIONNAIRE

Selon certains interlocuteurs ayant participé aux *focus groups*, les résultats du « rapport élève » pourraient avoir une influence indirecte sur la motivation des élèves face à l'école en général. Il ressort des discussions avec les élèves qu'ils sont contents et fiers s'ils ont de bons résultats ÉpStan. Ils estiment que les ÉpStan montrent éventuellement qu'ils savent faire plus que ce qu'ils croyaient et que cela leur donne plus de motivation pour faire des efforts par la suite. En revanche, si les résultats sont moins bons, cela peut constituer un signal d'alarme pour certains élèves qui se disent qu'il faudrait travailler plus. D'après les enseignants, l'impact des ÉpStan sur les élèves varie en fonction des résultats : de bons résultats feraient que les élèves se sentent valorisés alors que de mauvais résultats entraîneraient frustration et tristesse. En effet, les élèves de 9<sup>e</sup> se sentiraient dévalorisés face à de mauvais résultats comparatifs, ce qui pourrait les rendre inquiets et démotivés par rapport à l'école en général, d'autant plus s'ils sont orientés à la fin de l'année scolaire.

De même, lors de l'enquête par questionnaire, la grande majorité des enseignants enquêtés (cycle 3.1 et V<sup>e</sup>/9<sup>e</sup>) sont d'accord pour dire qu'un rapport positif peut renforcer l'estime de soi et la motivation scolaire de certains élèves. À l'inverse, à peu près la moitié des enseignants, au fondamental comme au secondaire, affirment être d'accord avec l'idée qu'un rapport individuel négatif peut diminuer l'estime de soi et la motivation scolaire de certains élèves et que les « rapports élève » sont susceptibles de renforcer la stigmatisation de certains élèves faibles. Une petite ma-

ajorité d'enseignants (cycle 3.1 et V<sup>e</sup>/9<sup>e</sup>) soutiennent l'idée qu'un rapport individualisé négatif pourrait amener l'élève à une prise de conscience.

À la question de savoir s'ils ont modifié leurs pratiques d'enseignement suite au « rapport classe », une minorité d'enseignants ayant participé au questionnaire, au fondamental comme au secondaire, ont répondu par l'affirmative. À peu près un tiers des enseignants au cycle 3.1 et un cinquième des enseignants en V<sup>e</sup>/9<sup>e</sup> déclarent que les rapports issus des ÉpStan ont eu de réelles retombées positives sur le travail quotidien de leur(s) classe(s). Malheureusement les enseignants n'avaient pas la possibilité de préciser de quel type de retombées il s'agissait.

Certains interlocuteurs ayant participé aux *focus groups*, en particulier ceux de l'enseignement secondaire, jugent que les ÉpStan n'ont pas énormément d'impact au niveau de l'établissement scolaire. Des coordinateurs pensent même que l'impact des ÉpStan est inexistant, qu'il n'y a pas vraiment de retombées – par exemple sur les projets d'établissement – ou qu'il n'y a pas d'impulsion partant des ÉpStan pour initier des changements importants. Cette idée est partagée par certains professeurs qui s'attendaient par exemple à des retours dans les commissions de programme, qui n'ont cependant pas eu lieu à leur connaissance. Ils doutent que les ÉpStan puissent déclencher une dynamique de changement dans une école et jugent qu'il n'y a pas d'impact sur la qualité scolaire. ●

#### Quel est l'impact potentiel des ÉpStan aux niveaux de la classe et de l'établissement scolaire ?

Au niveau de la classe, les résultats ÉpStan constituent une source d'informations intéressantes pour l'enseignant car ils permettent de situer le niveau de ses élèves par rapport aux autres (perspective comparative) et par rapport au programme scolaire (perspective critériée). Les résultats ÉpStan sont susceptibles d'amener l'enseignant à une remise en question de ses pratiques d'enseignement et/ou d'évaluation.

Les résultats ÉpStan sont aussi utilisés par la division « du traitement de données sur la qualité de l'encadrement et de l'offre scolaire et éducative » du SCRIPT dans le cadre de

la documentation des établissements scolaires sur leurs situations respectives. Cette documentation est rassemblée dans un « classeur école » ou dans un « classeur lycée » structuré en quatre parties (données démographiques, performances scolaires, processus scolaires, perceptions des acteurs). Ces classeurs sont censés être utilisés, en complément d'autres outils, pour définir des stratégies de développement de la qualité scolaire dans chaque établissement, notamment au travers des Plans de développement de l'établissement scolaire (PDS). Dans ce sens, les ÉpStan contribuent au niveau de l'établissement scolaire à une meilleure compréhension des situations scolaires particulières et aux processus de développement de la qualité scolaire qui sont conduits sur le terrain.

← Encadré 14 : Impact potentiel des ÉpStan au niveau des classes et des établissements scolaires

## 4.6 DIFFICULTÉS LIÉES À L'ACCÈS AUX DIFFÉRENTS RAPPORTS

### 4

### CONSTATS ISSUS DES DONNÉES DES FOCUS GROUPS

Lors des discussions dans les *focus groups*, certains interlocuteurs ont abordé les difficultés d'accès et de distribution qu'ils ont rencontrées concernant les rapports ou les résultats. Le problème de distribution concerne surtout les « rapports élève », car au fondamental, des enseignants disent ne pas savoir s'il faut les montrer ou les donner aux parents. Au secondaire, certains enseignants admettent ne pas avoir remis les rapports aux élèves, soit parce qu'ils ne savaient pas qu'ils devaient le faire, soit parce qu'ils ne voulaient pas le faire. Les raisons avancées pour la non-remise des rapports sont la stigmatisation des élèves faibles ou l'anticipation d'une situation où ils ne pourraient pas répondre aux questions des élèves.

De même, les questions d'accès aux résultats ont été discutées : dans certains lycées, les directions ne donnent pas accès au « rapport école » par crainte d'actes de malveillance de la part des professeurs (c'est-à-dire la communication des résultats en dehors de l'école). Certains coordi-

nateurs pensent que les directions devraient avoir accès aux « rapports classe » et/ou « rapports élève », tandis que les directions reconnaissent la raison d'être de cette restriction. Au fondamental, des pressions pour accéder au « rapport école » sont exercées par deux types d'acteurs : d'un côté, les parents – et plus spécifiquement les représentants de parents – qui sont censés contribuer au PRS et, de l'autre côté, les communes ou les commissions scolaires qui souhaiteraient en savoir davantage sur la performance de leur(s) établissement(s) scolaire(s). Des représentants de syndicats attendent à cet égard un positionnement de la part du Ministère qui devrait dire clairement que les « rapports école » servent uniquement en interne au développement de la qualité scolaire et ne doivent pas être publiés. Les inspecteurs estiment que l'accès aux « rapports élève » et/ou « rapports classe » serait intéressant pour eux, mais quelques-uns jugent pourtant légitime qu'ils n'ont pas accès à ces rapports. ●

#### Pourquoi l'accès à la plupart des rapports ÉpStan est-il restreint ?

À l'heure actuelle, il y a des restrictions d'accès aux différents rapports, excepté pour le rapport national qui est public. Ainsi, le « rapport école » est adressé au directeur d'école, à l'inspecteur, au président du comité d'école. La division « du traitement de données sur la qualité de l'encadrement et de l'offre scolaire et éducative » du SCRIPT y a également accès. Le « rapport classe » peut uniquement être téléchargé par l'enseignant, c'est-à-dire par l'instituteur ou le professeur de langue (allemand, français) ou de mathématiques concerné par les ÉpStan. Il en est de même pour le « rapport élève » qui peut uniquement être téléchargé par l'enseignant, puis distribué à chaque élève.

Ces modalités d'accès aux rapports ont été fixées délibérément pour permettre un travail serein et constructif à chaque niveau (établissement scolaire, classe et élève) et pour « protéger » les concernés d'un emploi délicat, voire

abusif des résultats (cf. encadré 17).

Dans une logique *bottom up*, rien n'empêche par contre, dans des écoles où la coopération se fait dans la confiance, de décider de commun accord entre enseignants et responsables d'école de partager les données respectives pour en tirer plus d'informations.

#### Comment la distribution des « rapports élève » devrait-elle se faire ?

La distribution du « rapport élève » est censée se faire par l'enseignant qui remet le rapport à l'élève, après l'avoir téléchargé via [feedback.epstan.lu](https://feedback.epstan.lu). Cette distribution est obligatoire depuis l'année scolaire 2016/17. Le retour du résultat à l'élève et à ses parents peut bien sûr constituer une incitation pour l'élève à faire l'épreuve avec le sérieux nécessaire (cf. encadré 9). Idéalement, lors de l'entretien avec les parents une discussion autour du résultat ÉpStan devrait avoir lieu.

## 4.7 QUESTIONS ET REVENDEICATIONS DES CONCERNÉS À PROPOS DES RÉSULTATS

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS*

Plusieurs questions et demandes adressées aux évaluateurs reviennent de façon récurrente. Certains demandent comment travailler pratiquement avec les rapports ÉpStan sur le terrain. Les concernés veulent également comprendre quelles conclusions sont tirées à partir des ÉpStan au niveau national et ils s'attendent à ce que ces conclusions soient communiquées. Des directives plus précises sont attendues à plusieurs niveaux. Certains inspecteurs aimeraient en avoir quant à la distribution et l'utilisation des « rapports élève », quant à leur responsabilité lors du travail avec les résultats face aux comités d'école

et quant à la communication des résultats de l'école à la commission scolaire. Des instituteurs voudraient des recommandations concernant l'échange avec les collègues. Des coordinateurs et des directeurs insistent sur l'importance de donner un suivi aux ÉpStan, c'est-à-dire un accompagnement pour la compréhension et l'utilisation des résultats : ceci augmenterait l'acceptation du dispositif et la motivation des responsables d'écoles à travailler avec les résultats ÉpStan. Selon eux, sans suivi, les ÉpStan sont perçues comme une fin en soi. ●

#### Quelles conclusions sont tirées au niveau national à partir des résultats ÉpStan ?

Les résultats ÉpStan sont publiés dans des rapports nationaux et dans le « Bildungsbericht Luxemburg ». Dans ces rapports, les scores moyens atteints dans les différentes matières testées sont présentés en fonction de plusieurs variables : statut socio-économique et migratoire des familles, genre des élèves et langues parlées à la maison. Pour chaque analyse, les résultats sont comparés à ceux des années précédentes. Ces rapports donnent une vue d'ensemble de l'évolution du système éducatif et permettent d'évaluer par exemple l'impact des réformes entreprises.

Sur le plan de la recherche, les résultats permettent d'ores et déjà de tirer des conclusions et d'orienter de futures études dans le domaine de l'éducation. Dans le dernier rapport national (Martin, Ugen & Fischbach, 2015), les chercheurs constatent que le système scolaire luxembourgeois doit surmonter des défis de taille qui sont notamment liés aux trois facteurs suivants : (1) un environnement de vie et d'apprentissage qui a fondamentalement changé avec l'émergence de la société de la connaissance où les nouvelles technologies et les services jouent un rôle important ; (2) une population scolaire particulièrement hétérogène d'un point de vue social, culturel et linguistique ; (3) une situation linguistique spécifique au Luxembourg et par conséquent un curriculum extrêmement complexe, notamment à cause de l'apprentissage des langues..

#### Est-il possible d'investir davantage dans le suivi des ÉpStan au niveau des établissements scolaires ?

Un accompagnement des écoles se fait déjà actuellement par les instituteurs spécialisés en développement scolaire du SCRIPT. La mission des instituteurs spécialisés est d'assister les écoles dans leur développement scolaire et de soutenir les enseignants qui demandent une assistance personnalisée dans leur travail pédagogique.

Le rôle de l'Université se situe dans le domaine de la recherche et son soutien ne pourra être que scientifique. Ainsi pourrait-elle assurer l'explication de données scientifiques, par exemple par l'intermédiaire de formations continues ou lors d'interventions dans des écoles, mais également mettre en place, en cas de besoins ou de demandes du terrain, des recherches collaboratives pour suivre et favoriser le processus du développement scolaire.

## 4.8 RISQUES ET APPRÉHENSIONS DES CONCERNÉS CONCERNANT LES RAPPORTS

### CONSTATS ISSUS DES DONNÉES DES *FOCUS GROUPS*

Certaines appréhensions et risques liés aux résultats des ÉpStan ont parfois été décrits, tels que des pressions ou des attentes face aux enseignants en provenance des directions ou des parents, le risque de concurrence (entre enseignants, entre filières, entre lycées) avec comme corollaire la peur du ranking et du tourisme scolaire, ou encore

le contrôle du travail de l'enseignant (par la direction/l'inspecteur, par le Ministère). Les interlocuteurs y apportent néanmoins une nuance importante en disant que ces situations sont souvent largement dépendantes de l'atmosphère de travail et de l'attitude des collègues. ●

#### Les appréhensions des concernés face aux ÉpStan sont-elles justifiées ?

Certaines de ces appréhensions sont sûrement justifiées, mais une partie d'entre elles sont liées à un accès aux données qui n'est pas prévu par le dispositif, comme par exemple l'accès d'une direction aux rapports « classe » de son école. Il serait souhaitable que les inspecteurs et les directeurs d'école se donnent une stratégie commune pour faire face à cette sorte de situations en rapport avec l'échange de données à l'intérieur de l'école afin de déjouer les appréhensions des enseignants (cf. encadré 15).

En ce qui concerne les appréhensions comme la concurrence entre les écoles, le ranking, le tourisme scolaire etc., elles présupposent un partage de données vers l'extérieur de l'école. Comme ce dernier n'est pas prévu dans le dispositif actuel, on peut conclure que ces craintes ne sont pas vraiment justifiées.

← Encadré 17 : Appréhensions face aux résultats des ÉpStan

## LISTE DES RÉFÉRENCES

**Bamberg, M., Dickes, P., & Schaber, G. (1977).** *Étude Magrip : Premier rapport de synthèse.* Walferdange : Institut pédagogique.

**Baumert, J., & Demmrich, A. (2001).** Test motivation in the assessment of student skills : The effects of incentives on motivation and performance. *European Journal of Psychology of Education*, 16, 441-462.

**Brunner, M., & Martin, R. (Dir.) (2011).** *Die MAGRIP-Studie (1968-2009). Wie beeinflussen sozio-kognitive Merkmale von Kindern im Grundschulalter und ihre Bildungswege ihr späteres Leben als Erwachsene in Luxemburg ?* Luxembourg : Université Luxembourg, Forschungseinheit EMACS. En ligne [http://magrip.lu/docs/unilu\\_MAGRIP\\_\\_web.pdf](http://magrip.lu/docs/unilu_MAGRIP__web.pdf)

**Dierendonck, C., Kafai, A., Fischbach, A., Martin, R., & Ugen, S. (2015).** Les Épreuves Standardisées. Élément-clé du pilotage du système éducatif luxembourgeois. *Education et Formations*, 86-87, 61-81. En ligne [http://cache.media.education.gouv.fr/file/revue\\_86-87/59/1/depp-2015-EF-86-87-epreuves-standardisees-element-cle-pilotage-systeme\\_424591.pdf](http://cache.media.education.gouv.fr/file/revue_86-87/59/1/depp-2015-EF-86-87-epreuves-standardisees-element-cle-pilotage-systeme_424591.pdf)

**Dierendonck, C., Milmeister, M., Milmeister, P., Weis, C., Reichert, M., Ugen, S., Keller, U., Sonnleitner, P., & Martin, R. (2015).** Mise en place d'une démarche qualité pour le dispositif ÉpStan. In R. Martin, S. Ugen, & A. Fischbach (Dir.), *Épreuves Standardisées : Bildungsmonitoring für Luxemburg. Nationaler Bericht 2011 bis 2013* (pp. 99-109). Esch/Alzette : University of Luxembourg, LUCET. En ligne <http://hdl.handle.net/10993/21046>

**Dierendonck, C., Sonnleitner, P., Ugen, S., Keller, U., Fischbach, A., & Martin, R. (2013).** La mesure de la motivation et de l'effort des élèves dans le cadre des Épreuves Standardisées au Luxembourg. *Actes du congrès AREF 2013.*

**Fischbach, A., Ugen, S., & Martin, R. (Dir.) (2014).** *Épstan Technical Report.* Luxembourg : Ministère de l'Éducation nationale et de la Formation professionnelle, Université du Luxembourg/ECCS/LUCET. En ligne <http://hdl.handle.net/10993/15802>

**Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., ..., Vollmer, H. J. (2007).** *Zur Entwicklung nationaler Bildungsstandards. Expertise. Bundesministerium für Bildung und Forschung (BMBF).* Bonn, Berlin : Bundesministerium für Bildung und Forschung (BMBF). En ligne [https://www.bmbf.de/pub/Bildungsforschung\\_Band\\_1.pdf](https://www.bmbf.de/pub/Bildungsforschung_Band_1.pdf)

**Loi du 6 février 2009 portant organisation de l'enseignement fondamental, MEMORIAL A – N°187 du 3 septembre 2009.**  
En ligne <http://www.legilux.public.lu/leg/a/archives/2009/0020/a020.pdf#page=4>

**Loi du 6 février 2009 portant restructuration du SCRIPT, MEMORIAL A – N°19 du 16 février 2009.**  
En ligne <http://www.legilux.public.lu/leg/a/archives/2009/0019/a019.pdf#page=2>

**Martin, R., Ugen, S., & Fischbach, A. (Dir.) (2015).** *Épreuves Standardisées : Bildungsmonitoring für Luxemburg. Nationaler Bericht 2011 bis 2013.* Esch/Alzette : University of Luxembourg, LUCET. En ligne <http://hdl.handle.net/10993/21046>

**Ministère de l'Éducation nationale et de la Formation professionnelle (2007).** *Die Steuerung des Luxemburger Schulwesens.* Courrier de l'Éducation nationale, numéro spécial. Luxembourg : Ministère de l'Éducation nationale et de la Formation professionnelle, SCRIPT. En ligne <http://www.men.public.lu/catalogue-publications/themes-transversaux/cen/cens/steuerung-luxemburger-schulwesens/de.pdf>

**Ministère de l'Éducation nationale, de l'Enfance et de la Jeunesse & Université du Luxembourg (Dir.).** *Bildungsbericht Luxemburg 2015. Band II : Analysen und Befunde.* Luxembourg : Ministère de l'Éducation nationale, de l'Enfance et de la Jeunesse, Université du Luxembourg. En ligne <http://www.men.public.lu/catalogue-publications/themes-transversaux/statistiques-analyses/bildungsbericht/2015/band-2.pdf>

**Wolf, L., & Smith, J. (1995).** The consequence of consequence : Motivation, anxiety and test performance. *Applied Measurement in Education*, 8, 227-242.

### SITES INTERNET :

**Site Internet de LUCET :** [http://www.en.uni.lu/recherche/flshase/luxembourg\\_centre\\_for\\_educational\\_testing\\_lucet/](http://www.en.uni.lu/recherche/flshase/luxembourg_centre_for_educational_testing_lucet/)

**Site Internet des ÉpStan :** <https://www.epstan.lu>

## Liste des abréviations

ADQS	Agence pour le développement de la qualité scolaire
CDS	Cellule de développement scolaire
ÉpCom	Épreuves Communes
ÉpStan	Épreuves Standardisées
ES	Enseignement Secondaire
EST	Enseignement Secondaire Technique
EST-PRE	Enseignement Secondaire Technique – Régime Préparatoire
FAPEL	Fédération des associations de parents d'élèves du Luxembourg
LUCET	Luxembourg Centre for Educational Testing
MAGRIP	Matière grise perdue (grande enquête psycho-socio-éducationnelle de 1968 destinée à comprendre les mécanismes du passage primaire-secondaire, répétée en 2009)
MENJE	Ministère de l'Éducation nationale, de l'Enfance et de la Jeunesse
OCDE	Organisation de coopération et de développement économiques
PDS (ancien format)	Plan de développement scolaire (pour le secondaire)
PDS (nouveau format)	Plan de développement de l'établissement scolaire (pour le fondamental et le secondaire)
PISA	Programme for International Student Assessment
PRS	Plan de réussite scolaire (pour le fondamental)
QCM	Questionnaire à choix multiple
SCRIPT	Service de Coordination de la Recherche et de l'Innovation pédagogiques et technologiques





[www.epstan.lu](http://www.epstan.lu)



[www.epstan.lu](http://www.epstan.lu)

LES ÉPREUVES STANDARDISÉES :  
COMMENT SONT-ELLES PERÇUES PAR LES ACTEURS CONCERNÉS ?  
PRINCIPAUX CONSTATS ET CLARIFICATIONS



UNIVERSITY OF LUXEMBOURG  
Luxembourg Centre for  
Educational Testing (LUCET)