



## Évaluation des compétences des élèves genevois

État des lieux, enjeux et perspectives

*Carl Denecker*

*Marion Dutrévis*

*Franck Petrucci*

*Oliver Prosperi*

# Sommaire

<b>Introduction.....</b>	<b>3</b>
<b>L'évaluation des compétences dans le canton de Genève.....</b>	<b>4</b>
Du point de vue de l'élève : quelles évaluations pour quelles conséquences ? .....	4
Le parcours évaluatif « classique » au primaire .....	4
Le parcours évaluatif « classique » au CO .....	7
Du point de vue du système : quelles évaluations pour quel monitoring ? .....	9
Au niveau cantonal : les TAF.....	9
Au niveau national : l'enquête COFO .....	9
Au niveau international : l'enquête PISA .....	12
Que nous disent ces outils de monitoring des compétences des élèves genevois ?.....	15
<b>Les facteurs qui influencent les compétences des élèves .....</b>	<b>16</b>
Les compétences des élèves selon les caractéristiques individuelles .....	16
Les effets du contexte sur les compétences des élèves .....	18
<b>Les coulisses de l'évaluation .....</b>	<b>23</b>
Du point de vue des disciplines enseignées : des évaluations de quelles compétences ? .....	24
Évaluation des compétences des élèves : regard critique.....	24
Le jugement sur la valeur scolaire des élèves ou les biais de la notation .....	25
Vers une évaluation pour l'apprentissage ? .....	28
<b>Perspectives.....</b>	<b>30</b>
Les limites des différentes formes d'évaluation et de leur usage .....	30
L'école obligatoire versus l'enseignement secondaire II .....	30
Les « privés » d'évaluation .....	31
L'évaluation comme enjeu de communication.....	31
<b>Bibliographie.....</b>	<b>34</b>
Travaux du SRED... ..	34
... et d'ailleurs .....	35

## Introduction

Comme il est stipulé dans l'article 10 de la Loi sur l'instruction publique (LIP, C 1 10), l'école publique a pour mission fondamentale de donner aux élèves le moyen d'acquérir les meilleures connaissances et compétences en vue de leurs activités futures. En plus d'être chargée de développer les compétences des élèves, l'école doit également attester de l'atteinte de celles-ci, en recourant à des évaluations (LIP, art. 62 et 73). En contexte scolaire, compétences et évaluation sont étroitement liées. Dès lors, comment évaluer l'acquisition des compétences par les élèves ? Et quelles sont les compétences à évaluer, à quel(s) moment(s) et dans quels buts ? Telles sont les questions qui vont structurer ce texte.

Depuis de nombreuses années, le SRED réalise des études visant à rendre compte des compétences acquises par les élèves genevois, et ce du début de la scolarité jusqu'à l'obtention des premières certifications. Ces travaux portent principalement sur la scolarité obligatoire. Ce focus sur l'enseignement obligatoire, s'il peut être en partie un choix institutionnel et scientifique, est aussi une question de disponibilités des informations. En effet, si à chaque niveau du cursus scolaire les élèves genevois sont soumis à des évaluations, leur nombre est particulièrement important durant la scolarité obligatoire. Au primaire et au cycle d'orientation s'ajoutent aux évaluations menées par les enseignantes et les enseignants dans leur classe l'évaluation commune (en 4P, 8P et 11<sup>e</sup>) et les tests d'attentes fondamentales (TAF) en 11<sup>e</sup>. En 11<sup>e</sup> toujours, certaines et certains élèves se voient tirés au sort tous les trois ans (quatre ans à partir de 2025) pour participer aux enquêtes internationales PISA. Enfin, les enquêtes nationales COFO, qui ont démarré en 2016, interrogent également des échantillons d'élèves de 4P, 8P ou 11<sup>e</sup>. Vient ensuite le temps de l'orientation vers le secondaire II où les contenus évaluatifs sont régis par des ordonnances fédérales pour chaque filière de formation. Dans ce degré d'enseignement, le SRED fournit des chiffres et analyses centrées principalement sur les taux de certification et de réussite aux examens des élèves.

*Il s'agit à la fois de prendre conscience des limites, voire des biais d'évaluation (...) et de penser un modèle combinatoire d'évaluations qui limite les biais et contribue à construire un regard le plus objectif possible sur les compétences des élèves.*

Les évaluations constituent donc une ressource largement mobilisée pour monitorer l'école genevoise. Elles donnent bien sûr à voir le niveau de compétences de la population scolaire, mais également les inégalités de réussite entre élèves ou groupes d'élèves, les écarts entre types d'établissements, ou encore les compétences disciplinaires évaluées (et donc considérées comme des apprentissages essentiels par l'institution). Pour autant, les évaluations comportent certaines limites qui font qu'une évaluation donnée ne reflète pas forcément les compétences réelles des élèves. Il s'agit donc à la fois de prendre conscience des limites, voire des biais d'évaluation, que ceux-ci concernent le travail enseignant au quotidien ou le système dans son ensemble, et de penser un modèle combinatoire d'évaluations qui limite les biais et contribue à construire un regard le plus objectif possible sur les compétences des élèves.

Si l'évaluation est largement pensée et mobilisée comme un outil pour rendre compte des apprentissages, les systèmes scolaires tout comme la recherche mettent de plus en plus en avant l'évaluation pour apprendre. Cette posture soulève de nouvelles questions et incite à penser un élargissement des travaux en la matière. Comment, par exemple, mobiliser les résultats des diverses évaluations proposées aux élèves pour qu'elles participent réellement au processus d'enseignement-apprentissage ? Dans certains cas, comment comprendre et justifier l'absence d'évaluation (tout du moins rendue visible) des élèves les plus en difficultés ? En effet, si l'évaluation est au service de l'apprentissage, ne serait-ce pas priver ces élèves d'opportunités d'apprentissage que de les priver d'évaluations ? Des premiers travaux menés aux SRED commencent à s'emparer de cette question.

Comme indiqué plus haut, du point de vue du système, les évaluations nous aident jusqu'ici principalement à monitorer l'école genevoise, essentiellement en s'assurant du niveau de compétences des élèves par rapport à un plan d'études donné, de même qu'en comparaison nationale ou internationale. Mais la réflexion sur notre système scolaire peut être élargie.

Quelles que soient les évaluations utilisées, ce sont presque toujours les mêmes disciplines qui sont évaluées. Ce constat dit beaucoup des compétences attendues de la part de l'école (et de la société), mais entre en contradiction avec les discours qui valorisent les autres compétences, qu'elles soient disciplinaires ou transversales. De même, de nouvelles compétences intègrent les plans de formation, comme les compétences numériques, mais leur évaluation reste encore à être pensée. Là aussi, des premiers travaux ont eu lieu mais devront sans doute s'inscrire dans la durée. L'évaluation des compétences devrait également être un enjeu crucial pour la formation initiale et continue des enseignantes et des enseignants : que ces évaluations soient internes aux classes ou aux établissements, qu'elles soient cantonales ou intercantionales, il s'agit de connaître leurs fonctions, leurs règles de construction et l'utilité, voire l'impact qu'elles peuvent avoir pour le parcours scolaire des élèves et le système dans son ensemble.

## L'évaluation des compétences dans le canton de Genève

Dès son entrée à l'école obligatoire à l'âge de 4 ans révolus au 31 juillet de l'année en cours, chaque enfant débute son parcours scolaire, qui peut être interprété comme un « parcours évaluatif » dans la mesure où il est jalonné par différents moments évaluatifs. Ceux-ci englobent l'évaluation que pratiquent les enseignantes et enseignants au quotidien dans la classe, des travaux d'évaluation plus formalisés qui peuvent éventuellement façonner la suite de la trajectoire scolaire de l'élève, notamment au moment des transitions, jusqu'à une possible participation à des enquêtes qui servent le monitoring du système éducatif.

### Du point de vue de l'élève : quelles évaluations pour quelles conséquences ?

Lorsqu'un enfant atteint l'âge de scolarité obligatoire, elle ou il est inscrit à l'école obligatoire et entre dans le degré primaire. Les détails de l'évaluation scolaire des élèves du degré primaire sont fixés dans les articles 39 à 55 du Règlement genevois de l'enseignement primaire (REP, C 1 10.21). La directive sur l'évaluation des compétences et des connaissances des élèves (EP) (2022) complète et précise les dispositions dudit règlement. En effet, depuis la rentrée scolaire 2022, le projet d'innovation pédagogique EP21 mené à l'échelle cantonale propose des changements importants au niveau de l'organisation de l'année scolaire, de la communication aux parents et de l'évaluation du Français, notamment au cycle moyen. Ce projet, déployé sur deux ans, a été évalué à l'issue de la première année (2022-23) par le SRED (Dutrévis, Wettstein & Spagnolo, 2024). Le cas échéant, la LIP et le REP seront adaptés pour une implémentation pérenne dès la rentrée 2024.

### *Le parcours évaluatif « classique » au primaire*

Le degré primaire dure huit ans et comprend deux cycles d'une durée de quatre ans chacun, à savoir le cycle élémentaire et le cycle moyen. Les apprentissages de l'élève dans les disciplines et la vie scolaire y sont évalués régulièrement en référence au plan d'études romand (PER) et à ses spécificités cantonales. Cette évaluation peut prendre diverses formes,

notamment certificative, informative ou formative<sup>1</sup>. Suivant le règlement de l'enseignement primaire, l'évaluation est initialement trimestrielle et communiquée aux parents trois fois par année au moyen du bulletin scolaire. Avec le projet EP21, l'année scolaire est désormais découpée en deux semestres, au profit d'une plus grande cohérence entre pratiques d'enseignement et évaluation.

Au cycle élémentaire (1P à 4P), l'évaluation des apprentissages se veut essentiellement formative et porte sur la progression de l'élève en regard des objectifs du PER. La directive sur l'évaluation des compétences et des connaissances des élèves (EP) laisse le nombre et la forme des activités d'évaluation à la libre appréciation des équipes enseignantes. La production de l'oral, l'observation critériée et le recueil de traces sont privilégiés pour objectiver l'évaluation des apprentissages. Durant les deux premières années de scolarité, les titulaires de classe évaluent chaque semestre la progression dans l'acquisition des connaissances et des compétences requises et communiquent celle-ci aux parents de l'élève<sup>2</sup>. En fin de 2P, la progression de l'élève dans les domaines disciplinaires fait l'objet d'appréciations en termes de « très satisfaisante », « satisfaisante » et « peu satisfaisante ». En 1P et 2P, il n'y a en principe pas de redoublement.

Dès la 3P, le passage d'une année à l'autre n'est plus automatique et le nombre de travaux d'évaluation écrits augmente progressivement. Dès le premier semestre, la progression en Français et en Mathématiques fait l'objet d'une évaluation selon les appréciations citées plus haut ; au deuxième semestre de la 3P, d'autres disciplines<sup>3</sup> s'y ajoutent. À la fin de la 3P, les élèves ayant une appréciation de progression « (très) satisfaisante » en Français et en Mathématiques sont admis en 4P. Les élèves ayant une appréciation de progression « peu satisfaisante » dans l'une des deux ou les deux disciplines sont admis, avec des mesures d'accompagnement. Un redoublement peut toutefois être envisagé sur la base d'un bilan

*La grande majorité des élèves atteint le seuil de réussite aux EC de fin de 4P, les Mathématiques étant typiquement mieux réussies que le Français.*

pédagogique, d'une évaluation pédagogique complémentaire et de l'avis des parents.

Enfin, au terme de la dernière année du cycle élémentaire, les titulaires de classe établissent, pour chaque discipline, un bilan certificatif annuel en référence aux attentes de fin de 4P définies dans le PER. L'évaluation individuelle des acquis s'effectue également à travers des épreuves cantonales (EC), organisées par le département à la fin de la 4P. La grande majorité des élèves atteint le seuil de réussite aux EC de fin de 4P, les Mathématiques étant typiquement mieux réussies que le Français. Ainsi, le taux de réussite pour l'EC de Français était de plus de 80% en 2018 et 2019 et de plus de 70% en 2022, alors que près de 90% des élèves atteignaient le seuil de réussite aux épreuves en Mathématiques (Denecker, Petrucci & Prospero, 2023a). Des différences de réussite apparaissent toutefois en fonction de certaines caractéristiques individuelles et contextuelles (cf. plus loin : *Les facteurs qui influencent les compétences des élèves*).

Le bilan certificatif inclut à la fois les évaluations faites par les enseignantes et enseignants titulaires des classes et les résultats des élèves aux évaluations communes de fin de cycle. Les élèves ayant « atteint avec aisance » ou « atteint » le niveau de connaissances et de compétences requis en Français et en Mathématiques sont admis en 5P. Les élèves ayant « presque atteint » le niveau de connaissances et de compétences requis en Français et/ou en Mathématiques sont promus par tolérance, avec des mesures d'accompagnement. Une décision de redoublement en 4P peut être prononcée pour les élèves n'ayant « pas atteint » le niveau requis dans l'une des deux ou les deux disciplines ; la direction peut également décider leur admission par dérogation en 5P, avec des mesures d'accompagnement.

Au cycle moyen (5P à 8P), en articulation avec l'évaluation formative, l'évaluation du niveau d'acquisition des connaissances et des compétences scolaires<sup>4</sup> est réalisée sur la base de

<sup>1</sup> Les différentes formes que peut prendre l'évaluation des apprentissages scolaires sont sommairement présentées plus loin (cf. *Évaluation des compétences des élèves : regard critique*).

<sup>2</sup> Avec l'introduction d'EP21, ce n'est plus le cas.

<sup>3</sup> Sciences de la nature, Sciences humaines et sociales, Éducation physique, Activités créatrices et manuelles – Arts visuels et Musique.

<sup>4</sup> En tant que notions pédagogiques, *connaissances* et *compétences* scolaires ont inspiré une littérature abondante et un débat permanent, tant elles sont difficiles à définir et discerner de manière entièrement satisfaisante. Par souci de commodité, nous retenons ici seulement la définition de Perrenoud (1999), qui caractérise la compétence comme « une capacité d'action efficace face à une famille de situations, qu'on arrive à maîtriser parce qu'on dispose à la fois des connaissances nécessaires et de la

deux types d'activités d'évaluation certificative : les travaux significatifs (TS) et, uniquement en Français et en Mathématiques, les travaux assimilés (TA)<sup>5</sup>. Notons que ces derniers peuvent aussi avoir une visée formative.

Selon la discipline et l'année de scolarité, les travaux significatifs font l'objet d'une note ou d'une appréciation. Les disciplines dont l'évaluation est notée sont : Français I, Français II et Mathématiques. Toutefois, Genève étant le seul canton romand à distinguer Français I (communication) de Français II (fonctionnement de la langue), le projet EP21 introduit désormais le passage de deux disciplines à une seule. Certaines disciplines font l'objet d'appréciations au début du cycle moyen et ne sont notées que plus tard : Allemand (dès la 6P), Anglais et Sciences de la nature (dès la 7P) ainsi que Sciences humaines et sociales (dès la 8P). Les disciplines non notées concernent l'Éducation physique, les Activités créatrices et manuelles – Arts visuels, ainsi que la Musique.

Notons que le projet EP21 prévoit un nombre minimal d'évaluations qui varie selon la discipline et l'année de scolarité. Ce nombre a baissé étant donné le passage à une note unique pour le Français. Le découpage de l'année scolaire en deux semestres de la 1P à la 8P entend renforcer la qualité des évaluations, aussi bien formatives que certificatives, en cohérence avec les pratiques d'enseignement. Par exemple, au moins trois travaux significatifs (TS) par semestre sont attendus en Français et en Mathématiques tout au long du cycle moyen, alors qu'un seul TS par semestre est attendu en Musique et en Éducation physique. De même, en Allemand et en Sciences de la nature, au moins un TS par semestre est attendu au début du cycle moyen, mais on passe à respectivement trois et deux TS par semestre dès qu'elles deviennent des disciplines notées.

*capacité de les mobiliser à bon escient, en temps opportun, pour identifier et résoudre le problème » (p. 79).*

<sup>5</sup> Les *travaux significatifs* (TS) reposent principalement sur des tâches dont la complexité exige la mobilisation en situation de connaissances et de compétences acquises par les élèves. Les *travaux assimilés* (TA), de moindre ampleur, permettent de vérifier l'acquisition de connaissances déclaratives ou procédurales devant être automatisées pour réduire la charge cognitive des élèves dans des situations complexes (Directive sur l'évaluation des compétences et connaissances des élèves (EP), 2022).

Les travaux des élèves dans les disciplines non notées font l'objet d'appréciations en termes de « objectifs de l'activité atteints » et « objectifs de l'activité non atteints, à reprendre ». À la fin de l'année scolaire, les titulaires indiquent dans le bilan certificatif annuel si le niveau de connaissances et de compétences défini pour la discipline et pour l'année de scolarité est « atteint avec aisance », « atteint », « presque atteint » ou « pas atteint ». Pour les disciplines notées, les résultats aux TS s'inscrivent sur une échelle allant de 6 (maximum) à 1,5 (minimum). La note 4 correspond au seuil de suffisance. La note 1 est réservée à un travail rendu sans réponse. Pour les évaluations notées, une moyenne est calculée à la fin de chaque trimestre (EP21 : semestre). La moyenne annuelle, obtenue par la moyenne des trois trimestres (EP21 : deux semestres) et arrondie au dixième, est renseignée dans le bilan certificatif annuel.

*Le découpage de l'année scolaire en deux semestres de la 1P à la 8P entend renforcer la qualité des évaluations, aussi bien formatives que certificatives, en cohérence avec les pratiques d'enseignement.*

À la fin de la 5P, de la 6P et de la 7P, les élèves ayant obtenu au moins 4,0 de moyenne annuelle en Français et en Mathématiques sont promus à l'année de scolarité sui-

vante. Les élèves n'ayant pas obtenu 4, mais au moins 3,0 de moyenne annuelle en Français et/ou en Mathématiques, sont promus par tolérance. Lorsqu'une ou un élève a obtenu une moyenne annuelle inférieure à 3, la direction d'établissement décide, en accord avec les enseignants et après consultation des parents, de son redoublement<sup>6</sup> ou de son admission par dérogation respectivement en 6P, 7P ou 8P. Aussi bien la promotion par tolérance que l'admission par dérogation sont obligatoirement assorties de mesures d'accompagnement.

Au terme de la 8P, dernière année du cycle moyen, les titulaires de classe établissent un bilan certificatif annuel qui indique la moyenne annuelle de l'élève dans chacune des disciplines notées, les appréciations dans les autres disciplines, et les résultats aux épreuves cantonales (EC), qui comptent pour 20% de la moyenne du deuxième semestre. La part des élèves qui atteignent le seuil de réus-

<sup>6</sup> À noter que le redoublement, comme décision issue d'un processus évaluatif, a fait l'objet de très nombreuses recherches depuis des décennies (Goos, Pipa & Peixoto, 2021 ; Jackson, 1975 ; Jimmerson, 2001). Ces analyses tendent à conclure à une absence d'effet positif du redoublement. Les données genevoises montrent que le redoublement est associé à des parcours scolaires complexes (cf. les *indicateurs du décrochage scolaire* et ceux de *pédagogie spécialisée*).



site aux EC est variable selon l'année et la discipline. En 2019, elle était plus élevée en Français I (90%) et Français II (83%) qu'en Mathématiques (76%). En 2022, on observe une tendance inverse. Le taux de réussite en Mathématiques s'élevait à 82%, alors que les élèves ont dans l'ensemble moins bien réussi les épreuves de Français (76% et 60%). Comme pour les épreuves cantonales de fin de 4P, des différences de réussite apparaissent en fonction de certaines caractéristiques individuelles et contextuelles sur lesquelles nous reviendrons plus loin (Denecker et al., 2023a).

Au terme de cette dernière année du degré primaire, l'évaluation du niveau d'acquisition des connaissances et des compétences dans les disciplines de passage conditionne l'admission et l'orientation au secondaire I. Ce passage est défini dans les articles 53 et 54 du Règlement de l'enseignement primaire (REP) et dans les articles 29 et 30 du Règlement du cycle d'orientation (RCO, C 1 10.26). Le projet EP21 déroge pour l'heure à certaines dispositions du REP concernant l'orientation au CO.

### Le parcours évaluatif « classique » au CO

La première année du CO (9<sup>e</sup> année) est organisée en trois regroupements d'élèves, appelés regroupement 1 (R1), regroupement 2 (R2) et regroupement 3 (R3). Les mêmes disciplines y sont enseignées à toutes et tous les élèves (pour faciliter les réorientations au cours et à l'issue de la 9<sup>e</sup> année), mais les niveaux d'attente ainsi que les dotations horaires varient d'un regroupement à l'autre.

Les élèves sont admis et dirigés vers l'un de ces trois regroupements sur la base de leurs moyennes annuelles en Mathématiques et en Français en fin de 8P<sup>7</sup>. Les normes d'accès aux trois regroupements pour les élèves promus sont les suivantes :

- pour accéder au R1, un total minimal de 9,0 est exigé, avec chaque note de passage égale ou supérieure à 3,0 ;
- pour accéder au R2, un total minimal de 11,5 est exigé, avec chaque note de passage égale ou supérieure à 3,5 ;
- pour accéder au R3, un total minimal de 14,0 est exigé, avec chaque note de passage égale ou supérieure à 4,0.

<sup>7</sup> Cette année, avec la mise en œuvre du projet d'innovation pédagogique EP21, le Français est enseigné et évalué comme une discipline unique. Dans cette période transitoire, la décision de promotion des élèves s'appuie sur un doublement de la note de Français.

En 2019 et 2022, environ 70% des élèves avaient des moyennes annuelles leur permettant d'aller en R3, autour de 20% sont admis en R2 et environ 6% en R1 (Denecker, Petrucci & Proserpi, 2023b).

En Suisse, chaque canton fait des choix différents sur la façon de regrouper les élèves et les normes d'accès au secondaire I ne sont pas les mêmes partout (CSRE, 2023). Le canton de Vaud, par exemple, ne propose que deux types de regroupements au secondaire I. L'accès des élèves à la voie pré-gymnasiale (VP) ou générale (VG) y est principalement conditionné par leurs résultats dans un groupe restreint de disciplines composé du Français, des Mathématiques et de l'Allemand. Un total minimal de 15 pour ces trois disciplines est exigé pour accéder à la voie dont les exigences scolaires sont les plus élevées et les résultats aux épreuves cantonales de référence (ECR) sont pris en compte dans le calcul de la moyenne annuelle à hauteur de 30%. En juin 2021, 51% des élèves vaudois de 8P ont été orientés en VG et 46% en VP. À première vue, l'accès au regroupement le plus exigeant paraît alors plus ouvert à Genève (70% vs 51%). Toutefois, dans le même temps, cet accès plus ouvert peut se traduire par des implications moins heureuses quant à l'image et la composition des regroupements aux exigences plus faibles, avec le risque de créer des environnements d'apprentissage moins favorables, voire de la ségrégation scolaire (Felouzis, Liot & Perroton, 2005).

À Genève, les 10<sup>e</sup> et 11<sup>e</sup> années sont organisées en trois sections : communication et technologie (CT), langues vivantes et communication (LC) avec profils Allemand/Anglais (AA) ainsi que Sciences appliquées (SA), et littéraire-scientifique (LS) avec profil latin, langues vivantes ou sciences. Les élèves y ont accès en fonction de leurs choix d'orientation et des résultats obtenus à la fin de leur 9<sup>e</sup> ou de leur 10<sup>e</sup> année. Alors que l'intitulé des sections laisse penser à une logique disciplinaire, celle-ci reflète en réalité davantage des niveaux d'exigences scolaires plus ou moins élevés. L'enseignement dispensé dans les 10<sup>e</sup> et 11<sup>e</sup> années se répartit entre des disciplines communes à toutes les sections et des disciplines qui caractérisent la section ou le profil. Ces disciplines peuvent être spécifiques ou constituer un approfondissement de disciplines communes. Durant les trois années du CO, les disciplines principales, reconnues comme fondamentales dans l'apprentissage des élèves et dont les notes interviennent spécifiquement dans les conditions de promotion, sont le

Français, les Mathématiques et l'Allemand pour l'ensemble des élèves. En 10<sup>e</sup> et en 11<sup>e</sup> années, une quatrième note – qui varie en fonction des sections – figure dans les disciplines principales. Les niveaux d'attente de chaque discipline sont déterminés pour chaque regroupement et section, sur la base des objectifs d'apprentissage, des progressions et des attentes fondamentales énoncés dans le PER romand et ses spécificités cantonales.

Les détails de l'évaluation et du parcours scolaire des élèves du secondaire I sont fixés dans les articles 35 à 58 du Règlement du cycle d'orientation (RCO). L'évaluation au CO peut prendre diverses formes, notamment certificative, formative ou diagnostique. L'évaluation certificative est réalisée notamment au travers des travaux exécutés en classe, des interrogations écrites ou orales, des travaux personnels ou de groupe. Les notes attribuées aux travaux s'inscrivent sur une échelle de 1 à 6<sup>8</sup>. Le seuil de suffisance est fixé à 4,0.

Toutes les disciplines enseignées font l'objet d'une évaluation chiffrée, à l'exception de certaines disciplines<sup>9</sup> du domaine de la formation générale. À la fin de chaque trimestre, une moyenne des notes est établie pour chaque discipline sur la base de l'ensemble des divers travaux effectués. Les moyennes annuelles sont établies sur la base des moyennes trimestrielles.

Pour être promu, une moyenne annuelle de 4,0 est requise pour chaque discipline. Des tolérances sont toutefois admises. Les élèves promus des regroupements 1, 2 et 3 de 9<sup>e</sup> année sont admis respectivement dans les sections CT, LC et LS. Les élèves promus de 10<sup>e</sup> sont admis dans la même section. En fonction des résultats de l'élève, une réorientation peut intervenir au passage d'une année à l'autre (et, plus exceptionnellement, en cours d'année). Une ou un élève peut être admis directement l'année suivante dans une section aux attentes immédiatement plus élevées (orientation promotionnelle directe), si, en fin de 9<sup>e</sup> ou de 10<sup>e</sup>, elle ou il est promu avec une

moyenne générale d'au moins 5,0 et avec au plus une seule moyenne annuelle insuffisante (à l'exclusion du Français et des Mathématiques). Une ou un élève promu peut également demander à redoubler dans un regroupement ou une section aux niveaux d'attente immédiatement plus élevés (redoublement promotionnel) à condition qu'elle ou il n'ait pas déjà redoublé une année au CO. À l'inverse, un changement vers un regroupement ou une section aux attentes immédiatement moins élevées peut avoir lieu en cas de non-promotion et en l'absence d'un redoublement. L'élève non promu peut toutefois demander à redoubler son année dans le même regroupement ou dans la même section.

En 2021, 66% des élèves de 11<sup>e</sup> se trouvent en LS, toutes options confondues, 22% des élèves sont dans la section LC et enfin, la section CT ainsi que les classes-ateliers comptent 12% des élèves de la volée (Rastoldo & Mouad, 2022). Près de sept élèves sur dix ont connu un parcours linéaire en trois ans avec un niveau d'exigences identique en 11<sup>e</sup> année à celui de leur regroupement à l'entrée au CO en 2018. Les réorientations (promotionnelles, sélectives ou neutres) intervenues durant le cursus au CO concernent environ 20% de cette volée d'élèves, en premier lieu pour des parcours sélectifs (12%) et promotionnels (5%). Dans un cas sur trois, les élèves ne parviennent pas à se maintenir dans le nouveau regroupement ou la nouvelle section à exigences scolaires supérieures (parcours neutres). Le redoublement concerne 4% de la volée d'élèves entrée au CO en 2018 et vise le plus souvent au maintien dans le regroupement ou la section.

À 15 ans et à la fin de sa dernière année au CO, chaque élève qui remplit les conditions de promotion de sa section peut accéder directement à une filière certifiante de l'enseignement secondaire II correspondant à sa section<sup>10</sup>. L'élève peut également accéder directement à une filière de formation plus exigeante de l'enseignement secondaire II si elle ou il est promu avec une moyenne générale d'au moins 5,0, a au plus une seule moyenne annuelle insuffisante (à l'exclusion du Français et des Mathématiques) et répond aux conditions fixées par les règlements des écoles du degré se-

*Les réorientations (promotionnelles, sélectives ou neutres) intervenues durant le cursus au CO concernent environ 20% de cette volée d'élèves, en premier lieu pour des parcours sélectifs (12%) et promotionnels (5%).*

<sup>8</sup> La note 0 est réservée aux cas de (tentative de) fraude ou de plagiat.

<sup>9</sup> L'évaluation notée concerne uniquement les domaines disciplinaires. La Formation générale et les Capacités transversales ne donnent pas lieu à des notes (pour une présentation d'ensemble du PER, voir [plandetudes.ch](http://plandetudes.ch)).

<sup>10</sup> Les élèves qui ne remplissent pas les conditions peuvent passer par des classes de transition ou préparatoires.



conadaire II établis en accord avec la direction générale. Pour certaines filières professionnelles, la passation de concours d'entrée spécifiques est demandée en sus de la promotion ordinaire.

Les orientations au secondaire II sont donc fortement structurées par le profil scolaire (section fréquentée, notes et statut de promotion) des élèves. Ainsi, en 2020, 82% des élèves promus de la section LS se sont dirigés vers la filière gymnasiale, bien qu'elles et ils aient accès à toutes les formations de niveau secondaire II. À l'inverse, environ six élèves sur dix non promus de LC et plus de huit élèves sur dix de section CT (promus ou non) n'ont pas d'autre choix que d'entrer dans une formation préqualifiante qui leur permettrait de rejoindre ultérieurement une filière certifiante. Enfin, quel que soit le statut de promotion, l'orientation vers l'apprentissage dual est limitée et varie de 2% (LS bien promus) à 9% (LC promus) des élèves (Rastoldo & Mouad, 2021).

Cette première section a montré les différentes pratiques évaluatives qui structurent le parcours scolaire des élèves de l'enseignement obligatoire. Mais l'évaluation peut servir d'autres fonctions.

### Du point de vue du système : quelles évaluations pour quel monitoring ?

Des évaluations élaborées par les enseignantes et enseignants dans la classe pour évaluer individuellement les élèves, passons maintenant aux évaluations standardisées à large échelle, conçues par des organismes externes à la classe, pour obtenir des informations sur les performances de l'ensemble du système éducatif ou d'une partie clairement définie de celui-ci. Plutôt que d'effectuer une distinction entre les élèves à partir de leurs performances individuelles ou de certifier l'acquisition individuelle de connaissances, les évaluations standardisées à large échelle visent à décrire dans quelle mesure les élèves d'un système scolaire dans son ensemble ont acquis des savoirs et savoir-faire visés.

Les élèves dans le canton de Genève participent à plusieurs évaluations de ce type, organisées au niveau cantonal (TAF), national (enquête COFO) et international (enquête PISA).

### Au niveau cantonal : les TAF

Les Tests d'attentes fondamentales (TAF) ont été introduits en 2018 afin de disposer d'un outil de monitoring du système éducatif cantonal. Prévus dans un premier temps en 11<sup>e</sup> année, les TAF ont été élaborés en ôtant des évaluations communes (EC) de Mathématiques, Français et Allemand les tâches relatives à la vérification des attentes fondamentales définies dans le PER pour les compiler dans une épreuve indépendante. Les TAF visent à vérifier que les attentes fondamentales, communes à l'ensemble des élèves, sont maîtrisées. Les tests ne sont pas notés, mais leurs résultats sont consignés sous la forme d'une mention « atteint/non atteint » dans le bulletin scolaire de l'élève. Les attentes fondamentales ne constituent pas l'objectif à atteindre, la finalité de l'enseignement, mais sont plutôt un seuil en dessous duquel aucune et aucun élève ne devrait en principe se trouver. Autrement dit, elles doivent théoriquement être maîtrisées par toutes et tous les élèves.

Pour l'année scolaire 2018-19, le Français est la seule discipline dans laquelle la proportion d'élèves de 11<sup>e</sup> qui réussit au TAF se rapproche de cet objectif de (quasi-)exhaustivité : avec plus de neuf élèves sur dix qui atteignent les attentes fondamentales, la réussite se révèle meilleure qu'en Mathématiques et en Allemand où les taux de réussite ne sont respectivement que de 76% et 70% (Petrucci, 2020). Plus récemment, pour l'année scolaire 2021-22, il apparaît que cet objectif est en grande partie atteint en Français et en Allemand. Dans ces deux disciplines, la proportion d'élèves de 11<sup>e</sup> qui réussit au TAF est en effet respectivement de 89% et 86%. En revanche, en Mathématiques, la réussite se révèle nettement plus faible puisque seule la moitié des élèves atteint les attentes fondamentales (Petrucci, 2023). Pour les deux années en question, la réussite en 11<sup>e</sup> au TAF est différenciée en fonction de l'origine sociale, de la première langue parlée à la maison ou de la situation scolaire (en retard dans la scolarité). Ces liens entre réussite des élèves et caractéristiques sociodémographiques et scolaires seront discutés plus en détail plus bas (cf. *Les facteurs qui influencent les compétences des élèves*).

### Au niveau national : l'enquête COFO

Depuis 2006, les cantons sont tenus d'harmoniser à l'échelle suisse un certain nombre de dimensions (âge de l'entrée à l'école et durée des degrés d'enseignement) et d'objectifs. Les objectifs nationaux de forma-

tion pour la scolarité obligatoire, que la CDIP a adoptés en 2011, décrivent les compétences fondamentales dans la langue de scolarisation, dans la deuxième langue nationale, en anglais, en mathématiques et en sciences naturelles, que toutes et tous les élèves doivent avoir acquis à un certain moment de leur parcours scolaire<sup>11</sup>. Les compétences fondamentales ont été intégrées aux trois plans d'études en vigueur dans les régions linguistiques de Suisse (PER, *Lehrplan 21, Piano di studio*).

L'acquisition des compétences fondamentales est considérée comme une condition essentielle pour la poursuite des parcours éducatifs et, au-delà, pour la future participation des élèves à la vie sociale et professionnelle. Dans ce contexte, les cantons vérifient, par le biais d'enquêtes intitulées « Vérification de l'atteinte des compétences fondamentales (COFO) », dans quelle mesure les élèves suisses atteignent une partie des objectifs nationaux de formation. Ces enquêtes COFO s'inscrivent dans le cadre de l'utilisation des enquêtes à large échelle pour piloter les systèmes scolaires. Leur but est d'évaluer les performances au niveau du système suisse de formation et non pas de juger les écoles, les enseignants ou les élèves.

*Les compétences fondamentales ont été intégrées aux trois plans d'études en vigueur dans les régions linguistiques de Suisse (PER, Lehrplan 21, Piano di studio).*

Jusqu'à récemment, la Suisse ne disposait pas d'instruments nationaux de ce type, mais sa participation à l'enquête PISA (cf. plus loin : *Au niveau international : l'enquête PISA*) lui avait permis de disposer à la fois d'informations sur les performances et le fonctionnement de certains de ses systèmes éducatifs cantonaux. Ainsi, lors des enquêtes PISA de 2000 à 2012, la Suisse a constitué des sur-échantillons d'élèves en fin de scolarité obligatoire (11<sup>e</sup> année HarmoS) pour chaque région linguistique, ainsi que pour chaque canton qui le souhaitait. Ceci a permis de fournir des informations sur les compétences des élèves en Lecture, Sciences et Mathématiques, et ces dernières ont pu être mises en relation avec les caractéristiques des élèves ainsi qu'avec l'organisation scolaire. Depuis 2015, la Suisse ne participe plus qu'au volet international de PISA (élèves de 15 ans au niveau national) et nous ne disposons dès lors plus d'informations pour les

élèves de 11<sup>e</sup> année. Ce type d'information est dorénavant fourni par la vérification de l'atteinte des compétences fondamentales nationales, qui a déjà eu lieu en 2016 et 2017 (Consortium COFO, 2019a&b).

## COFO 2016

En 2016, l'enquête COFO a vérifié l'atteinte des compétences fondamentales en Mathématiques en 11<sup>e</sup> année, à la fin de la scolarité obligatoire (Consortium COFO, 2019a). Le modèle de compétences des Mathématiques, qui constitue la base des enquêtes COFO, est un modèle de compétences pluridimensionnel, dans lequel différents aspects et facteurs importants pour la description des compétences mathématiques sont différenciés et classés dans un ordre systématique. Dans ce modèle sont pris en compte des aspects de compétence<sup>12</sup> (se référant à des actions), des domaines de compétence<sup>13</sup> (se référant à des contenus), différents niveaux de compétence, une dimension évolutive (4<sup>e</sup>, 8<sup>e</sup> et 11<sup>e</sup> année de la scolarité) et des dimensions non cognitives (telles les facettes motivationnelles et sociales) (CDIP 2011, p. 6). Les domaines mathématiques choisis pour l'enquête permettent de couvrir l'ensemble des mathématiques abordées durant la scolarité obligatoire<sup>14</sup>.

Le modèle de compétences en Mathématiques est conçu pour servir de système de référence pour les standards minimaux. Ceux-ci doivent être compris comme des attentes accessibles à toutes et tous élèves. Il est donc adapté aux niveaux de performances les plus faibles.

<sup>12</sup> Huit aspects de compétence, à savoir « Savoir, reconnaître et décrire », « Appliquer des procédures et utiliser des techniques », « Présenter et communiquer », « Mathématiser et modéliser », « Argumenter et justifier », « Utiliser des instruments et des outils », « Interpréter et analyser des résultats » et « Explorer et essayer ».

<sup>13</sup> Cinq domaines de compétence, à savoir « Nombres, opérations et algèbre », « Espace », « Grandeurs et mesures », « Fonctions » et « Analyse de données et probabilités ».

<sup>14</sup> Pour certains domaines et pour certaines années, il ne saurait être question de définir des compétences fondamentales, ni d'en traiter toutes les dimensions. En effet, bien que soient prévues pour chaque année de scolarité des activités didactiques couvrant tous les contenus d'apprentissage, une grande partie d'entre elles est tout d'abord constituée d'activités de sensibilisation et d'introduction, et ce n'est que plus tard qu'elles deviennent objet de la compétence. Ainsi, des compétences fondamentales ont été développées pour la 4<sup>e</sup> année uniquement dans les domaines « Nombres, opérations et algèbre » et « Espace », et pour la 8<sup>e</sup> année, que le domaine « Analyse de données et probabilités » n'est pas développé.

<sup>11</sup> Les objectifs sont fixés pour la fin de la 4<sup>e</sup>, de la 8<sup>e</sup> et de la 11<sup>e</sup> année, à l'exception de ceux concernant la deuxième langue nationale et l'anglais, qui valent uniquement pour la 8<sup>e</sup> et la 11<sup>e</sup>.

Les élèves participant aux enquêtes COFO ont rempli un cahier de test qui ne contenait qu'un sous-ensemble de toutes les tâches du test. L'utilisation de cahiers de test a permis d'évaluer toutes les tâches, bien que chaque élève n'ait travaillé que sur une partie (4/13) de tout le matériel de test. Les tests de compétence standardisés ont été complétés par un questionnaire contextuel aux élèves sur leurs caractéristiques individuelles et leur contexte scolaire et familial qui fournit des indicateurs clés visant à expliquer les différences de réussite.

La population cible de l'enquête COFO 2016 (élèves de 11<sup>e</sup> année) regroupait plus de 80 000 élèves dans toute la Suisse. Dans la plupart des cantons participants, des écoles et des élèves ont été sélectionnés selon une procédure d'échantillonnage aléatoire en une ou deux étapes avec stratification<sup>15</sup>. Dans l'ensemble de la Suisse, 1 305 sessions de test ont été conduites et 22 423 élèves ont participé à l'enquête avec un échantillon pour chaque canton qui permet d'obtenir des résultats au niveau des différentes populations testées (Verner & Helbling, 2019a). Les passations ont été administrées selon des procédures standardisées par des administratrices et administrateurs de tests formés pour cette tâche. La passation s'est déroulée sous forme électronique en ligne en utilisant les infrastructures informatiques des écoles.

Dans toute la Suisse, 62% des élèves atteignent les compétences fondamentales sur l'échelle globale des Mathématiques. Les pourcentages au niveau cantonal varient entre 44% et 83%. La proportion dans le canton de Genève (61%) ne se distingue pas de manière statistiquement significative de l'ensemble de la Suisse (Consortium COFO, 2019a). On observe des effets de petite à moyenne ampleur de certaines caractéristiques individuelles des élèves (statut migratoire, langue[s] parlée[s] à la maison et origine sociale) sur la probabilité d'atteindre les compétences fondamentales. Les différences de composition des popula-

*COFO 2016 : en Suisse, 62% des élèves atteignent les compétences fondamentales sur l'échelle globale des Mathématiques ; cette proportion est comparable (61%) dans le canton de Genève.*

tions cantonales d'élèves expliquent peu les différences de résultats entre les cantons. En d'autres termes, les différences cantonales dans les proportions d'élèves atteignant les compétences fondamentales sont plus importantes que ce à quoi on pourrait s'attendre compte tenu de la composition des populations d'élèves.

Pour expliquer les différences cantonales dans l'atteinte des compétences fondamentales, d'autres caractéristiques doivent être prises en compte en plus de la composition de la population d'élèves. En complément des caractéristiques individuelles des élèves, le temps d'enseignement est analysé comme exemple de caractéristique au niveau du système scolaire, sur laquelle les cantons peuvent avoir une action possible. Bien qu'on observe un effet du temps d'enseignement sur l'acquisition des compétences fondamentales, cela n'explique qu'une petite partie des différences de performance entre les cantons ; nous y reviendrons.

#### COFO 2017

En 2017, l'enquête COFO a vérifié l'atteinte des compétences fondamentales en langue de scolarisation et dans la première langue étrangère en 8<sup>e</sup> année, à la fin du degré primaire (Consortium COFO, 2019b).

Pour la langue de scolarisation, les domaines de la lecture et de l'orthographe ont été évalués. Pour la première langue étrangère, la compréhension écrite et la compréhension orale ont été évaluées. Selon les cantons, la première langue étrangère correspond à l'Allemand, au Français ou à l'Anglais. Dans les cantons de Suisse romande, celle-ci correspond à l'Allemand.

Comme pour l'enquête précédente, les élèves participant à l'enquête COFO 2017 ont chacune et chacun répondu à un cahier de test pour chaque matière et chaque domaine de compétence, qui ne contient qu'un sous-ensemble de toutes les questions du test. Chaque élève n'a donc travaillé que sur une partie (4/13) de l'ensemble d'une matière et d'un domaine de compétence. Les cahiers de test ont été constitués de telle sorte que certains contiennent les mêmes groupes de questions. Ainsi, les difficultés et les performances des élèves peuvent toujours être présentées sur une échelle de mesure commune.

<sup>15</sup> Dans les plus petits cantons (AI, AR, BE-fr, GL, JU, NW, OW, SH, UR, VS-d et ZG), aucun échantillon n'a été prélevé au profit d'un recensement complet. Dans le groupe des cantons de taille moyenne (dont Genève), on a eu recours à une procédure d'échantillonnage en une seule étape. Dans le groupe des grands cantons (AG, BE-d, LU, SG, VD et ZH), on a eu recours à une procédure d'échantillonnage en deux étapes.

La population cible de l'enquête COFO 2017 (élèves de 8<sup>e</sup> année) comptait plus de 80 000 élèves dans toute la Suisse. Dans les cantons où les élèves étaient relativement peu nombreux (AI, AR, BE-fr, BS, FR-d, JU, NW, GL, SH, OW, UR, VS-d et ZG), une procédure d'échantillonnage en une seule étape a été utilisée. Toutes les écoles avec une classe de 8<sup>e</sup> année ont été invitées à participer et au sein de ces écoles, une certaine proportion d'élèves a été sélectionnée par échantillonnage. Dans les cantons les plus peuplés (dont Genève), en raison du nombre élevé d'écoles et d'élèves, une procédure d'échantillonnage en deux étapes (d'abord au niveau des écoles, puis au niveau des élèves des écoles sélectionnées) a été utilisée. Au total, 20 177 élèves de 8<sup>e</sup> année ont participé à l'enquête, avec un échantillon pour chaque canton qui permet d'obtenir des résultats au niveau des différentes populations testées (Verner & Helbling, 2019b). Les passations, au format électronique (tablettes), ont été administrées selon des procédures standardisées par des administratrices et administrateurs de tests formés pour cette tâche.

En Suisse, 88% des élèves atteignent les compétences fondamentales en lecture dans la langue de scolarisation, les pourcentages au niveau cantonal variant entre 80% et 91%. La proportion dans le canton de Genève (91%) ne se distingue pas

*COFO 2017 : 88% des élèves suisses atteignent les compétences fondamentales en lecture dans la langue de scolarisation ; ces compétences sont atteintes par une proportion similaire (91%) d'élèves genevois.*

de manière statistiquement significative de l'ensemble de la Suisse. En raison du contenu différent des tests entre régions linguistiques, les résultats en orthographe dans la langue de scolarisation ainsi qu'en compréhension orale et écrite dans la première langue étrangère sont présentés pour la Suisse romande. Dans ces cantons, 89% des élèves ont atteint les compétences fondamentales en orthographe ; les proportions cantonales varient entre 87% et 92%. À Genève, la proportion s'élève à 89%. La proportion d'atteinte des compétences fondamentales en Allemand dans les cantons romands est de 72% pour la compréhension de l'écrit (de 64% à 79% selon le canton) et de 88% pour la compréhension de l'oral (de 84% à 92% selon le canton). Les proportions correspondant au canton de Genève (64% et 84%) ne diffèrent pas significativement de ces valeurs (Consortium COFO, 2019b).

Au printemps 2023, les langues ont à nouveau été testées, mais cette fois-ci auprès des élèves de 11<sup>e</sup> année. Par rapport à 2017, la deuxième langue étrangère a également été testée. La

publication des résultats de l'enquête COFO 2023 est prévue pour fin 2024. Aussi en 2024, mais au printemps, ce sont les compétences fondamentales des élèves de la 4<sup>e</sup> année Har-moS en langue de scolarisation et en Mathématiques qui seront évaluées.

### **Au niveau international : l'enquête PISA**

Dans les années 1950, un groupe de chercheurs a entrepris pour la première fois une évaluation standardisée internationale à large échelle, qui rassemblait seulement 12 pays et concernait uniquement l'évaluation de la performance des élèves en Mathématiques (en postulant que cette discipline était la plus facile à traduire en différentes langues), avec l'espoir qu'à travers de telles évaluations les pays pouvaient apprendre les uns des autres (Foshay, Thorndike, Hotyat, Pidgeon, & Walker, 1962). À partir des années 1990, les évaluations internationales à large échelle ont gagné en envergure et en visibilité lorsque les décideurs politiques ont réalisé l'utilité de disposer de données éducatives fiables, basées sur des

preuves scientifiques, pour informer et guider la prise de décision en matière de politiques éducatives (Hastedt & Rocher, 2020).

Actuellement, les programmes d'évaluation les plus connus sont

ceux pilotés par l'Association internationale pour l'évaluation de la réussite éducative (IEA) – TIMSS, PIRLS, ICCS et ICILS<sup>16</sup> – et l'enquête PISA<sup>17</sup>, coordonnée par l'Organisation de coopération et de développement économiques (OCDE). La Suisse n'a participé que très occasionnellement aux programmes de l'IEA<sup>18</sup>. En revanche, le pays – et par-là même le canton de Genève – participe à PISA depuis la première enquête en 2000, et ceci tous les trois ans.

<sup>16</sup> TIMSS : *Trends in International Mathematics and Science Study*; PIRLS : *Progress in International Reading Literacy Study*; ICCS : *International Civic and Citizenship Education Study*; ICILS : *International Computer and Information Literacy Study*.

<sup>17</sup> PISA : Programme international pour le suivi des acquis des élèves.

<sup>18</sup> La Suisse, et plus précisément la partie alémanique, a participé une seule fois lors de la première édition de l'enquête TIMSS 8, en 1995. Elle a également participé lors de la première édition de l'enquête ICCS en 2009. De même, l'enquête ICILS a été menée une seule fois en Suisse, en 2013. La Suisse n'a par contre jamais participé à l'enquête PIRLS.



Bien que les différents programmes des deux organismes présentent des similarités, ils ont aussi des différences substantielles. Tout d'abord, l'enquête PISA est une enquête qui vise à évaluer les systèmes scolaires par l'intermédiaire de la mesure des compétences des élèves de 15 ans, indépendamment de leur parcours scolaire, alors que les échantillons des études IEA sont basés sur les niveaux scolaires (degrés 4, 8 ou 12 selon la classification CITE définie par l'UNESCO) pour mieux refléter la structure du programme scolaire. En plus, les études IEA cherchent à mesurer ce qui est enseigné dans les écoles et les contextes de l'apprentissage. Les programmes des pays participants sont donc au centre des analyses et ceux-ci développent ensemble le cadre d'évaluation et les matériaux de test. Plutôt que d'évaluer ce que les écoles enseignent, l'OCDE a décidé, quant à elle, de se focaliser sur les compétences attendues pour travailler dans une « économie moderne ». L'enquête PISA cherche donc à évaluer l'acquisition de compétences sélectionnées par les élèves vers la fin de leur scolarité obligatoire. Ce sont des experts et expertes de l'OCDE qui déterminent quelles compétences les élèves doivent avoir maîtrisées (pour les utiliser plus tard dans la vie) et qui élaborent les matériaux pour l'évaluation (Hastedt & Roher, 2020).

Les domaines examinés sont la Lecture, les Mathématiques et les Sciences. Tous les trois ans, l'un des trois domaines de compétences constitue le domaine principal et fait l'objet de tests plus approfondis. Ainsi, la Lecture a été le domaine principal en 2000, 2009 et 2018. Il est à noter que l'expression « compréhension de l'écrit » a été préférée au terme de « lecture », parce qu'elle est susceptible de mieux traduire l'idée de ce que l'enquête va évaluer. En effet, par « lecture », on entend souvent un simple décodage, ou même le fait de lire à haute voix, alors que « la compréhension de l'écrit » fait appel à un large éventail de compétences cognitives, qui vont de la faculté élémentaire de décoder l'écrit à la connaissance du vocabulaire, de la grammaire et plus largement des structures et caractéristiques linguistiques et textuelles, en passant par la connaissance du monde. Ce concept tient notamment compte du fait que les habitudes de lecture ont considérablement changé ces der-

*L'enquête PISA vise à évaluer les systèmes scolaires par l'intermédiaire de la mesure des compétences des élèves de 15 ans, indépendamment de leur parcours scolaire (...)  
L'OCDE a décidé de se focaliser sur les compétences attendues pour travailler dans une « économie moderne ».*

nières années en raison de l'apparition des médias numériques. Tous ces aspects sont pris en considération dans la définition PISA de la compréhension de l'écrit en 2018 (OCDE, 2019, p. 35): « *Comprendre l'écrit, c'est non seulement comprendre et utiliser des textes écrits, mais aussi réfléchir à leur propos et s'y engager. Cette capacité devrait permettre à chacun de réaliser ses objectifs, de développer ses connaissances et son potentiel, et de prendre une part active dans la société.* »

De même, dans l'enquête PISA, l'expression « culture mathématique » a été préférée au terme de « Mathématiques ». La culture mathématique a été le domaine principal en 2003, 2012 et 2022. Son cadre d'évaluation a été revu et actualisé en 2012. La culture mathématique dans le cadre de l'enquête PISA se définit désormais comme suit (OCDE, 2013, p. 27): « *dans l'enquête PISA 2015, "la culture mathématique" est l'aptitude d'un individu à formuler, employer et interpréter des mathématiques dans un éventail de contextes, soit à se livrer à un raisonnement mathématique et à utiliser des concepts, procédures, faits et outils mathématiques pour décrire, expliquer et prévoir des phénomènes. Elle aide les individus à comprendre le rôle que les mathématiques jouent dans le monde et à se comporter en citoyens constructifs, engagés et réfléchis, c'est-à-dire à poser des jugements et à prendre des décisions en toute connaissance de cause.* »

Cette définition englobe trois aspects interdépendants: (1) les processus mathématiques qui décrivent ce que font des individus pour établir un lien entre le contexte du problème et les mathématiques et, donc, pour résoudre le problème, ainsi que les facultés qui soutiennent ces processus, (2) les contenus mathématiques qui doivent être utilisés pour répondre aux items et (3) les contextes dans lesquels les items s'inscrivent.

Par définition, la culture mathématique renvoie à la capacité des individus à formuler, employer et interpréter les mathématiques. Ces trois verbes sont chacun associés à l'un de trois processus mathématiques. Le verbe *formuler* renvoie à la capacité des individus à identifier et à reconnaître des possibilités d'utiliser les mathématiques dans le contexte



d'un problème, puis à structurer sous forme mathématique un problème présenté jusqu'à un certain point sous une forme contextualisée. Le verbe *employer* renvoie à la capacité des individus à appliquer des concepts, faits, procédures et raisonnements mathématiques pour résoudre des problèmes énoncés de façon mathématique afin d'aboutir à des conclusions mathématiques. Le verbe *interpréter* renvoie à la capacité des individus à réfléchir à des solutions, des résultats ou des conclusions mathématiques, et à les interpréter dans le cadre de problèmes tirés du monde réel.

Enfin, les Sciences ont été le domaine principal en 2006 et 2015, et le seront à nouveau lors de l'enquête PISA 2025. Tout comme pour les autres domaines, l'expression « culture scientifique » a été préférée au terme « sciences » pour souligner l'importance que les épreuves PISA accordent à l'application des connaissances scientifiques dans des situations tirées de la vie courante. Son cadre théorique a été révisé en 2015 et la culture scientifique dans le cadre de l'enquête PISA se définit désormais comme la capacité (1) d'expliquer des phénomènes de manière scientifique, (2) d'évaluer et concevoir des recherches scientifiques et (3) d'interpréter des données et des faits de manière scientifique. L'exploitation de ces trois compétences requièrent à leur tour trois types de connaissances : scientifiques, procédurales et épistémiques (OCDE, 2016, p. 24).

PISA est une enquête d'une envergure exceptionnelle non seulement en termes de domaines de compétences couverts et du nombre d'items évalués, mais également du nombre d'élèves testés. En 2000, 250 000 élèves sur 31 pays, dont plus de 13 000 en Suisse et environ 5 000 en Suisse romande, ont participé à la première enquête (Consortium PISA romand, 2001). Leur nombre était de 600 000 pour l'enquête PISA 2018. Cet échantillon représente environ 32 000 000 de jeunes de 15 ans dans 79 pays participants (37 pays membres de l'OCDE et 42 pays ou économies partenaires). Rappelons que la Suisse ne constitue plus des sur-échantillons d'élèves en fin de scolarité obligatoire (11<sup>e</sup> année HarmoS) depuis 2015 et qu'elle ne participe plus qu'au volet international de PISA (élèves de 15 ans au niveau national). Dès lors, le nombre d'élèves suisses participant à l'enquête est moins important. En 2018, un

échantillon de 6 511 élèves représentant l'ensemble de la population des élèves de 15 ans de Suisse<sup>19</sup> a été tiré, dont 1 412 pour la Suisse romande. La proportion d'élèves de 15 ans fréquentant l'enseignement obligatoire diffère grandement d'une région linguistique à l'autre ; elle s'élevait à près de 80% et 60% en Suisse alémanique et Suisse romande respectivement, alors qu'en Suisse italienne, près de 80% des élèves de l'échantillon suivaient une formation dans le secondaire II (Consortium PISA.ch, 2019).

Les scores PISA n'ont pas de signification en soi, mais sont déterminés en fonction de la variation des résultats de toutes et tous les élèves aux épreuves. Pour en faciliter la lecture, les résultats sont mis à l'échelle pour produire une distribution normale où la moyenne des pays de l'OCDE est fixe à 500 points et l'écart-type à 0 points. Environ deux tiers des élèves ont alors un score situé entre 400 et 600 points et en théorie, il n'y a ni score maximum, ni score minimum aux épreuves. En plus, pour interpréter les scores des élèves, les échelles PISA sont également divisées en niveaux de compétence. Les tâches se répartissent entre huit degrés de difficulté (1c, 1b, 1a, 2, 3, 4, 5 et 6). Les tâches classées au niveau 1c de l'échelle de compétence sont les plus faciles. Les tâches au niveau 2 correspondent au niveau minimal de compétences pour participer effectivement à la vie courante. Ainsi, une ou un élève situé au niveau 1c est susceptible de mener à bien des tâches de ce niveau, mais pas des tâches situées à des niveaux supérieurs. Une ou un élève de niveau 6 est capable de mener à bien les tâches les plus difficiles et nécessitant un niveau de compétence plus élevé, mais également toutes les autres tâches du domaine d'évaluation visé. Chaque niveau de compétence représente de l'ordre de 80 points (OCDE, 2019).

*Les performances des jeunes Suisses peuvent être globalement considérées comme satisfaisantes. Elles et ils obtiennent en effet systématiquement de meilleurs résultats que la moyenne des pays de l'OCDE en Mathématiques et en Sciences.*

<sup>19</sup> Plus précisément, les résultats de PISA 2018 permettent de se prononcer sur l'ensemble de la population des élèves de 15 ans, à l'exception qui ont été exclus selon des règles définies par la direction internationale du projet. En Suisse, cela concerne les élèves (1) des écoles spéciales, (2) des écoles internationales, (3) qui souffrent d'une déficience cognitive ou fonctionnelle, ou (4) qui ont un très faible niveau de connaissance de la langue du test.

En partant des enquêtes PISA<sup>20</sup>, et comparées à d'autres pays, les performances des jeunes Suisses peuvent être globalement considérées comme satisfaisantes. Elles et ils obtiennent en effet systématiquement de meilleurs résultats que la moyenne des pays de l'OCDE en Mathématiques et en Sciences (cf. *Tableau 1*). Ces écarts sont statistiquement significatifs. Le bilan est plus mitigé en lecture. Comparé à la moyenne de l'OCDE, la Suisse a obtenu un résultat moyen en 2000. Si elle a su se redresser en 2009, en obtenant un résultat statistiquement supérieur à la moyenne de l'OCDE, on observe une diminution des performances moyennes depuis 2012. Malgré cela, la Suisse fait partie des dix-huit pays qui ont obtenu, en 2022, des résultats supérieurs à la moyenne OCDE dans les trois domaines (Erzinger, Pham, Prosperi, & Salvisberg, 2023).

**Tableau 1** : Évolution des moyennes de l'OCDE, de la Suisse (CH) et de la Suisse romande (CH-fr) dans les trois domaines (Lecture, Mathématiques et Sciences – domaines principaux en gras)

		2000	2003	2006	2009	2012	2015	2018	2022
OCDE	L	500	494	492	493	496	493	487	476
	M	500	500	498	497	494	490	489	472
	S	500	500	500	501	501	493	489	485
CH	L	494	499	499	501	506	492	484	483
	M	529	527	527	534	530	521	515	508
	S	496	513	512	517	515	506	495	503
CH-fr	L	504	499	496	506	509	--	--	--
	M	535	528	533	530	523	--	--	--
	S	497	509	505	500	500	--	--	--

On constate qu'en Suisse, la performance générale des élèves de 15 ans s'améliore légèrement à partir de la première enquête, mais qu'elle a particulièrement diminué depuis 2012. On observe la même tendance à la baisse des valeurs moyennes de performance dans la plupart des pays de référence et ce, de manière encore plus prononcée<sup>21</sup>. Cette évolution globale est comparable aux deux extrémités du spectre de performance. En mathématiques, par exemple, le pourcentage d'élèves très performants a diminué de 4,4 points de pourcentage entre 2012 et 2018 et le pourcentage d'élèves peu performants a augmenté dans la même mesure. Entre 2009 et 2018, le pourcentage d'élèves de 15 ans issus de l'immigration

en Suisse a augmenté de 10 points de pourcentage, l'une des hausses les plus fortes dans les pays de l'OCDE. Comme en Suisse, les élèves issus de l'immigration ont obtenu en compréhension de l'écrit un score inférieur de 50 points environ à celui des élèves autochtones en 2009 et en 2018, l'évolution du pourcentage d'élèves issus de l'immigration et d'élèves autochtones peut à elle seule expliquer une variation de 5 points environ, soit un tiers des 17 points de différence, de l'écart de score moyen entre les évaluations PISA de 2009 et de 2018 (OCDE, 2019).

Lorsque l'on compare les résultats des élèves des trois régions linguistiques (de 2000 à 2012), la première enquête positionne la Suisse romande en première place dans les trois domaines évalués. À l'issue des enquêtes suivantes, ce sont les élèves de la Suisse alémanique qui ont obtenu les meilleurs résultats dans les trois domaines, sauf pour la lecture en 2009. Lors des cinq enquêtes en considération, la Suisse italienne a obtenu les moins bonnes moyennes du pays. Si l'on passe à un niveau d'analyse encore plus fine et que l'on s'intéresse à la comparaison entre cantons romands, on observe qu'en 2000, 2003 et 2006, Genève a obtenu les moyennes les plus faibles pour chaque domaine des cantons romands et se distinguait de l'ensemble des autres cantons. À partir de 2009, les résultats des élèves genevois s'apparentent davantage aux résultats observés dans d'autres cantons (notamment Berne, Neuchâtel et Vaud), bien que Genève reste en queue du peloton romand.

### Que nous disent ces outils de monitoring des compétences des élèves genevois ?

Les enquêtes PISA d'abord et les enquêtes COFO après 2015 nous apprennent que, de manière générale, les résultats moyens du canton ne se distinguent pas de ceux observés en Suisse ou en Suisse romande, quels que soient le domaine de compétence évalué et l'année de scolarité considérée. Ainsi, en mathématiques, 61% des élèves genevois maîtrisent les compétences fondamentales en 11<sup>e</sup> (contre 62% en Suisse) et 91% des élèves de 8<sup>e</sup> atteignent les compétences fondamentales en lecture dans la langue de scolarisation (contre 88% en Suisse). Si on peut se réjouir d'une réussite conforme à la moyenne nationale, il n'en demeure pas moins que les résultats du canton apparaissent préoccupants à certains égards, notamment en mathématiques. En effet, tout d'abord seuls six élèves

<sup>20</sup> Les publications issues des différentes enquêtes PISA peuvent être consultées sur le site de PISA Suisse à l'adresse suivante : [www.pisa-suisse.ch/](http://www.pisa-suisse.ch/)

<sup>21</sup> Les pays de référence concernent les pays limitrophes, plus la Belgique, le Canada, la Finlande et le Luxembourg.

sur dix atteignent les compétences fondamentales dans ce domaine. Par ailleurs, l'observation de la réussite par filière a mis en lumière des chiffres particulièrement inquiétants en LC et CT où les taux d'atteinte des compétences fondamentales ne sont respectivement que de 26% et 11%, soit des proportions extrêmement éloignées des objectifs de (quasi) exhaustivité affichés dans COFO. Les tests cantonaux d'attentes fondamentales (TAF), bien que ne mesurant pas exactement le même type de compétences, sont venus confirmer, à plusieurs reprises, les constats préoccupants tirés de COFO. En 2021-22, par exemple, seulement 50% des élèves de 11<sup>e</sup> maîtrisaient les attentes fondamentales en mathématiques. Ces chiffres diffèrent en outre sensiblement d'une filière à l'autre puisque 68% des élèves de LS réussissent au TAF contre respectivement 15% et 3% des élèves de LC et de CT. Les difficultés des élèves de ces deux dernières filières sont d'ailleurs très largement corroborées par les études sur les parcours de formation puisqu'ils constituent la quasi-totalité des 15% de jeunes qui sont orientés chaque année vers le préqualifiant à l'issue de la scolarité obligatoire (environ 70% des élèves de CT et 17% des élèves de LC ; Rastoldo & Mouad, 2021).

La visée résolument comparative de l'enquête COFO a également permis d'établir, sur la base du calcul de taux de réussite « ajustés », que compte tenu de la composition sociodémographique de sa population scolaire, Genève obtient une réussite en lecture (langue de scolarisation) en fin de 8<sup>e</sup> tout à fait conforme à ce que l'on peut espérer. En revanche, les résultats en mathématiques en fin de scolarité obligatoire pourraient être un peu plus élevés (Consortium COFO, 2019). En d'autres termes, cela signifie que d'autres cantons ayant une population scolaire de composition similaire ont obtenu de meilleurs résultats et qu'une partie des résultats du canton est bel et bien liée à la manière dont l'école est organisée et fonctionne. Conformément aux résultats de la littérature sur l'effet des filières, l'enquête COFO a notamment permis de vérifier qu'à Genève, comme ailleurs, l'assignation à un groupe de niveau faible a malheureusement des effets négatifs sur de nombreuses dimensions liées aux attitudes et à la motivation face aux apprentissages scolaires (cf. *Les effets du contexte sur les compétences des élèves*).

## Les facteurs qui influencent les compétences des élèves

Cette note se concentre essentiellement sur les élèves de l'enseignement obligatoire régulier pour qui une ou plusieurs évaluations sont disponibles, sous forme de notes ou d'appréciations de type « acquis / non acquis ». Pour les élèves de l'enseignement spécialisé et les élèves des classes d'accueil, de tels éléments d'information n'existent pas. Nous reviendrons sur ce point en conclusion de ce travail.

### Les compétences des élèves selon les caractéristiques individuelles

Les résultats des élèves aux évaluations faites en classe et aux épreuves externes (TAF, COFO et PISA) sont largement marqués par les caractéristiques sociales et scolaires des élèves. Le genre, l'origine sociale et le statut migratoire, appréhendé principalement à travers la nationalité ou l'allophonie, sont autant de facteurs de risque pour le parcours scolaire des élèves, et ce depuis toujours – ou tout du moins depuis que des données existent pour en rendre compte (Denecker et al., 2023a et b ; Petrucci, 2023 ; SRED, 2022 ; CSRE, 2023 ; OCDE, 2019). Il en va de même avec les facteurs scolaires. Ainsi, les élèves qui ont connu un redoublement dans leur parcours sont plus à risque de rencontrer des difficultés.

Les élèves genevois n'échappent pas à ces constats. Les repères et indicateurs statistiques qui rendent compte des compétences des élèves de l'enseignement obligatoire mettent systématiquement en évidence un marquage scolaire et social des résultats aux épreuves cantonales (Denecker et al., 2023a et b ; Petrucci, 2023). Dès la 4<sup>e</sup>, si la très grande majorité des élèves atteignent les seuils de réussite en Français et en Mathématiques (respectivement 71.9% et 88.1%), certains groupes d'élèves obtiennent des taux de réussite moins élevés. C'est le cas des élèves de milieu modeste, des élèves allophones et, en Français uniquement, des garçons. Les élèves qui ont déjà redoublé et celles et ceux qui sont « en retard » sont également concernés. Ces résultats observés en 4<sup>e</sup> se retrouvent en 8<sup>e</sup> (Denecker et al., 2023b) puis en 11<sup>e</sup> aux tests d'attentes fondamentales (Petrucci, 2023). Les élèves de milieu modeste, allophones et en retard dans leur parcours obtiennent en moyenne des résultats inférieurs à

ceux des élèves de milieu plus favorisé, francophones et sans retard dans leur parcours. Concernant les filles et les garçons, les résultats dépendent de la discipline : les filles ont en moyenne des résultats supérieurs en Français, alors que la tendance inverse s'observe pour les Mathématiques. Par ailleurs, plus les élèves avancent dans leur scolarité, plus les écarts se creusent.

En 2023, le dernier rapport national sur l'éducation relève, sur la base du suivi d'une cohorte d'élèves scolarisés en 3P en 2012-2013, que les élèves étrangers et de faible statut socioéconomique – tout comme les élèves romands de faible statut socioéconomique – ont connu des parcours scolaires moins linéaires que les autres élèves (CSRE, 2023).

Ces résultats trouvent-ils un écho dans les résultats des enquêtes nationales (tests COFO) et internationales (PISA) ? Comme le relèvent Petrucci et Roos (2020) concernant l'atteinte des compétences fondamentales de Mathématiques en 11<sup>e</sup> en 2016, « l'enquête COFO a [...] permis de vérifier une nouvelle fois qu'il existe un lien entre l'acquisition des compétences et les caractéristiques sociodémographiques usuelles des élèves que sont le genre, l'origine sociale, le statut migratoire ou la langue parlée à la maison » (p. 27).

Au niveau international, les enquêtes PISA interrogent également de manière systématique l'équité des systèmes éducatifs. Pour chaque enquête, une publication se concentre spécifiquement sur les écarts de performance liés au genre, à l'origine sociale, et au statut migratoire, sur l'évolution de ces écarts et sur les disparités d'équité en éducation d'un pays à l'autre, avec des constats en grande partie identiques à ceux observés à Genève et en Suisse, mais avec une variation de l'ampleur des effets d'un pays à l'autre (OCDE, 2013, 2016, 2019). Notons à cet égard que le dernier rapport national PISA (Erzinger et al., 2022) confirme la relation entre l'origine sociale et la réussite scolaire. Si les élèves d'origine sociale défavorisée obtiennent en moyenne de moins bons résultats en mathématiques dans toutes les enquêtes PISA, l'écart entre les élèves socialement défavorisés et les élèves privilégiés n'a jamais été aussi important que dans PISA 2022.

*Si les élèves d'origine sociale défavorisée obtiennent en moyenne de moins bons résultats en mathématiques dans toutes les enquêtes PISA, l'écart entre les élèves socialement défavorisés et les élèves privilégiés n'a jamais été aussi important que dans PISA 2022.*

Dans une des publications liées à l'enquête PISA 2015 (OCDE, 2016), on voit poindre une possible double lecture des raisons de ces différences de résultats. D'une part, il est écrit d'emblée que « [le] Canada, le Danemark, l'Estonie, Hong-Kong (Chine) et Macao (Chine) se distinguent par des niveaux élevés de performance et par un rendement de l'éducation plus équitable » (OCDE, 2016, p. 20). Mais d'autre part, « [environ] 29% des élèves défavorisés sont considérés comme résilients : ils déjouent les pronostics et obtiennent des scores élevés en sciences » (OCDE, 2016, p. 20). D'un côté, ce seraient donc les systèmes éducatifs nationaux qui permettraient plus d'équité, de l'autre ce serait le fruit de la résilience de certaines et certains élèves. Cette double lecture fait porter la responsabilité des compétences acquises plutôt sur l'école ou plutôt sur les élèves (et leur environnement social et familial). Mais la recherche montre largement que les apprentissages des élèves dépendent – au moins en partie – du contexte

et des politiques scolaires (voir ci-dessus ; voir aussi OCDE, 2020). La partie suivante sera consacrée aux effets de contexte. Mais arrêtons-nous un instant sur les effets liés aux caractéristiques des élèves.

La littérature scientifique sur cette thématique est

riche. Qu'elles soient ancrées en sociologie, psychologie sociale, sciences de l'éducation ou économie par exemple, les recherches visent une meilleure compréhension des plus faibles performances et, de fait, des parcours scolaires plus complexes d'une partie de la population scolaire. Deux axes de travail sont évoqués ici.

- Une première manière de lire les compétences des élèves renvoie à l'écart entre la culture, les normes et les valeurs scolaires d'une part, et la culture, les normes et les valeurs familiales d'autre part. Cette distance culturelle expliquerait en partie pourquoi les élèves de milieu défavorisé ou issus de la migration obtiennent en moyenne des performances systématiquement inférieures aux élèves de milieu plus favorisé et autochtones. Si les données mentionnées plus haut renseignent les écarts de réussite entre élèves, les travaux menés au SRED et ailleurs suggèrent que l'entrée à l'école marque déjà une différence entre les élèves familiaux avec la culture scolaire et les autres (Cecchini, Dutrévis, Guilley & Jaunin,



2023, à *paraître*). Cet écart est d'autant plus difficile à combler que l'école fonctionne avec de nombreux implicites qui sont de fait plus facilement décodés par celles et ceux qui sont proches de la culture scolaire (Dutrévis & Rastoldo, 2019; Rayou & Bautier, 2009). Notons néanmoins que le projet d'innovation pédagogique EP21 cherche à réduire cet écart avec un important travail autour de l'explicitation des attentes institutionnelles.

- Une seconde lecture a trait aux préjugés et stéréotypes concernant certains groupes sociaux et les compétences qu'ils seraient à même de démontrer. Dans le domaine scolaire, certains groupes d'élèves sont réputés moins intelligents, moins aptes à apprendre ou encore moins « (pré-)destinés » à certaines disciplines et professions (Dutrévis, Soussi & Genoud, 2018). Ces réputations ne sont pas sans conséquences en termes de développement et de démonstration des compétences pour les élèves qu'elles visent. Ainsi, confrontés de manière récurrente à une image de leur groupe, les élèves peuvent intérioriser un sentiment de compétence général à l'école ou spécifique à certains domaines (Cvencek, Meltzoff & Greenwald, 2011; Plante, de la Sablonnière, Aronson & Théorêt, 2013). Ces réputations peuvent aussi être véhiculées, très souvent de manière inconsciente, en contexte scolaire (Denessen, Hornstra, van den Bergh & Bijlstra, 2022; Durante & Fiske, 2017). Et, même lorsque les élèves n'adhèrent pas aux images utilisées pour les décrire, celles-ci constituent une chape de plomb qui plane au-dessus de leur tête : démontrer ses compétences en contexte évaluatif devient alors un enjeu d'autant plus essentiel qu'ils et elles doivent, en plus de démontrer leurs compétences, contredire la réputation qui les concerne et les faibles attentes qui en découlent (voir notamment les travaux sur la menace du stéréotype, Steele, 1997)<sup>22</sup>.

À travers ces deux exemples, on entrevoit déjà – et nous y reviendrons plus tard – le risque que les évaluations ne reflètent pas toujours de manière fidèle les compétences des élèves. En effet, ces compétences se formalisent dans un contexte social et scolaire qui comprend plus ou moins d'obstacles vers la réussite selon les caractéristiques sociales des élèves. Certains doivent décoder l'école et ses normes, s'octroyer le droit de choisir certaines voies ou filières, ou encore résister

à l'image que d'autres se font de leurs capacités à réussir.

Après ces premiers éléments sur les caractéristiques individuelles des élèves, abordons maintenant les effets du contexte scolaire. Il s'agit alors de voir comment l'école accueille et organise l'apprentissage des élèves et comment cette organisation scolaire peut, à différents niveaux, influencer les compétences des élèves.

### Les effets du contexte sur les compétences des élèves

Depuis les années 1960, l'intérêt pour la problématique de l'effet du contexte scolaire sur les acquisitions des élèves n'a cessé de s'accroître. De très nombreuses recherches ont essayé de voir dans quelle mesure l'apprentissage d'une ou d'un élève est influencé par le contexte scolaire, ce dernier pouvant tour à tour correspondre à l'établissement ou la classe (Bressoux, 1994; Crahay, 2006; Dumay & Dupriez, 2009). Aujourd'hui, il existe au sein de la communauté scientifique un relatif consensus sur l'importance de facteurs contextuels dans l'explication des variations d'acquisitions de compétences des élèves. C'est notamment pour cette raison que certaines évaluations standardisées à large échelle comme PISA ou COFO collectent, en plus des estimations sur les performances, une multitude de renseignements sur les élèves, leurs familles et leurs écoles, autant d'éléments qui doivent conduire à une meilleure compréhension de l'impact des politiques et des pratiques sur les résultats des élèves (Rutkowski et al., 2014).

D'une manière générale, la question des niveaux contextuels les plus pertinents à considérer s'est posée rapidement. Ainsi, à la suite des premiers travaux de Coleman, Campbell, Hobson, Mcpartland, Mood, Weinfeld, & York (1966) centrés sur l'effet-établissement, d'autres recherches se sont davantage intéressées au contexte de la classe et ont permis de constituer un savoir spécifique sur la taille des différents effets. La supériorité explicative de l'effet-classe par rapport à l'effet-établissement est désormais clairement établie (Muijs, Kyriakides, van der Werf, Creemers, Timperley, & Earl, 2014). C'est bien au niveau de la classe que se joue la plus grande part de la variabilité des acquisitions dans la plupart des pays. Ainsi, selon Bianco et Bressoux (2009), « *la plupart des études montrent que la classe explique entre 10% et 20% de la variance des acquisitions des élèves [...]. On re-*

<sup>22</sup> Pour un accès plus complet aux travaux dans ce champ : <https://www.reducingstereotypethreat.org/>



lève, certes, des variations dans ces effets d'un pays à l'autre, mais dans la plupart des cas, l'effet-classe se révèle plus important que l'effet-école, généralement estimé aux alentours de 4 à 5% » (p. 36). Ce constat s'explique en grande partie par le fait que les facteurs les plus proches de l'expérience scolaire de l'élève sont les plus directement liés aux résultats, l'effet de facteurs plus éloignés étant plus faible (Muijs & Reynolds, 2002). Les élèves ne font pas l'expérience d'une école dans son ensemble, mais elles et ils font davantage l'expérience de la classe au sein de celle-ci (Reynolds et al., 2014). La théorie des systèmes écologiques stipule d'ailleurs qu'il existe une relation proportionnellement inverse entre l'ampleur de l'effet de contexte et la distance qui sépare ce dernier de l'individu (Leroy, 2009). Dans ce cadre, l'établissement, même s'il constitue une unité pertinente d'analyse des différences d'acquisition de compétences, apparaît comme un contexte de scolarisation plus éloigné de l'élève que ne l'est la classe. Ceci fait également écho aux conclusions de Van Damme et al. (2009) qui indiquent qu'il convient de « *rappeler que l'influence des écoles est limitée lorsqu'on la compare à l'influence des caractéristiques des élèves et de leurs familles. Il semble par ailleurs clair que les variables "école" qui ont le plus d'impact sur les performances des élèves sont les variables les plus directement connectées au processus d'apprentissage – et donc les méthodes d'enseignement et la qualité de l'éducation –, par comparaison avec l'effet de variables plus distales, comme la gestion de l'établissement* » (p. 34).

En ce qui concerne l'effet-classe, un grand nombre de données empiriques permettent d'affirmer aujourd'hui que ce dernier est majoritairement un effet-enseignant. Les caractéristiques morphologiques de la classe (nombre d'élèves, nombre de cours...) et les caractéristiques du public accueilli (niveau moyen d'acquisitions, pourcentage d'élèves défavorisés, pourcentage d'étrangers, hétérogénéité...) ont certes une influence sur les acquisitions de compétences des élèves par l'imposition de contraintes spécifiques dans l'organisation et la gestion de la classe, mais la plupart du temps, elles ne permettent pas d'expliquer toute l'ampleur de l'effet-classe

*Des décennies de recherche sur l'efficacité en éducation ont non seulement montré que le niveau de la classe peut expliquer une plus grande partie de la variance des résultats des élèves que le niveau de l'école, mais aussi qu'une grande partie de cette variance au niveau de la classe peut être expliquée par ce que les enseignantes et enseignants font en classe.*

(Bressoux, 1995). Des décennies de recherche sur l'efficacité en éducation ont non seulement montré que le niveau de la classe peut expliquer une plus grande partie de la variance des résultats des élèves que le niveau de l'école, mais aussi qu'une grande partie de cette variance au niveau de la classe peut être expliquée par ce que les enseignantes et enseignants font en classe (Muijs & Reynolds, 2017). Les approches centrées sur les pratiques d'enseignement ont en effet produit des résultats nettement plus probants que celles portant sur d'autres caractéristiques de l'enseignante ou enseignant, telles que les traits de personnalité, le genre, l'âge, le niveau de formation ou les années d'expérience (Borrich, 1996 ; Scheerens, 2016). Autrement dit, comme l'a souligné Talbot (2012), il faut davantage s'attacher à ce que *fait* la ou le professeur plutôt qu'à ce qu'elle ou il est.

L'une des principales conclusions des revues de littérature sur l'enseignement efficace est que ce dernier correspond davantage au fait d'être capable de faire correctement un grand nombre de « petites » choses plutôt que de faire correctement un petit nombre de « grandes » choses (Reynolds et al., 2014). Au Royaume-Uni, par exemple, près de soixante pratiques enseignantes différentes ont été identifiées comme étant liées à l'amélioration des performances des élèves de l'enseignement primaire au cours de

l'année scolaire dans l'évaluation du *Mathematics Enhancement Programme* (Muijs & Reynolds, 2000). En se basant sur des recherches utilisant un modèle dit « processus-produit »<sup>23</sup>, une multitude d'études a permis d'établir des listes de pratiques enseignantes efficaces positivement liées aux résultats des élèves en matière de compétences de base (Brophy & Good, 1986 ; Doyle, 1986 ; Muijs & Reynolds, 2002, 2017).

<sup>23</sup> Les travaux réalisés dans le cadre du paradigme « processus-produit » visent à examiner la relation entre le comportement des enseignants et les résultats des élèves. Ils impliquent d'utiliser un pré-test d'une certaine mesure de la réussite scolaire comme la lecture, un post-test administré plus tard, et de faire une observation approfondie des différents comportements, attitudes et climats en classe de grands échantillons d'enseignants différents, dont les « gains » en termes de niveaux de réussite de leurs élèves sont liés à leurs caractéristiques (Bressoux, 1994 ; Crahay, 2006 ; Muijs & Reynolds, 2017).

En synthèse, parmi les pratiques évoquées de façon récurrente, on retrouve notamment :

- La *clarté de l'enseignement et des routines* de gestion de classe (Rosenshine, 1970).
- Des *opportunités d'apprentissage élevées* qui résultent de la couverture du plan d'études. Les enseignantes et enseignants efficaces offrent à leurs élèves de meilleures opportunités d'apprentissage en termes de temps d'apprentissage effectif, de cohérence entre le curriculum prévu, mis en œuvre et réalisé, de rapport approprié entre la matière couverte et la durée de la leçon, de fidélité aux objectifs du plan d'études et de pertinence du matériel pédagogique (Scheerens, 2016 ; Stallings, 1985).
- La *structuration de la leçon* qui facilite entre autres la mémorisation des informations et permet aux élèves de les comprendre comme un tout intégré, en reconnaissant les relations entre les parties (Creemers & Kyriakides, 2008). Selon Rosenshine et Stevens (1986), les performances des élèves sont maximisées lorsque les enseignantes et enseignants structurent la leçon en commençant par une vue d'ensemble et/ou une révision des objectifs, qu'ils soulignent le contenu à couvrir et signalent les transitions entre les parties de la leçon, attirent l'attention sur les idées principales et renvoient ces dernières à la fin du cours.
- Une *gestion de classe* qui maximise l'attention des élèves et le temps qu'elles et ils consacrent aux tâches (temps d'apprentissage actif). La classe est organisée et gérée comme un environnement d'enseignement dans lequel les activités se déroulent sans heurts, les transitions sont brèves et peu de temps est consacré à l'organisation ou à la gestion des comportements difficiles (Brophy & Good, 1986). L'étude de Day, Sammons, Kington, Regan, Ko, Brown, Gunraj, & Robertson (2008) a par ailleurs montré que les enseignants efficaces font preuve d'une gestion de classe proactive, en utilisant des pratiques et des stratégies éducatives visant prioritairement à prévenir les écarts de conduite des élèves.
- Un *enseignement actif* qui apporte le contenu du plan d'études aux élèves. Les enseignantes et enseignants efficaces passent la plupart de leur temps en classe à enseigner ou à superviser les élèves plutôt que de les laisser travailler seuls (la discussion dirigée par l'enseignant domine par rapport au travail individuel). L'enseignant efficace joue un rôle actif et ne se contente pas de « faciliter » l'apprentissage des élèves. Le recours aux

exemples est fréquent et les présentations sont vivantes et attrayantes pour maximiser les gains des élèves (Brophy & Good, 1986 ; Lampert, 1988 ; Borich, 1996).

- Un *questionnement des élèves de qualité et avec des questions de haut niveau* : les enseignants efficaces posent beaucoup de questions et font participer les élèves aux discussions en classe. Ceci incite les élèves à rester impliqués dans la leçon et permet, dans le même temps, à l'enseignant d'évaluer leurs progrès et leur compréhension. Le niveau cognitif des questions doit être varié en fonction des compétences à maîtriser. La stratégie la plus judicieuse semble être d'utiliser un mélange de questions sur les « produits » (appelant une réponse unique des élèves) et les « processus » (appelant des explications de la part des élèves). Il a été démontré que les enseignants efficaces posent davantage de questions sur les processus (Brophy & Good, 1986 ; Evertson, Anderson, Anderson, & Brophy, 1980).
- Des *feedbacks* qui ont été identifiés comme étant un des nombreux facteurs d'efficacité ayant un effet sur la performance des élèves (Black & Wiliam, 1998 ; Scheerens & Bosker, 1997). La littérature a montré qu'un feedback efficace est fréquent (Crahay, 2006) et « instrumental » (lié à la tâche). Les corrections sont affectivement neutres. Ce n'est pas la personne de l'élève qui est l'objet du feedback, mais uniquement la validité de sa réponse (Bressoux, 1994). Par ailleurs, un feedback efficace doit être immédiat : les erreurs font partie du processus d'apprentissage et sont source de progrès à condition qu'elles conduisent à une correction immédiate afin d'éviter qu'elles ne deviennent systématiques. Selon Scheerens (2016), un feedback efficace comporte également des pistes pour combler les lacunes d'apprentissage. Enfin, les enseignantes et enseignants efficaces incitent aussi plus souvent et plus longtemps les élèves à chercher la réponse exacte et à reformuler une réponse après le signalement d'une erreur (Bressoux, 1994).
- Un *climat de classe positif* : le climat de classe renvoie au développement d'un environnement dans lequel l'enseignement et l'apprentissage peuvent avoir lieu. La recherche a montré l'importance du climat de la classe, non seulement pour la réussite des élèves, mais aussi pour l'estime de soi et la participation aux cours (Scheerens & Bosker, 1997 ; Muijs & Reynolds, 2000, 2002, 2017). Selon Muijs et al. (2014), les études sur l'efficacité de l'enseignement menées au cours des deux der-

nières décennies (p. ex. Košir, 2005 ; Rohrbeck, Ginsburg-Block, Fantuzzo & Miller, 2003 ; Slavin & Cooper, 1999) ont relevé l'importance d'étudier la contribution des enseignants à la création d'un environnement d'apprentissage dans leur classe en tenant compte des interactions enseignant-élèves, des interactions entre élèves, du traitement des élèves par l'enseignant, de la compétition et la collaboration entre élèves et du désordre dans la classe (Creemers & Kyriakides, 2008). Selon Hattie (2017), établir un climat de classe positif consiste essentiellement à créer un refuge où l'erreur est accueillie favorablement. Dans un tel cadre, les erreurs sont considérées comme des occasions de s'améliorer et les élèves se sentent à l'aise avec le fait de se tromper et de montrer leur ignorance. Le climat de classe intègre également la dimension des attentes de l'enseignant envers les élèves. Les enseignants efficaces ont des attentes élevées et considèrent que chaque élève est capable de réussir.

En ce qui concerne l'effet-établissement, il existe, selon Scheerens et Bosker (1997 ; cités par Van Damme et al., 2009), treize facteurs d'efficacité sur la performance des élèves dans les matières de base : *« les attentes des enseignants en termes de performance scolaire, un leadership éducatif fort, du consensus et de la cohésion au sein de l'équipe enseignante, un curriculum de grande qualité, d'importantes opportunités d'apprentissage, un climat d'école ordonné et structuré, une utilisation de l'évaluation à des fins d'apprentissage, un haut degré d'implication des parents dans l'école, un climat de classe favorable, un haut temps effectif d'apprentissage, un enseignement structuré, l'encouragement de l'apprentissage autonome, un feedback régulier et fort, et la différenciation de l'enseignement »* (p. 31). Les études de Teddlie et Reynolds (2000) ainsi que l'examen de la littérature réalisé par Marzano (2003, 2007) ont abouti à des résultats très similaires. Bien que les courants de recherche sur l'efficacité des classes et des établissements se soient historiquement développés en parallèle, ils ont produit des résultats qui coïncident fortement puisque les facteurs associés aux performances des classes sont en partie les mêmes que ceux associés aux performances des établissements. Comme l'a relevé Bressoux (1994), *« ceci traduit le fait*

*que les acquisitions des élèves s'opèrent fondamentalement au sein des classes et que ce qui est observé au niveau de l'école en ce domaine ne peut que refléter, au moins pour partie, ce qui se passe au niveau des classes »* (p. 113). Dans les recherches sur l'efficacité des établissements, l'unité d'analyse est l'école. La classe n'est alors qu'une sous-partie de cette entité dont le fonctionnement est affecté par ce qui se produit à un niveau plus macro. Selon Duru-Bellat (2003), l'idée principale est ici que *« le niveau organisationnel supérieur joue sur l'efficacité du niveau inférieur en facilitant les formes d'organisation et de fonctionnement qui s'avèrent les plus porteuses d'efficacité »* (p. 191). Il est important de relever que l'effet-établissement ne correspond pas à une simple agrégation d'effets-classes. L'établissement est une organisation sociale à part entière qui définit son propre ensemble de rôles, normes, évaluations et attentes vis-à-vis des élèves et développe son propre système de relations entre acteurs (Bressoux, 1994). L'effet-établissement est bien souvent modeste sur les acquisitions des élèves, mais il peut en revanche se révéler relativement fort sur d'autres dimensions, comme par exemple le moral des enseignants

et enseignantes, leur satisfaction au travail ou leur implication, qui sont autant de variables pouvant avoir une influence sur leurs pratiques (Addimando, 2019).

Enfin, au-delà des classes et des établissements, l'organisation d'un système éducatif est, elle aussi, suscep-

tible d'affecter les résultats d'apprentissage. Le contexte spécifique à chaque système scolaire conditionne l'efficacité de l'enseignement et certains systèmes éducatifs sont plus enclins que d'autres à faire émerger des effets contextuels relativement importants (Van Damme et al., 2009). À Genève, à l'issue de l'enseignement primaire, les élèves sont orientés dans trois filières distinctes sur la base de leurs résultats scolaires. L'avantage escompté de ce choix organisationnel est la création de conditions dans lesquelles les enseignants peuvent cibler efficacement l'enseignement en fonction des besoins des élèves pour parvenir à de meilleurs apprentissages (Dupriez, 2010). En d'autres termes, la structuration de l'enseignement et de l'apprentissage est organisée autour des capacités différenciées des

*L'effet-établissement est bien souvent modeste sur les acquisitions des élèves, mais il peut en revanche se révéler relativement fort sur d'autres dimensions, comme par exemple le moral des enseignants et enseignantes, leur satisfaction au travail ou leur implication, qui sont autant de variables pouvant avoir une influence sur leurs pratiques.*

élèves (Mazenod, 2021). Ce moyen de gérer l'hétérogénéité, retenu également dans plusieurs pays d'Europe tels que l'Allemagne, l'Autriche, la Hongrie et, à des degrés divers, dans la plupart des cantons suisses<sup>24</sup>, n'est pourtant pas exempt de critiques dans la littérature, notamment du point de vue de l'efficacité et de l'équité. En effet, la recherche a montré que dans ce type d'organisation, les élèves des filières les plus exigeantes ont tendance à accroître leurs avantages en termes de résultats par rapport à leurs camarades scolarisés dans les regroupements les moins exigeants (Gamoran, 2010). À Genève, l'évaluation de la dernière réforme de l'enseignement secondaire I (mise en place en 2011) a offert l'opportunité de vérifier qu'à performances scolaires comparables en fin de primaire la fréquentation d'un regroupement moins exigeant a des effets négatifs sur les compétences et les orientations des élèves (Petrucci, Fouquet-Chauprade, Charmillot, & Felouzis, 2021). Selon Evrard, Petrucci et Rastoldo (2019), « les élèves qui ont connu, après la réforme, une orientation à l'entrée du CO dans un regroupement moins exigeant que leurs homologues de même niveau scolaire avant la réforme, n'ont pas

bénéficié de cette nouvelle organisation (enseignement plus adapté à leur niveau scolaire et possibilités accrues de réorientations promotionnelles). Leurs résultats scolaires ont davantage divergé (en devenant un peu plus faibles), leurs orientations après le CO n'ont pas été plus ambitieuses, ni plus souvent dans une filière certifiante (pour les plus faibles scolairement), ni encore prélude à une entrée plus souvent linéaire au secondaire II » (p. 86).

Les effets des filières sont dans une large mesure déterminés par des différences d'enseignement (Gamoran, 2010 ; Van Houtte, 2004). Les filières peuvent notamment avoir une influence sur le type d'interactions entre enseignants et élèves (Watanabe, 2008), l'organisation des activités en classe (Forquin, 1984 ; Keddie, 1971), les opportunités d'apprentissage (Bressoux & Pansu, 2003 ; Grisy, 1993), le rythme des leçons (Gamoran, 2004) ou encore les attentes des enseignants

envers les élèves (Jamil, Larsen & Hamre, 2018 ; Rosenthal & Jacobson, 1968). De plus, l'assignation à un groupe de niveau faible peut également influencer sur un certain nombre de dimensions individuelles relatives aux attitudes et à la motivation face aux apprentissages scolaires. On relève, entre autres, des effets délétères sur le rapport à l'école, le sentiment d'auto-efficacité, le concept de soi, la motivation, l'absentéisme, la discipline ou encore, l'engagement dans le travail et les perspectives scolaires (Berends, 1995 ; Carbonaro, 2005 ; Cherkaoui, 1979 ; Müller & Hofmann, 2016).

Les filières constituent des contextes sociaux au sein desquels les élèves évaluent leurs propres performances, intériorisent les normes scolaires et apprennent à nourrir telle ou telle ambition concernant leurs performances à venir. L'univers scolaire est un univers de comparaison et dans ce cadre, les élèves visent des objectifs qui correspondent à la filière qu'elles

et ils fréquentent. L'assignation à un groupe « faible » fournit à l'élève des informations qui le contraignent à se situer dans la catégorie des « mauvais élèves », avec le risque que cela ne devienne une part définitive et non

modifiable de son identité. La théorie du *labeling* (ou processus d'étiquetage) stipule quant à elle que c'est la stigmatisation elle-même qui crée les comportements d'échec ou de déviance, en engendrant une dynamique d'attentes et d'auto-dévalorisation qui tend à s'entretenir d'elle-même.

Les résultats sur les effets-classes et les effets-établissements, tout comme ceux sur les effets-filières, ne sont pas universels et varient en fonction des pays ou régions. De nombreux effets des facteurs scolaires sont contingents et peuvent se produire ou non en fonction du contexte. Les travaux réalisés à ce jour par le SRED ont permis d'illustrer le fait qu'il existe à Genève, comme dans beaucoup d'autres systèmes éducatifs, un effet du contexte de scolarisation sur les acquisitions des élèves et plus précisément un effet de la classe et de l'établissement. Cet effet se révèle toutefois assez variable d'une discipline ou d'un domaine de compétences à l'autre et d'un degré à l'autre. Ainsi, la part des différences de résultats imputable à la classe oscille entre 2% et 20% et celle imputable à l'établissement varie

*Les filières constituent des contextes sociaux au sein desquels les élèves évaluent leurs propres performances, intériorisent les normes scolaires et apprennent à nourrir telle ou telle ambition concernant leurs performances à venir. L'univers scolaire est un univers de comparaison et dans ce cadre, les élèves visent des objectifs qui correspondent à la filière qu'elles et ils fréquentent.*

<sup>24</sup> Pour une présentation des différentes organisations du secondaire I, voir le rapport national sur l'éducation (CSRE, 2023).



de 0% (non significatif) à 7% (Petrucci, 2010 ; Soussi, Nidegger, Dutrévis, & Crahay, 2012 ; Petrucci, Ambrosetti, Crotta, & Nidegger, 2020 ; Ambrosetti, Petrucci, Fenaroli, & Egloff, 2022). Ces résultats confirment également qu'à Genève l'effet-classe est plus important que l'effet-établissement. On notera, par ailleurs, que le système éducatif cantonal semble assez peu propice à l'émergence d'effets-établissement. Selon Dumay, Dupriez et Maroy (2010), « *la ségrégation entre écoles risque davantage de produire des effets sur les élèves si les écoles bénéficient d'une forte autonomie pédagogique, susceptible d'entraîner une forme d'adaptation des écoles à leur public et une différenciation des objectifs en fonction des compétences (présumées) des élèves fréquentant un établissement* » (p. 479). À Genève, il n'existe pas de réelle possibilité de développement de la ségrégation entre écoles, car ces dernières n'ont pas la possibilité de sélectionner leurs élèves (seul le lieu de résidence détermine le lieu de scolarisation). De plus, il existe un cadre relativement contraignant en matière de plan d'études et de moyens d'enseignement qui contribue à rendre les contextes de scolarisation assez homogènes, et ce quand bien même il existe une certaine « liberté pédagogique » des enseignantes et enseignants.

Concernant la nature des effets de contexte, les travaux du SRED ont principalement porté sur les effets de composition (notamment sociale et académique) définis par Dumay et Dupriez (2009) comme l'impact des caractéristiques agrégées des élèves après que leur effet sur le plan individuel a été pris en compte. Dans ce cadre, quelques facteurs expliquant partiellement les différences de résultats imputables à la classe ou à l'établissement ont pu être identifiés (p. ex. la composition sociale du public accueilli dans la classe ; Petrucci, 2010), la participation des enseignants à la gestion de l'établissement ou la taille des structures (Ambrosetti et al., 2022), mais, là encore, ces facteurs diffèrent à la fois d'un domaine de compétence / d'une discipline à l'autre et d'un degré à l'autre rendant, à ce jour, la nature de l'effet-classe et de l'effet-établissement encore insaisissable dans le canton. Les facteurs globalement efficaces sont rares et la recherche « du bon maître » ou de « la bonne école » est perdue d'avance, mais il existe bien des régularités, des facteurs qui, dans certains contextes, sont plus efficaces que d'autres (Bresoux, 1994). Les données à disposition et les designs de recherche mis en œuvre à Genève n'ont, pour le moment, pas permis de dégager ces régularités.

En ce qui concerne l'effet des filières, la vérification de l'atteinte des compétences fondamentales définies au niveau suisse (enquête COFO) a permis d'identifier plusieurs aspects de l'engagement des élèves à l'égard de l'école et de l'apprentissage corrélés à l'atteinte de ces compétences de base : les attitudes positives envers l'école, l'auto-efficacité scolaire, le concept de soi académique, la motivation instrumentale ou encore l'absentéisme. Les effets de ces différents facteurs ont déjà été abondamment documentés dans la littérature et ne constituent, en eux-mêmes, ni une nouveauté ni une réelle surprise. En revanche, il est intéressant de relever qu'à Genève les profils des élèves de 11<sup>e</sup> année sur ces différentes dimensions apparaissent particulièrement contrastés en fonction des filières. Pour résumer, les élèves des filières les plus exigeantes (LS) font preuve, en moyenne, de plus d'engagement envers l'école et l'apprentissage que ne le font leurs camarades scolarisés dans les groupes les moins performants (CT et LC) (Petrucci & Roos, 2020).

Des travaux en cours devraient bientôt permettre d'étudier également l'effet-maître et l'effet des pratiques d'enseignement sur l'acquisition des compétences fondamentales en Mathématiques et en Français à l'intérieur du canton (Petrucci, Rastoldo & Roos, à paraître). Dans ce cadre, il sera possible de voir dans quelle mesure certaines dimensions de l'enseignement comme les opportunités d'apprentissage, les attentes envers les élèves ou le climat de classe sont impactées par l'existence des filières. Enfin, des informations collectées au niveau des établissements devraient permettre de tirer des enseignements concrets sur les modes de pilotage pédagogique locaux et leurs effets sur la réussite des élèves. Il s'agira notamment, en d'autres termes, de mieux comprendre comment ce qui se passe au niveau de l'établissement affecte les processus qui prennent place au sein des classes.



## Les coulisses de l'évaluation

### Du point de vue des disciplines enseignées : des évaluations de quelles compétences ?

En 2015, Delevay (p. 11) relevait que « [d]ans le monde scolaire, un apparent consensus semblait jusqu'alors exister autour du "quoi enseigner", alors que le "comment enseigner" faisait davantage question, et ce depuis longtemps ». Force est de constater que ces dernières années, se profile la disparition de certaines disciplines jadis réputées, alors que d'autres savoirs, connaissances ou compétences dites d'avenir « frappent à la porte » des salles de classe et des plans d'études, et pour certains et certaines d'entre elles y sont déjà entrées.

Ainsi, et en adéquation avec le PER, les élèves genevois découvrent désormais l'anglais dès la 7P (Soussi & Brüderlin, 2018). L'étude du fait religieux a également été introduite au programme d'histoire cantonal dès 2018, et ce, dès le cycle élémentaire. Ces deux exemples traduisent, à des degrés divers, l'harmonisation intercantonale de l'offre scolaire (HarmoS, 2011). Plus récemment encore, et plus ou moins installé selon les degrés de scolarité, l'enseignement des compétences numériques a fait son entrée dans le PER. Plus exactement, l'éducation numérique, à laquelle les élèves étaient exposés au travers des MITIC<sup>25</sup>, est amenée à devenir un domaine disciplinaire à part entière, au même titre que les Langues, les Mathématiques et les Sciences de la nature, les Sciences humaines et sociales, les Arts et le domaine Corps et mouvement. L'implémentation de l'éducation numérique est perçue comme une visée prioritaire tant dans l'enseignement primaire que dans le secondaire I et le secondaire II (CIIP, 2018 ; Educa, 2021).

La question de l'enseignement des compétences transversales se pose également avec de plus en plus d'acuité (Dutrévis & Rastoldo, 2020 ; Gaussel, 2018). Si elles sont présentes depuis le début dans le PER par exemple, elles semblent traverser la scolarité des élèves sans faire l'objet d'un enseignement ou d'un travail explicite avec les élèves. Pour autant, une même tendance à la revalorisation de ces

compétences s'observe tant dans le domaine scolaire que dans le domaine professionnel (CIIP, 2022 ; OCDE, 2016 ; SEFRI, 2017).

Comme on le voit, de nouvelles compétences émergent et d'autres compétences semblent bénéficier d'une revalorisation. À l'école, plusieurs indices renseignent sur la classification et le statut des savoirs enseignés. Citons notamment le nombre d'heures par semaine consacrées à telle ou telle discipline, les évaluations ou encore le rôle joué dans la promotion ou l'orientation des élèves. Selon ces trois indices, la hiérarchie des domaines d'apprentissage reste très stable dans le temps. Langue de scolarisation et Mathématiques (ou littératie ou numératie) restent en effet les disciplines (ou les compétences) les plus enseignées, les plus évaluées et les plus influentes dans le parcours scolaire des élèves (OCDE, 2021 ; PER, 2011 ; REP, 1993 ; RCO, 2010). À titre d'exemple, dans le cadre des *apprentissages dans la vie scolaire*, l'autonomie des élèves ainsi que la collaboration sont des compétences qui font l'objet d'appréciations tout au long du degré primaire, mais sans avoir pour autant une réelle incidence sur le parcours scolaire. À ce stade, les compétences transversales ne sont que rarement évaluées pour elles-mêmes dans la plupart des systèmes et des tests internationaux. Au mieux, elles sont prises en considération à titre d'éléments contextuels (Petrucci & Roos, 2020).

En revenant à Delevay (2015, p. 11), si « *le champ du "quoi enseigner", depuis Durkheim, renvoyait à une trilogie faisant de l'école le lieu de formation du citoyen, de la personne, et du travailleur* », les injonctions faites à l'école ne sont donc pas forcément différentes aujourd'hui, mais on attendra des citoyens, des personnes et des travailleurs de demain d'autres expertises – qu'elles soient en plus ou complémentaires – de celles acquises durant les décennies précédentes.

### Évaluation des compétences des élèves : regard critique

À ce jour se dessine un éventail toujours plus englobant de dispositifs, d'outils et de pratiques d'évaluation qui touchent à tous les domaines du système éducatif. Qu'il s'agisse de l'évaluation d'une enseignante ou d'un enseignant (dans une visée de développement professionnel), de l'efficacité d'une innovation pédagogique, ou encore de la performance d'un système éducatif, le but est d'insuffler une véri-

<sup>25</sup> Médias, Images, Technologies de l'Information et de la Communication – Inscrit dans la Formation générale du PER.

table culture de l'évaluation pour, *in fine*, améliorer la qualité de l'éducation.

L'évaluation des apprentissages scolaires se présente sous différentes formes et se résume généralement à trois grandes catégories selon leur finalité (Bonora, 1996 ; Muijs & Reynolds, 2018). Premièrement, l'*évaluation sommative* ou *certificative* est censée donner une image de la performance d'une ou d'un élève (ou d'un groupe d'élèves) sur une période donnée par rapport à un ensemble d'objectifs d'apprentissage dans une matière particulière et par ce biais, certifier des acquis. Elle remplit ainsi une fonction administrative ou sociale. À noter que le recours à la notation, bien qu'omniprésent dans les systèmes éducatifs, n'est pas un impératif de ce type d'évaluation. Un deuxième type concerne l'*évaluation pronostique*. Comme la précédente, elle sert à faire un bilan des compétences et des connaissances, mais plutôt en amont des apprentissages. Elle permet notamment de fonder des décisions d'orientation ou d'admission des élèves en fonction de leur aptitude présumée, à l'exemple de la transition entre l'école primaire et le cycle d'orientation. L'*évaluation formative*, quant à elle, a une fonction pédagogique. Elle est conçue pour informer l'enseignante ou l'enseignant sur les performances, les connaissances et les compétences antérieures de ses élèves (évaluation diagnostique), et ces informations sont ensuite utilisées pour planifier des leçons ou des mesures correctives (remédiation) afin d'améliorer les performances des élèves.

Ces trois catégories d'évaluation vont accompagner les élèves tout au long de leur parcours scolaire, avec des apparences, des intervalles et des finalités différentes. Et pourtant, souvent encore, l'évaluation des apprentissages des élèves est considérée comme un processus qui consiste à attribuer une note. Cette conception restrictive de l'évaluation reste ancrée dans les représentations (tant des parents que des professionnels de l'éducation) et les pratiques professionnelles, malgré l'abondance de travaux qui contestent le postulat de l'« objectivité » de la notation, les enjeux particulièrement élevés en termes de décisions de passage ou d'orientation, et les effets potentiellement néfastes sur le bien-être des enfants en contexte scolaire.

### Le jugement sur la valeur scolaire des élèves ou les biais de la notation

Les spécialistes en évaluation scolaire ont depuis longtemps alerté sur les sources principales d'erreur ou de biais qui entachent la notation par les enseignantes et enseignants (Leclercq, Nicaise & Demeuse, 2004). Une première concerne l'influence du *contexte scolaire* sur le processus d'évaluation. Ainsi, la classe dans laquelle se trouvent les élèves n'est pas sans conséquence. L'« effet Posthumus » (Grisay, 1984), par exemple, décrit la tendance des enseignantes et enseignants à établir, voire à exacerber des différences de performance entre les élèves d'une classe, ce qui les amène à produire systématiquement une distribution gaussienne des résultats, avec toujours quelques élèves « faibles », une bonne partie des élèves « dans la moyenne » et quelques élèves « forts ». Dans la même veine, la « *constante macabre* » (Antibi, 2003) désigne le phénomène d'une proportion assez stable de « mauvaises notes » liée à la manière d'évaluer les élèves et ce, quel que soit le sujet, la qualité de l'enseignement ou le véritable niveau de la classe. Dans les deux cas, la note attribuée à une ou un élève particu-

*La note attribuée à une ou un élève particulier ne dépend pas uniquement de sa propre performance dans l'absolu, mais aussi de sa performance par rapport à celles des autres élèves de sa classe.*

lier ne dépend pas uniquement de sa propre performance dans l'absolu, mais aussi de sa performance par rapport à celles des autres élèves de sa classe. Notre système d'évaluation semble moins enclin à évaluer réellement la réussite des apprentissages qu'à établir un classement des élèves au sein duquel certaines et certains sont inévitablement destinés à l'échec. Bien que cet effet de microcosme puisse aussi se faire ressentir au niveau de l'établissement scolaire (par exemple, lorsqu'un établissement veut se bâtir une réputation d'excellence en pratiquant une sélection par « écrémage »), il est surtout à l'œuvre dans les classes (Bressoux, 1994). À Genève, on estime qu'environ 7% des différences de notes en Mathématiques et en Français, en fin de primaire, sont liées à la classe (Zanolla & Petrucci, *en préparation*). La représentation qu'une enseignante ou un enseignant se construit du niveau scolaire d'une ou d'un élève s'effectue en réalité sur la base d'un étalon propre à sa classe et non sur un étalon propre à toutes les classes. Plus le public accueilli au sein d'une classe est performant et plus le jugement de l'enseignant est

sévère (Bressoux & Pansu, 2003 ; Dompnier, Pansu & Bressoux, 2011).

Aussi banal et aléatoire qu'il puisse paraître, l'ordre de correction des « copies » peut également induire des biais de notation. Ainsi, une même copie peut être notée différemment selon la qualité de la copie qui la précède (effet de contraste). De même, elle risque d'être différemment notée selon qu'elle se trouve parmi les premières ou les dernières copies à corriger (Leclercq, Nicaise & Demeuse, 2004). Dans une étude récente, on a également relevé que les enseignantes et enseignants plus expérimentés ont eu tendance à délivrer, « toutes choses égales par ailleurs », des notes plus faibles à leurs élèves (Zanolla & Petrucci, en préparation).

Les notes ont un impact considérable sur la scolarité des élèves en termes de décisions de passage ou d'orientation. Elles sont également susceptibles d'avoir un effet sur l'investissement des élèves dans leur travail scolaire, en les motivant ou en les démotivant (Madon, Jussim & Eccles, 1997) et peuvent influencer sur les décisions des familles d'investir plus ou moins dans la scolarité de leur enfant (Lievore & Triventi, 2022). Pour toutes ces raisons, de nombreux travaux de recherche se sont intéressés à l'attribution des notes. La note est le résultat d'un processus complexe qui n'est ni la simple application d'un barème de correction ni une simple activité de « mesurage » des performances effectives (Chevallard, 1986). Elle est, en réalité, l'expression d'un *jugement scolaire* qui va bien au-delà de la seule réalité objective des performances des élèves (Dompnier, Pansu & Bressoux, 2011). Les connaissances « réelles » de ces derniers apportent bien une part importante d'explication des différences de notes, mais d'autres caractéristiques du contexte, des enseignants et des élèves eux-mêmes ont aussi une influence (Duru-Bellat & Mingat, 1988). Pour le dire autrement, la note est le fruit d'un processus de fabrication entaché de « biais » (Leclercq, Nicaise & Demeuse, 2004).

L'étude des différences de « traitement » réservé aux élèves en matière de notation peut se faire au moyen d'un rapprochement systématique entre les moyennes annuelles délivrées par les enseignants et les résultats à des

tests standardisés, comme les épreuves cantonales, pour une même année de scolarité<sup>26</sup>. À Genève, à la fin de l'enseignement primaire<sup>27</sup>, on peut observer que les connaissances « réelles » des élèves (c.-à-d. les résultats à l'épreuve cantonale) constituent bien la principale explication des différences des moyennes annuelles délivrées par l'enseignant<sup>28</sup> en Mathématiques et en Français. Plus le score à l'épreuve cantonale est élevé et plus la moyenne annuelle délivrée par l'enseignant est élevée et inversement. Les jugements des enseignants sont donc globalement assez valides même s'il existe une grande variabilité d'un enseignant à l'autre.

Des « biais de notation » existent cependant dans le canton comme ailleurs.

De façon plus précise, à niveau de connaissances « réelles » comparable, le fait d'être en retard dans la scolarité est une caractéristique individuelle qui tend à faire

systématiquement baisser la note délivrée par l'enseignante ou l'enseignant. Ceci s'explique, en partie, par des effets d'« étiquetage » (ou *labelling* en anglais), les enseignants ayant tendance à incorporer des informations ou des indices sur les élèves dans l'élaboration de leur jugement (McMillan & Nash, 2000). Pour le cas spécifique des redoublants, certains travaux ont montré qu'ils sont jugés plus sévèrement que les non-redoublants en raison du fait qu'ils ont suivi le programme de l'année (au moins) deux fois. Les enseignants sont alors plus exigeants à leur égard.

On relève aussi, à Genève, un effet négatif associé à la profession faiblement qualifiée des parents : les élèves de milieu populaire sont, à niveau de connaissance identique, plutôt moins bien notés que leurs homologues issus de milieux plus favorisés. Ce résultat, déjà observé dans de nombreux contextes, est notamment lié à l'existence de croyances so-

*La note est le résultat d'un processus complexe qui n'est ni la simple application d'un barème de correction ni une simple activité de « mesurage » des performances effectives (...). Elle est, en réalité, l'expression d'un jugement scolaire qui va bien au-delà de la seule réalité objective des performances des élèves.*

<sup>26</sup> Les connaissances « réelles » des élèves sont mesurées par les résultats aux tests standardisés (p. ex. PISA, COFO, TAF, EC). Ces derniers constituent un critère externe de mesure de la valeur scolaire des élèves couramment utilisé dans les travaux sur les « biais » de la notation.

<sup>27</sup> Les analyses portent sur les années scolaires 2015-16 et 2021-22.

<sup>28</sup> Entre 75% et 80% des différences de notes entre élèves à l'intérieur des classes sont expliquées par les performances des élèves aux épreuves cantonales.

ciales négatives à l'égard des élèves les plus modestes (Duru-Bellat & Mingat, 1993 ; Duru-Bellat, 2015 ; cf. section sur l'influence des caractéristiques individuelles). À l'inverse, les élèves issus de milieux favorisés sont notés plus favorablement par les enseignants en Mathématiques comme en Français. Bien qu'ils évoluent plus fréquemment dans des contextes (classe et établissement) de meilleur niveau scolaire qui incitent à une notation plus sévère, ces élèves savent aussi faire preuve de compétences particulières pour « négocier » leurs notes avec les enseignantes et enseignants. Merle (1996) parle d'« arrangements évaluatifs ». De façon générale, les élèves peuvent en effet jouer sur l'adéquation avec les normes sociales pour influencer favorablement les jugements de leurs enseignants, pour autant qu'ils disposent d'une connaissance adéquate de ces normes. Py et Somat (1991) parlent de « clairvoyance normative ». Plus que les autres, les élèves issus des milieux favorisés sont capables d'ajuster leurs attitudes, jugements et comportements aux normes sociales en vigueur pour obtenir de leurs enseignants des jugements scolaires plus favorables.

D'autres travaux de recherche ont montré que la conduite en classe de l'élève peut également influencer sur la notation. Lorsque le comportement de l'élève est perçu comme bon par l'enseignant, le jugement émis sur le niveau scolaire tend très souvent à être meilleur (Bennett, Gottesman, Rock, & Cerullo, 1993). À Genève, ce résultat se retrouve avec les appréciations portées par l'enseignante ou l'enseignant sur la progression de l'élève dans différentes dimensions des apprentissages de la vie scolaire. De façon plus précise, un élève qui fait preuve d'une progression peu satisfaisante dans la prise en charge de son travail personnel ou dans la collaboration avec ses camarades sera moins bien noté. Outre l'acquisition de compétences et connaissances scolaires, l'école vise aussi la transmission de principes ancrés dans le fonctionnement social au sens large. Les enseignantes et enseignants valorisent, par une plus-value sur la note, les élèves dont les comportements, attitudes ou jugements sont en accord avec les principes véhiculés par le système éducatif (Dompnier et al., 2011).

D'autres « biais de notation » sont plus spécifiquement liés à une discipline. Ainsi, à Genève,

à niveau de connaissances « réelles » identique, les filles sont systématiquement avantagées en Français et parfois désavantagées en Mathématiques. En Français, les élèves allophones semblent parfois soumis à un jugement plus sévère que leurs homologues dont la langue maternelle est le français. On peut également mettre en évidence un effet de halo qui correspond à la situation dans laquelle les performances des élèves dans une discipline donnée influencent le jugement des enseignants dans d'autres disciplines (à Genève, la performance en Français influe sur la note de Mathématiques et la performance en Mathématiques influe sur la note de Français).

Les facteurs qui influencent l'attribution des notes sont donc, comme nous venons de le voir, plus nombreux et complexes que le simple niveau « réel » de compétence des élèves. Il apparaît dès lors légitime de s'interroger sur l'ampleur des différents « biais » qui affectent la notation. Sur la base d'une mesure statistique standardisée<sup>29</sup>, qui permet de comparer la taille des effets pour des variables exprimées sur des échelles différentes, on constate que les « biais » sont modestes à Genève si on les considère indépendamment les uns des autres. Le calcul de notes « typiques » pour des profils théoriques d'élèves et de classes à partir des estimations fournies par les modèles statistiques vient conforter ce résultat. Considérés séparément les uns des autres, chacun des « biais » de notation que nous venons de présenter a un effet de taille modeste. Le cumul de plusieurs d'entre eux peut toutefois avoir un impact plus important sur la note délivrée par l'enseignant. En effet, à connaissances « réelles » identiques, le cumul des « biais » de notation peut représenter jusqu'à environ un point de différence selon le profil considéré. Il s'agit d'une variation relativement importante pour des moyennes annuelles qui, dans la quasi-totalité des cas, sont comprises entre trois et six points (et même plutôt entre quatre et six points).

Le caractère multidimensionnel des notes et en particulier leur composante non cognitive a été largement souligné dans les recherches antérieures (Bressoux & Pansu, 2003 ; Crotta, Salvisberg, & Tanolla, 2020) et les travaux du SRED indiquent que Genève ne fait pas excep-

*À connaissances « réelles » identiques, le cumul des « biais » de notation peut représenter jusqu'à environ un point de différence selon le profil considéré.*

<sup>29</sup> Le  $f^2$  de Cohen (Cohen, 1988).



tion à la règle (Zanolla & Petrucci, *en préparation*). Le jugement scolaire, dont la note est une expression, est un jugement de valeur teinté par un ensemble de facteurs qui dépassent la seule performance des élèves. Il s'élabore à partir d'éléments sociocognitifs, contextuels et sociétaux qui sont indépendants des élèves eux-mêmes. Ceci doit conduire à s'interroger sur la validité des notes (Brookhart, 2015; Cross & Frary, 1999) et devrait sans doute inciter à relativiser leur portée. Rendre les enseignants plus clairvoyants sur les facteurs qu'ils mobilisent dans la construction de leur jugement scolaire pourrait contribuer à réduire la subjectivité lors de l'évaluation des apprentissages. Pour limiter les biais de jugement, il est aussi possible de recourir à certaines procédures telles que l'anonymisation des copies (comme vient de le mettre en place l'Université de Genève) ou la correction par des enseignantes et enseignants autres que celui ou celle de l'élève évalué (Leclercq et al., 2004; Merle, 2012). Par ailleurs, le processus de raisonnement des enseignants concernant leurs décisions de notation a été encore peu exploré. Conduire des recherches permettant de « pénétrer dans la tête » des enseignants pour découvrir ce qui influence leur prise de décision concernant les pratiques d'évaluation et de notation serait sans doute particulièrement intéressant et utile (McMillan & Nash, 2000). En Suisse, certains cantons comme le Tessin sont en train de s'engager dans cette voie.

### **Vers une évaluation pour l'apprentissage ?**

L'évaluation des apprentissages des élèves est l'une des principales parties du travail des enseignantes et enseignants et occupe une part importante de leur temps.

Encore trop souvent, l'attribution de notes reste la modalité principale d'évaluation des élèves alors même que, à Genève comme ailleurs, la fabrique de la note est perfectible à de nombreux égards puisqu'elle ne reflète que partiellement les performances objectives des élèves. Au-delà des problèmes de « biais » qui conduisent à s'interroger sur leur validité, les notes peuvent également avoir des effets négatifs sur les élèves en étant, par exemple, source d'anxiété ou de faible estime de soi.

Les jugements des enseignantes et enseignants genevois sur le niveau scolaire de leurs élèves sont globalement assez valides. Pour autant, il n'en demeure pas moins que la note reste le résultat d'un processus de fabrication susceptible de donner lieu à des biais, voire des distorsions. Les pratiques de notation ne reflètent pas seulement le niveau « objectif » de compétences des élèves, mais aussi la perception de l'effort, de la motivation et même du comportement. Les différences dans le « traitement » réservé aux élèves en matière de notation sont des possibles sources d'inégalité qui conduisent à se questionner sur les formes et rôles de l'évaluation au sein de notre système scolaire. La littérature a montré que, sous certaines formes, l'évaluation (qu'elle conduise ou non à l'attribution d'une note) peut apparaître comme un des facteurs

*Pour être efficace, l'évaluation au service de l'apprentissage doit être un élément central de la pratique en classe. Dans cette approche, l'objectif est d'intégrer l'évaluation dans la pratique pédagogique quotidienne et d'aider les élèves à évaluer où ils en sont, où ils doivent être et comment ils peuvent combler l'écart grâce à un retour d'information efficace.*

qui contribuent de façon significative à l'efficacité de l'enseignement (Scheerens, 2016). En effet, l'évaluation *formative* peut avoir un impact fort et positif sur les performances des élèves, en particulier si certaines stratégies de feedback sont mobilisées. C'est ce qu'on appelle l'évaluation au service de l'apprentissage

ou *assessment for learning* en anglais. Ce terme, initialement développé au Royaume-Uni (Black & William, 1998), a été utilisé pour désigner les pratiques évaluatives qui se sont avérées avoir un effet positif direct sur la réussite des élèves et les résultats d'apprentissage. Il s'agit principalement d'avoir recours à l'évaluation en classe pour améliorer les résultats des élèves. L'idée de base est qu'un élève progressera davantage s'il comprend le but de son apprentissage, où il se situe par rapport à ce but et comment il peut l'atteindre (ou combler l'écart dans ses connaissances) (QCDA, 2009).

De façon plus précise, pour être efficace, l'évaluation au service de l'apprentissage doit être un élément central de la pratique en classe. Dans cette approche, l'objectif est d'intégrer l'évaluation dans la pratique pédagogique quotidienne et d'aider les élèves à évaluer où ils en sont, où ils doivent être et comment ils peuvent combler l'écart grâce à un retour d'information efficace. L'évaluation au service de l'apprentissage doit être comprise comme faisant partie du processus d'enseignement, et pas seulement comme un complément à l'évaluation « traditionnelle ».



Elle correspond en réalité à une approche plus large de l'évaluation (Swaffield, 2011). Bien souvent, l'introduction de l'évaluation au service de l'apprentissage peut nécessiter un changement culturel dans les écoles et les systèmes éducatifs, en particulier là où l'évaluation a traditionnellement consisté en priorité à sélectionner et à répartir les élèves dans différentes filières, dans différents types d'écoles ou différents établissements d'enseignement supérieur (Kennedy, 2006). L'évaluation au service des apprentissages fait partie des objectifs du projet d'innovation pédagogique EP21, déployé dans les écoles primaires genevoises à la rentrée 2022. Il s'agit à la fois de trouver un meilleur équilibre entre enseignement et évaluation et d'inscrire les pratiques évaluatives dans le processus d'enseignement-apprentissage. Ce projet fait l'objet d'une évaluation accompagnante de la part du SRED.

Les principaux éléments de l'évaluation au service de l'apprentissage tels que décrits par Muijs et Reynolds (2017) sont les suivants.

▪ **Le partage des objectifs d'apprentissage.**

Les élèves doivent avoir connaissance des objectifs d'apprentissage d'une leçon ou d'un ensemble de leçons. Cela peut être facilité en les impliquant dans la définition même des objectifs d'apprentissage. Les critères d'évaluation doivent être clairs pour l'élève et doivent lui être communiqués. Les élèves doivent savoir ce que l'enseignante ou l'enseignant recherche dans un travail évalué ou une évaluation particulière, et sur quelle base les jugements seront portés. L'essentiel est de fournir aux élèves un retour d'information efficace sur leurs performances. Certains travaux ont par exemple démontré que les intentions d'apprentissage et les critères d'évaluation explicitement énoncés améliorent les résultats d'apprentissage des élèves (Hanover Research, 2014).

▪ **Un feedback centré sur le travail évalué.**

Donner un feedback formatif sur le travail évalué est un des éléments les plus importants de l'évaluation au service de l'apprentissage. Selon Black et William (1998), fournir un feedback écrit est beaucoup plus propice à l'apprentissage que la simple restitution d'une note. Les commentaires transmis à l'élève doivent fournir des suggestions d'amélioration, mais pas la solution complète. Le retour d'information doit permettre aux élèves de trouver eux-mêmes la bonne solution compte tenu de leur état de connaissances et de compétences. Les élèves doivent également être encouragés à trouver des solutions alterna-

tives, et un retour d'information doit être fourni au fil du temps sur les différentes étapes qu'elles et ils ont franchies pour trouver cette solution. Les commentaires doivent permettre aux élèves de savoir où ils en sont en termes de sujet, de tâche ou de connaissances sur lesquels ils travaillent, où ils doivent aller et comment ils peuvent combler tout écart entre les deux. Fournir des commentaires qui comparent les performances d'un élève à son propre niveau précédent plutôt qu'à celles d'autres élèves peut l'aider à voir son propre apprentissage et à développer des objectifs appropriés. Recourir à des commentaires écrits plutôt qu'à des notes atténue aussi les éventuels effets négatifs de la comparaison avec d'autres élèves. De plus, si l'évaluation devient une partie quotidienne de l'expérience et de l'apprentissage en classe, l'anxiété de performance qui accompagne les tests traditionnels pourrait diminuer.

▪ **L'utilisation de l'évaluation pour planifier l'apprentissage.**

L'évaluation au service de l'apprentissage ne concerne pas seulement ce que font les élèves, mais aussi la manière dont les enseignants et les enseignantes elles-mêmes utilisent l'évaluation. Ils et elles doivent utiliser l'évaluation comme un moyen de planifier les cours, de les adapter aux besoins des élèves et de reprendre l'enseignement là où les problèmes persistent. En d'autres termes, l'évaluation est un moyen de différencier l'enseignement. La fixation d'objectifs individuels, ambitieux et réalistes, basés sur des données d'évaluation est un élément particulièrement important. Le processus d'établissement d'objectifs doit comprendre à la fois des objectifs annuels à long terme et des objectifs à plus court terme basés sur les résultats de l'évaluation continue.

▪ **Le développement des compétences d'auto-évaluation des élèves et de l'évaluation par les pairs.**

S'auto-évaluer peut aider les élèves à devenir des apprenants plus auto-régulés et plus efficaces. Pour qu'ils puissent réaliser cette tâche, ils doivent disposer d'objectifs clairs, d'informations sur leur position par rapport à ces objectifs et d'une certaine compréhension de la manière dont ils peuvent combler l'écart entre leur situation et leurs objectifs. La notation par les pairs, où les élèves évaluent le travail des autres, est aussi un moyen de développer des compétences d'auto-évaluation qui a été utilisé avec succès dans les écoles (Noonan & Duncan, 2005 ; McCallum, 2000).

Bien entendu, comme pour toute autre intervention éducative, il convient de relever que l'évaluation au service de l'apprentissage ne fonctionne pas toujours pour tous les élèves. Toutes les études ne sont pas parvenues à identifier des effets positifs, bien souvent en raison de la très grande variabilité dans la mise en œuvre de ce type d'évaluation (Bennett, 2011). Pour que l'évaluation au service de l'apprentissage ait un impact, Wiliam (2011) a identifié deux conditions qui doivent être remplies. Premièrement, les élèves doivent recevoir des conseils sur comment s'améliorer. Et deuxièmement, elles et ils doivent mettre en œuvre et travailler ces conseils en utilisant le matériel fourni par l'enseignant, en allant vers l'enseignant pour obtenir de l'aide ou en travaillant avec d'autres élèves.

## Perspectives

### Les limites des différentes formes d'évaluation et de leur usage

En tant qu'instruments de monitoring, les enquêtes nationales COFO comme les TAF cantonaux, doivent aider à mieux comprendre ce qui se joue à l'intérieur du système éducatif genevois et à identifier des points de vigilance pour les autorités scolaires. Toutefois toutes les réponses ne se trouvent évidemment pas dans les données extraites de ces instruments de monitoring. Ces derniers ne renseignent que de façon assez « macro » sur le niveau de compétences de la population scolaire et doivent aussi conduire à la formulation de questions qui nécessiteront, à leur tour, d'autres récoltes d'informations et d'autres investigations. Des approfondissements ultérieurs doivent découler de l'utilisation de ces instruments comme le préconisent d'ailleurs les perspectives esquissées dans la conclusion du rapport national COFO 2016 : « *d'autres sources de données pourraient être utiles pour rechercher les causes des différences de performance entre cantons. Par exemple, le rôle des structures scolaires ou la qualité de l'enseignement pourraient être mieux analysés à l'aide d'enquêtes auprès des directions d'écoles* » (Consortium COFO, 2019, p. 83). Pour aller dans ce sens, le SRED a, entre autres, mis sur pied en 2020-21 une recherche au niveau des classes pour identifier d'autres déterminants de la réussite scolaire et investiguer, de manière plus générale, le champ de l'effet-maître et des pratiques d'enseignement (Petrucci, Rastoldo & Roos, à paraître ; cf. *Les effets du contexte sur les compétences des élèves*).

### L'école obligatoire versus l'enseignement secondaire II

La réflexion menée dans ce texte s'est principalement concentrée sur l'enseignement obligatoire. Durant l'enseignement primaire et secondaire I, l'ensemble des élèves bénéficie d'un enseignement basé sur le même plan d'études – le PER – et doit donc développer ses compétences dans les mêmes domaines. Par ailleurs, c'est aussi sur cette période de scolarité obligatoire que les grandes enquêtes nationales (COFO) ou internationales (PISA<sup>30</sup>) se concentrent. L'évaluation des compétences

<sup>30</sup> Plus précisément, PISA est destiné aux élèves de 15 ans, ce qui peut parfois étendre le périmètre de l'enquête à des élèves en début de secondaire II.

au secondaire II ne fait en revanche pas l'objet d'une attention aussi forte, à la fois dans les travaux du SRED et à l'échelle internationale. Au niveau cantonal, les données rendent compte des taux de certification (Le Roy-Zen Ruffinen & Mouad, 2021a), de l'obtention des diplômes (Le Roy-Zen Ruffinen & Mouad, 2021b), et de la réussite aux examens (Mouad & Le Roy-Zen Ruffinen, 2021). Les indicateurs du décrochage scolaire quant à eux décrivent les parcours complexes (SRED, 2022). Ces différences de logique dans le regard sur les compétences des élèves entre enseignement obligatoire et enseignement secondaire II s'expliquent en grande partie par des logiques systémiques elles-mêmes spécifiques à chacun de ces degrés d'enseignement (formation commune versus filières de formation spécifiques, prescriptions cantonales versus ordonnances fédérales, absence versus présence de certification...). Pour autant, ce qui a été développé dans ce texte concernant l'évaluation, ses plus-values et ses limites pourrait également faire l'objet d'une attention particulière au secondaire II. Ainsi, les effets des caractéristiques individuelles et des stéréotypes associés (cf. par exemple le projet FNS réalisé par le SRED et l'UNIL sur les *Trajectoires de formation professionnelle au prisme du genre et de l'orientation sexuelle*, mené dans le secondaire II genevois), le rôle du contexte scolaire, les biais de jugement dans l'évaluation ou, à l'inverse, l'inscription de l'évaluation dans le processus d'enseignement-apprentissage constituent des objets d'étude et d'attention tout aussi pertinents dans la formation post-obligatoire que dans les premières étapes de scolarité des élèves.

### Les « privés » d'évaluation

L'évaluation fait partie intégrante de la vie scolaire des élèves et, ce, tout au long de la scolarité. Les évaluations permettent de rendre compte des apprentissages des élèves, elles servent à les orienter, et elles permettent aussi de monitorer le système scolaire. Pour autant, certains élèves, à certains moments de leur scolarité ou de manière plus pérenne, ne sont pas évalués, parce que dans certaines formations, on n'évalue pas afin de « protéger » les élèves. Ainsi, un parcours scolaire effectué dans l'enseignement spécialisé se fait sans attribution de notes et d'évaluation certificative. De même, les élèves scolarisés dans des dispositifs d'accueil ne reçoivent pas non plus de notes (Dutrévis & Brüderlin, 2019). En soi, cette logique de protection peut s'entendre : il s'agit de ne pas exposer les élèves à des risques d'échec trop élevés, avec des résultats

qui les bloqueraient dans leur parcours et qui affecteraient l'image qu'ils ont d'eux-mêmes. Mais comme montré précédemment, de nombreux élèves sont malgré tout exposés à ces risques. Par ailleurs, cette dimension protectrice montre à quel point l'évaluation, dans sa forme sommative, vient sanctionner les apprentissages (ou les lacunes) d'un ou d'une élève. Si l'on renverse le raisonnement et que l'on conçoit l'évaluation comme étant au service des apprentissages, alors certaines et certains élèves ne sont plus « protégés » de l'évaluation, mais « privés » de ces étapes qui font partie du processus d'enseignement-apprentissage tout comme du processus d'orientation des élèves d'ailleurs. Bien sûr, le fait de ne pas avoir de notes ne signifie pas une absence totale d'évaluation, mais le discours autour de l'évaluation est plutôt celui de préserver les élèves et d'ajuster les contenus d'enseignement et l'évaluation des progressions sur la base de ses compétences initiales ou de ce qu'on envisage pour lui dans le futur (Jendoubi, Guilley, & Dutrévis, 2019). Le processus dans son ensemble ne s'ancre finalement que peu sur un référentiel commun, même si le PER reste, dans les textes, le référentiel sur lequel baser l'enseignement. L'absence d'évaluations notées constitue également un enjeu pour le parcours de l'élève. Ne pas avoir de notes, c'est ne pas pouvoir envisager de transition entre une structure spécialisée et une classe régulière puisque les transitions et les orientations se font principalement sur la base des résultats scolaires des élèves. L'absence de notes ou d'évaluations qui formalisent les acquis des élèves sur la base d'un référentiel commun peut par ailleurs péjorer les transitions, c'est un point sur lequel nous revenons plus bas dans la partie sur les enjeux de communication.

À noter pour finir que du point de vue du monitoring du système, l'absence de regard global qui inclut aussi les élèves avec des besoins spécifiques constitue une limite puisqu'il est difficile de monitorer un système lorsqu'on en exclut de fait une partie de ces actrices et acteurs. Au niveau international, la tendance actuelle dans bon nombre de pays est, à l'inverse, d'adapter les items des tests externes à ces populations d'élèves afin que toutes et tous puissent être évalués (Flip+, 2023).

### L'évaluation comme enjeu de communication

À bien des égards, l'évaluation est un sujet sensible lorsqu'il est question d'en communiquer les contenus, ou plutôt les résultats.

Communiquer les résultats à un élève ou à sa famille, communiquer les résultats d'une classe à son enseignant ou son enseignante, communiquer le résultat des élèves d'un canton à l'ensemble d'un département, ou d'un canton... autant de sources de satisfaction ou de tension entre actrices et acteurs. Pourtant, l'évaluation constitue sans doute une des thématiques d'intérêt majeur pour l'ensemble de la population. Dans le rapport famille-école, cela constitue même sans doute la source de communication la plus fréquente de l'école vers les parents. Sans les développer sont rapidement posés ci-dessous quelques points d'attention qu'il semble particulièrement important de garder à l'esprit.

### On évalue ce qui compte vraiment

Comme évoqué plus haut dans ce texte, les évaluations sommatives et certificatives se concentrent particulièrement sur certaines disciplines. Le Français, les Mathématiques, les Sciences et les Langues étrangères sont les disciplines au cœur de la majorité des évaluations et des compétences retenues pour orienter les élèves. Le message de l'institution (ou des institutions compte tenu de la consistance de ce constat à travers les différentes formes d'évaluations et à travers le temps) est clair sur ce qui compte dans le parcours des élèves et ce qui est considéré comme des compétences et des savoirs essentiels à acquérir (cf. les travaux déjà anciens de Forquin, 1991 ; Lenoir, Larose, Grenon, & Hasni, 2000 ; Dutrévis & Toczek-Capelle, 2007). La question se pose alors des « autres compétences ».

### Une évaluation binaire : « Que fait mon enfant à l'école ? Il réussit ou il échoue »

L'évaluation et ses supports peuvent être considérés comme l'interface principale de la relation école-familles. Interface principale au sens où les documents de communication à destination des parents relèvent souvent des acquis de leurs enfants. L'école informe sur ce que l'élève sait ou ne sait pas et, indirectement, sur les contenus d'enseignement-apprentissage dont elle ou il a bénéficié. Le fait que cette communication occupe une place si centrale est sans nul doute attribuable au souhait réciproque de connaissances à ce sujet. En 2006, la population genevoise a d'ailleurs estimé nécessaire un retour aux notes après que celles-ci aient disparu pendant quelques années dans l'école primaire. Savoir où en est son enfant dans ses apprentissages, mais aussi (voire surtout) comment son enfant se situe par rapport aux autres élèves correspond donc

(ou tout du moins était-ce le cas en 2006) à un besoin pour une majorité de familles. Mais cela pose un certain nombre de questions sur ce que l'on donne (et obtient) réellement comme informations sur le quotidien scolaire des élèves via ces supports.

### Les résultats des élèves comme reflet de leurs compétences ou de celles de leurs enseignantes et enseignants ?

Évaluer ses élèves quand on est enseignante ou enseignant, c'est de fait prendre connaissance de ce qu'ils et elles ont acquis et, par extension, au moins en partie de comment on leur a enseigné. À l'échelle de sa propre classe – et sans revenir sur les limites de l'évaluation en classe –, l'information peut être perçue comme constructive et informative pour reprendre certains points, ajuster les séquences d'enseignement suivantes... bref, réguler sa pratique. Cette lecture constructive de l'évaluation peut être altérée lorsque l'évaluation solde la fin de son travail avec un ou une élève puisque les enseignantes et enseignants suivants vont pouvoir mettre en perspective les compétences observées de l'élève et les notes qu'il a obtenues dans le passé. Ils et elles peuvent possiblement juger d'une forme de laxisme ou d'autoritarisme selon le décalage observé. La littérature sur les effets du redoublement montre ainsi que cette pratique est en partie imputable à la peur du jugement de l'enseignant ou de l'enseignante à qui on « refile » un élève qui n'avait pas les compétences pour accéder à l'année de scolarité supérieure.

La même logique de jugement sur les pratiques enseignantes peut intervenir dans le cas des épreuves externes, où les résultats des élèves d'une classe pourraient être lus comme la réussite ou l'échec des pratiques de l'enseignant ou de l'enseignante. Dans ces deux situations, on voit comment la communication et la visibilisation en dehors de la classe peuvent être source de tensions pour le corps enseignant (et la même logique s'applique à l'échelle de l'établissement puis du canton).

### Toutes ces évaluations pour savoir quoi ?

C'est finalement la question que l'on peut se poser à la lecture de ce qui précède. Et c'est également une question qui émerge parfois dans les différents travaux menés. Les évaluations sont soumises à un certain nombre de biais, elles peuvent générer de l'anxiété et du stress, ou à l'inverse susciter peu d'implications de la part des élèves, voire



d'autres actrices et acteurs lorsqu'elles n'ont pas de valeur sommative ou certificative.

Plusieurs études (p. ex. Rozenwajn & Dumay, 2014 ; Yerli, 2017) se sont à cet égard penchées sur la question de l'impact des évaluations externes sur les pratiques enseignantes. De manière générale, il en ressort que la contribution de ces évaluations à l'innovation pédagogique et à l'amélioration de la qualité de l'enseignement est encore trop limitée. La plupart des enseignantes et enseignants (surtout les plus anciens) semblent toujours faire plus confiance à leur propre évaluation et connaissance des élèves. Elles et ils ne se considèrent souvent pas les premiers destinataires des résultats, et regrettent le manque de feedback et, s'il y en a, un retour trop peu détaillé pour être opérationnel. Au mieux, les évaluations externes parviendraient alors, au travers de l'alignement de l'enseignement et du matériel pédagogique sur le contenu des tests, au couplage entre exigences institutionnelles et pratiques locales.

Or, lorsque l'on considère à la fois l'évaluation comme au service de l'apprentissage et comme faisant partie intégrante du processus d'enseignement-apprentissage, alors ces arguments perdent de l'importance. L'évaluation vient renseigner l'élève sur où elle ou il en est dans ses apprentissages, sur ses acquis, mais aussi sur ce qu'il lui reste à travailler. À nouveau, la symétrie s'opère avec le corps enseignant et, au-delà, le système. L'évaluation constitue un outil, une information au service de chacune et chacun. Mais cela implique une communication sur les résultats des évaluations, qui d'une part soit suffisamment élaborée pour renseigner les destinataires dans leurs pratiques et d'autre part s'inscrive dans une logique de progression (pour l'élève, l'enseignant et le système) plus que de sanction. ■

## Bibliographie

### Travaux du SRED...

- Ambrosetti, A., Petrucci, F., Fenaroli, S., & Egloff, M. (2022). The school effect on mathematics performance in PISA 2012: A comparison between two cantons in Switzerland. *Swiss journal of educational research*, 44(1), 90-102.
- Cecchini, A., Dutrévis, M., Guilley, E., & Jaunin, A. (2023). « De la petite enfance à la première rentrée scolaire. Portraits et enjeux de transition ». *Éclairages*, n° 3. Genève : SRED.
- Consortium COFO (éd.) (2019a). *Vérification de l'atteinte des compétences fondamentales. Rapport national COFO 2016 : mathématiques 1<sup>re</sup> année scolaire*. Berne et Genève : CDIP et SRED.
- Consortium COFO (éd.) (2019b). *Vérification de l'atteinte des compétences fondamentales. Rapport national COFO 2017 : langues 8<sup>e</sup> année scolaire*. Berne et Genève : CDIP et SRED.
- Consortium PISA.ch (2019). *PISA 2018 : Les élèves de Suisse en comparaison internationale*. Berne et Genève : SEFRI/CDIP et Consortium PISA.ch.
- Consortium romand PISA (éd.) (2001). *Compétences des jeunes romands. Résultats de l'enquête PISA 2000 auprès des élèves de 9<sup>e</sup> année*. Neuchâtel et Genève : IRDP et SRED.
- Denecker, C., Petrucci, F., & Prosperi, O. (2023a). Compétences des élèves en fin de 4P. *Repères et indicateurs statistiques F1*, n° 128. Genève : SRED.
- Denecker, C., Petrucci, F., & Prosperi, O. (2023b). Compétences des élèves en fin de 8P. *Repères et indicateurs statistiques F2*, n° 129. Genève : SRED.
- Dutrévis, M., & Brüderlin, M. (2019). *Étude sur les nouveaux dispositifs d'accueil pour élèves primo-arrivants allophones*. Genève : SRED.
- Dutrévis, M., & Rastoldo, F. (2020). *Compétences transversales: les bénéfices de sortir de l'implicite*. Genève : SRED.
- Dutrévis, M., Soussi, A., & Genoud, P. (2017). *Les attitudes et aspirations scientifiques des filles et des garçons à Genève. Enquête auprès des élèves du primaire et du cycle d'orientation*. Genève : SRED.
- Dutrévis, M., Wettstein, J., & Spagnolo, T. (2024). *Évaluation du projet d'innovation pédagogique EP21*. Genève : SRED.
- Evrard, A., Petrucci, F., & Rastoldo, F. (2019). *Les effets de la réforme du cycle d'orientation sur les parcours de formation des élèves*. Genève : SRED.
- Jendoubi, V., Guilley, E., & Dutrévis, M. (2019). *Évaluation des classes intégrées (CLI) au cycle d'orientation*. Genève : SRED.
- Le Roy-Zen Ruffinen, O., & Mouad, R. (2021a). Première et deuxième certification de niveau secondaire II. *Repères et indicateurs statistiques G1*, n° 102. Genève : SRED.
- Le Roy-Zen Ruffinen, O., & Mouad, R. (2021b). Type de diplôme secondaire II et âge à l'obtention. *Repères et indicateurs statistiques G2*, n° 103. Genève : SRED.
- Mouad, R., & Le Roy-Zen Ruffinen, O. (2021). Taux de réussite aux examens (secondaire II). *Repères et indicateurs statistiques G3*, n° 104. Genève : SRED.
- Petrucci, F. (2010). Compétences en lecture à Genève : influence des caractéristiques des élèves et des classes à différents moments de la scolarité obligatoire. *Mesure et évaluation en éducation*, 33(1), 79-106.
- Petrucci, F. (2020). Compétences des élèves en fin de 11<sup>e</sup>. *Repères et indicateurs statistiques F4*, n° 98. Genève : SRED.
- Petrucci, F. (2023). Compétences des élèves en fin de 11<sup>e</sup>. *Repères et indicateurs statistiques F4*, n° 130. Genève : SRED.
- Petrucci, F., & Roos, E. (2020). *Atteinte des compétences fondamentales dans le canton de Genève. Que nous enseignent les enquêtes COFO 2016 et 2017 ?* Genève : SRED.
- Petrucci, F., Ambrosetti, A., Crotta, F., & Nidegger, C. (2020). Impact des établissements sur les performances des élèves. Une étude basée sur les résultats de l'enquête COFO 2016. *Schweizerische Zeitschrift für Bildungswissenschaften*, 42(2), 305-322.
- Petrucci, F., Fouquet-Chauprade, B., Charmillot, S., & Felouzis, G. (2021). Tracking effects on achievement and opportunities of middle-high ability students: a case study in Switzerland. *School Effectiveness and School Improvement*, 33(1), 104-124.
- Rastoldo, F., & Mouad, R. (2021). Transitions vers l'enseignement secondaire II. *Repères et indicateurs statistiques D4*, n° 108. Genève : SRED.
- Rastoldo, F., & Mouad, R. (2022). Transitions à l'intérieur de l'enseignement secondaire I.

*Repères et indicateurs statistiques D3, n° 127.* Genève : SRED.

Soussi, A., Nidegger, C., Dutrévis, M., & Crahay, M. (2012). Un réseau d'enseignement prioritaire dans le canton de Genève : quels effets sur les élèves ? *Revue française de pédagogie*, 178, 53-66.

SRED (2022). *Indicateurs du décrochage de la formation.* Genève : SRED.

### ... et d'ailleurs

Addimando, L. (2019). *Le pratiche didattiche nella Scuola dell'obbligo.* Locarno : CIRSE.

Antibi, A. (2003). *La constante macabre.* Paris : Nathan.

Bennet, R.E. (2011). Formative assessment: a critical review. *Assessment in education: principles, policy and practice*, 18(1), 5-25.

Bennett, R., Gottesman, R., Rock, D., & Cerullo, F. (1993). Influence of behavior perceptions and gender on teacher's judgments of students' academic skill, *Journal of educational psychology*, 85(2), 347-356.

Berends, M. (1995). Educational stratification and students' social bonding to school. *British journal of sociology of education*, 16(3), 327-351.

Bianco, M., & Bressoux, P. (2009). Effet-classe et effet-maître dans l'enseignement primaire : vers un enseignement efficace de la compréhension ? Dans X. Dumay & V. Dupriez (Éds.), *L'efficacité dans l'enseignement. Promesses et zones d'ombre* (pp. 35-54). Bruxelles : De Boeck.

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in education: principles, policy & practice*, 5(1), 7-74.

Bonora, D. (1996). Les modalités de l'évaluation. *Revue internationale d'éducation de Sèvres [en ligne]*, 11.

Borich, G. (1996). *Effective teaching methods.* (3<sup>rd</sup> edition). New York: Macmillan.

Bressoux, P. (1994). Note de synthèse. Les recherches sur les effets-écoles et les effets-maîtres. *Revue française de pédagogie*, 108, 91-137.

Bressoux, P. (1995). Les effets du contexte scolaire sur les acquisitions des élèves : effet-école et effets-classes en lecture. *Revue française de sociologie*, 36(2), 273-294.

Bressoux, P., & Pansu, P. (2003). *Quand les enseignants jugent leurs élèves.* Paris : PUF.

Brookhart, S. (2015). Graded achievement, tested achievement, and validity. *Educational assessment*, 20(4), 268-296.

Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. Dans M.C. Wittrock (Éd.), *Handbook of research on teaching* (pp. 328-375). New York: MacMillan.

Carbonaro, W. (2005). Tracking, students' effort, and academic achievement. *Sociology of education*, 78(1), 27-49.

CCIP (2018). *La CIIP se donne un plan d'action en faveur de l'éducation numérique.* Communiqué de presse du 6 décembre 2018. Neuchâtel : CIIP.

CDIP (2011). *Compétences fondamentales pour les mathématiques. Standards nationaux de formation, adoptés par l'Assemblée plénière de la CDIP le 16 juin 2011.* Berne : CDIP.

CDIP (2021). *Utilisation des données dans le cadre de la vérification de l'atteinte des compétences fondamentales.* Berne : CDIP.

Cherkaoui, M. (1979). *Les paradoxes de la réussite scolaire.* Paris : PUF.

Chevallard, Y. (1986). Vers une analyse didactique des faits d'évaluation, in J.-M. De Ketele (éd.), *L'évaluation : approche descriptive ou prescriptive ?* (pp. 31-67). Bruxelles : De Boeck.

CIIP (2022). *Les capacités transversales. Bulletin CIIP N°6.* CIIP : Neuchâtel.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Coleman, J., Campbell, E., Hobson, C., Mcpartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity.* Washington, DC: GPO.

Crahay, M. (2006). *Un bilan des recherches processus-produit : l'enseignement peut-il contribuer à l'apprentissage des élèves et, si oui, comment ?* (Carnets des sciences de l'éducation). Genève : Université de Genève.

Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: a contribution to policy, practice and theory in contemporary schools.* London: Routledge.

Cross, L., & Frary, R. (1999). Hodgepodge grading: endorsed by students and teachers alike. *Applied measurement in education*, 12(1), 53-72.

Crotta, F., Salvisberg, M., & Zanolla, G. (2020). Note scolastiche e competenze fondamentali in matematica degli allievi dell'11 anno HarmoS in Ticino. *Schweizerische Zeitschrift für Bildungswissenschaften*, 42(2), 323-342.

- CSRE (2023). *L'éducation en Suisse – Rapport 2023*. Aarau : Centre suisse de coordination pour la recherche en éducation.
- Cvencek, D., Meltzoff, A.N., & Greenwald, A.G. (2011). Math-gender stereotypes in elementary school children. *Child Development, 82*, 766-779.
- Day, C., Sammons, P., Kington, A., Regan, E., Ko, J., Brown, E., Gunraj, J., & Robertson, D. (2008). *Effective classroom practice (ECP): a mixed-method study of influences and outcomes*. Report submitted to the Economic and Social Research Council (RES-000-23-1564). University of Nottingham.
- Delevay, M. (2015). *D'un programme de connaissances à un curriculum de compétences*. Bruxelles: De Boeck.
- Denessen, E., Hornstra, L., van den Bergh, L., & Bijlstra, G. (2022). Implicit measures of teachers' attitudes and stereotypes, and their effects on teacher practice and student outcomes: a review. *Learning and instruction, 1014-1037*.
- Directive du 25 août 2008 sur l'évaluation des compétences et des connaissances des élèves (EP) (D-E-DGEO-EP-SEE-12 ; état le 22 août 2022). Genève : SEE.
- Dompnier, B., Pansu, P., & Bressoux, P. (2011). L'évaluation scolaire : une activité multidéterminée. Dans F. Butera, C. Buchs & C. Darnon (Éds.), *L'évaluation, une menace ?* Paris : PUF.
- Doyle, W. (1986). Classroom organisation and management. Dans M.C. Wittrock (Éd.), *Handbook of research on teaching* (pp. 392-431). New York: MacMillan.
- Dumay, X., & Dupriez, V. (2009). Contexte d'établissement et apprentissage des élèves. Dans X. Dumay & V. Dupriez (Éds.), *L'efficacité dans l'enseignement. Promesses et zones d'ombre* (pp. 103-122). Bruxelles : De Boeck.
- Dumay, X., Dupriez, V., & Maroy, C. (2010). Ségrégation entre écoles, effets de la composition scolaire et inégalités de résultats. *Revue française de sociologie, 51(3)*, 461-480.
- Dupriez, V. (2010). *Séparer pour réussir ? Les modalités de groupement des élèves*. Paris : UNESCO.
- Durante, F., & Fiske, S.T. (2017). How social-class stereotypes maintain inequality. *Current Opinion in Psychology, 18*, 43-48.
- Duru-Bellat, M. (2003). Les apprentissages des élèves dans leur contexte : les effets de la composition de l'environnement scolaire. *Carrefours de l'éducation, 16*, 182-206.
- Duru-Bellat, M. (2015). *Les inégalités sociales à l'école : genèse et mythes*. Paris : PUF.
- Duru-Bellat, M., & Mingat, A. (1988). Le déroulement de la scolarité au collège : le contexte « fait des différences ». *Revue française de sociologie, 29(4)*, 649-666.
- Duru-Bellat, M., & Mingat, A. (1993). *Pour une approche analytique du fonctionnement du système éducatif*. Paris : PUF.
- Dutrévis, M., & Toczec-Capelle, M-C. (2007). Perception des disciplines scolaires et sexe des élèves : le cas des enseignants et des élèves de l'école primaire en France. *L'Orientation scolaire et professionnelle, 36(3)*, 379-400.
- Erzinger, A. B., Pham, G., Prospero, O., & Salvendy, M. (Éds.) (2023). *PISA 2022. La Suisse sous la loupe*. Université de Berne.
- Evertson, C., Anderson, C., Anderson, L., & Brophy, J. (1980). Relationships between classroom behaviors and student outcomes in junior high mathematics and English classes. *American educational research journal, 17(1)*, 43-60.
- Felouzis, G., Liot, F., & Perrotin, J. (2005). *L'apartheid scolaire. Enquête sur la ségrégation ethnique dans les collèges*. Paris : Seuil.
- Flip+ (2023, 8<sup>th</sup>-9<sup>th</sup> June). 6<sup>th</sup> FLIP+ Annual Event. Workshop *Inclusivity and SEND in digital assessment. Practice in supporting all learners through digitalization of assessment*. Cambridge, UK.
- Forquin, J.C. (1984). La sociologie du curriculum en Grande-Bretagne : une nouvelle approche des enjeux sociaux de la scolarisation. *Revue française de sociologie, 25(2)*, 211-232.
- Forquin, J-C. (1971). Savoirs scolaires, contraintes didactiques et enjeux sociaux. *Sociologie et sociétés, 23(1)*, 25-39.
- Foshay, A., Thorndike, R., Hotyat, F., Pidgeon, D., & Walker, D. (1962). *Educational achievements of thirteen-year-olds in twelve countries. Results of an international research project, 1959-1961*. Hamburg: UNESCO Institute for Education.
- Gamoran, A. (2004). Classroom organization and instructional quality. In M.C. Wang & H.J. Walberg (Eds.), *Can unlike students learn together? Grade retention, tracking, and grouping* (pp. 141-155). Greenwich, CT: Information Age Publishing.
- Gamoran, A. (2010). Tracking and inequality: new directions for research and practice. In M. Apple, S. Ball & L. Gandin (Eds.), *The Routledge international handbook of the sociology of education* (pp. 213-228). Abingdon: Taylor & Francis.



- Gaussel, M. (2018). *À l'école des compétences sociales*. Dossier de veille de l'IFé. Lyon : ENS de Lyon.
- Goos, M., Pipa, J., & Peixoto, F. (2021). Effectiveness of grade retention: a systematic review and meta-analysis. *Educational research review*, 34, 100401.
- Grisay, A. (1984). Les mirages de l'évaluation scolaire. Rendements en français, notes et échecs à l'école primaire ? *Revue de la direction générale de l'organisation des études*, XIX(5), 29-42.
- Grisay, A. (1993). *Le fonctionnement des collèges et ses effets sur les élèves de sixième et de cinquième*. (Les dossiers Éducation et Formations, n° 32). Paris : Ministère de l'éducation nationale.
- Hanover Research (2014). *The impact of formative assessment and learning intentions on student achievement*. Washington, DC: Hanover Research.
- Hastedt, D., & Rocher, T. (2020, septembre). Les évaluations internationales à grande échelle en éducation : un guide rapide. *IEA Compass : Les résumés en éducation*, 10.
- Hattie, J. (2017). *L'apprentissage visible pour les enseignants : connaître son impact pour maximiser le rendement des élèves*. Québec : PUQ.
- Jackson, G. (1975). The research evidence on the effects of grade retention. *Review of Educational Research*, 45(4), 613-635.
- Jamil, F., Larsen, R., & Hamre, B. (2018). Exploring longitudinal changes in teacher expectancy effects on children's mathematics achievement. *Journal for research in mathematics education*, 49(1), 57-90.
- Jimerson, S. (2001). A synthesis of grade retention research: Looking backward and moving forward. *California school psychologist*, 6, 47-59.
- Keddie, N. (1971). Classroom knowledge. In F.D. Young (Ed.), *Knowledge and control* (pp. 133-160). London: Collier Macmillan.
- Kennedy, K. (2006). *Barriers to innovative school practice: a socio-cultural framework for understanding assessment practices in asia*. Paper presented at the Redesigning pedagogy – Culture, understanding and practice Conference, Singapore.
- Košir, K. (2005). The influence of teacher's classroom management style on pupils' self-regulative behavior. *Studia psychologica*, 47, 119-143.
- Lampert, M. (1988). What can research on teacher education tell us about improving quality in mathematics education? *Teaching and teacher education*, 4(2), 157-170.
- Leclercq, D., Nicaise, J., & Demeuse, M. (2004). Docimologie critique : des difficultés de noter des copies et d'attribuer des notes aux élèves. Dans M. Demeuse (Éd.), *Introduction aux théories et aux méthodes de la mesure en sciences psychologiques et en sciences de l'éducation* (pp. 273-292). Liège : Université de Liège.
- Lenoir, Y., Larose, F., Grenon, V., & Hasni, A. (2000). La stratification des matières scolaires chez les enseignants du primaire au Québec : évolution ou stabilité des représentations depuis 1981. *Revue des sciences de l'éducation*, 26(3), 483-514.
- Leroy, N. (2009). *Impact du contexte scolaire sur la motivation et ses conséquences au plan des apprentissages* (Thèse de doctorat en Sciences de l'Éducation). Grenoble : Université Pierre-Mendès-France-Grenoble 2.
- Lievore, I. & Triventi, M. (2022). Do teacher and classroom characteristics affect the way in which girls and boys are graded? *British Journal of Sociology of Education*, 44(1), 97-122.
- Loi cantonale genevoise du 17 septembre 2015 sur l'instruction publique (LIP) (rsGE C 1 10 ; état le 30 avril 2022).
- Madon, S., Jussim, L., & Eccles, J. (1997). In search of the powerful self-fulfilling prophecy. *Journal of personality and social psychology*, 72(4), 791-809.
- Marzano, R. (2003). *What works in schools: translating research into action*. Alexandria, VA: ASCD.
- Marzano, R. (2007). Leadership and school reform factors. In T. Townsend (Ed.), *International handbook of school effectiveness and improvement* (pp. 597-614). New York: Springer.
- Mazenod, A. (2021). Classes de niveau : variations internationales dans les regroupements d'élèves et la constitution de classes au collège. *Revue française de pédagogie*, 212, 93-108.
- McCallum, B. (2000). *Formative assessment : implications for classroom practice*. London: Institute of Education.
- McMillan, J., & Nash, S. (2000). *Teacher classroom assessment and grading practices decision making*. Paper presented at the Annual Meeting of the National Council of Measurement in Education, New Orleans, LA.
- Merle, P. (1996). *L'évaluation des élèves. Enquête sur le jugement professoral*. Paris : PUF.

- Merle, P. (2012). L'évaluation par les notes : quelle fiabilité et quelles réformes ? *Regards croisés sur l'autonomie*, 12, 218-230.
- Muijs, D., & Reynolds, D. (2000). School effectiveness and teacher effectiveness in mathematics: some preliminary findings from the evaluation of the mathematics enhancement programme (primary). *School effectiveness and school improvement*, 11, 273-303.
- Muijs, D., & Reynolds, D. (2002). Teachers' beliefs and behaviors: what really matters? *The journal of classroom interaction*, 37(2), 3-15.
- Muijs, D., & Reynolds, D. (2018). *Effective teaching. Evidence and practice*. London: Sage.
- Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art. Teacher effectiveness and professional learning. *School effectiveness and school improvement*, 25(2), 231-256.
- Müller, C., & Hofmann, V. (2016). Does being assigned to a low school track negatively affect psychological adjustment? A longitudinal study in the first year of secondary school. *School effectiveness and school improvement*, 27(2), 95-115.
- Noonan, C. & Duncan, C. (2005). Peer- and self-assessment in high schools. *Practical assessment, research and evaluation*, 17(10), 12-20.
- OCDE (2013). *Cadre d'évaluation et d'analyse du cycle PISA 2012 : compétences en mathématiques, en compréhension de l'écrit, en sciences, en résolution de problèmes et en matières financières*. Paris : Éditions OCDE.
- OCDE (2016). *Cadre d'évaluation et d'analyse du cycle PISA 2015 : compétences en mathématiques, en compréhension de l'écrit, en sciences, en résolution de problèmes et en matières financières*. Paris : Éditions OCDE.
- OCDE (2019). *Country snapshots: Switzerland*. Paris : Éditions OCDE.
- OCDE (2019). *Résultats du PISA 2018 (Volume I) : savoirs et savoir-faire des élèves*. Paris : Éditions OCDE.
- Plante, I., de la Sablonnière, R., Aronson, J. M., & Théorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes : the role of competence beliefs and task values. *Contemporary educational psychology*, 38, 225-235.
- Py, J., & Somat, A. (1991). Normativité, conformité et clairvoyance : leurs effets sur le jugement éducatif dans un contexte scolaire. In J-L. Beauvois, R.V. Joule, & J.M. Monteil (Éds.), *Perspectives cognitives et conduites sociales*. Fribourg : Cousset.
- QCDA (2009). *Assessment for learning guidance*. London: The Stationery Office.
- Rayou, P., & Bautier, É. (2009). *Les inégalités d'apprentissage : programmes, pratiques et malentendus scolaires*. Paris : Presses universitaires de France.
- Règlement genevois du 7 juillet 1993 de l'enseignement primaire (REP) (rsGE C 1 10.21 ; état le 30 juin 2021).
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (EER): a state-of-the-art review. *School effectiveness and school improvement*, 25(2), 197-230.
- Rohrbeck, C., Ginsburg-Block, M., Fantuzzo, J., & Miller, T. (2003). Peer-assisted learning interventions with elementary school students: a meta-analytic review. *Journal of educational psychology*, 95, 240-257.
- Rosenshine, B. (1970). Evaluation of classroom instruction. *Review of educational research*, 40(2), 279-300.
- Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 376-391). New York: MacMillan.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart & Winston.
- Rozenwajn, E., & Dumay, X. (2014). Les effets de l'évaluation externe sur les pratiques enseignantes : une revue de la littérature. *Revue française de pédagogie*, 189, 105-138.
- Rutkowski, L., Von Davier, M., & Rutkowski, D. (2014). A brief introduction to modern international large-scale assessment. In L. Rutkowski, M. Von Davier & D. Rutkowski (Eds.), *Handbook of international large-scale assessment. Background, technical issues, and methods of data analysis*. Abingdon: Taylor & Francis.
- Scheerens, J. (2016). *Educational effectiveness and ineffectiveness: a critical review of the knowledge base*. Dordrecht: Springer.
- Scheerens, J., & Bosker, R. (1997). *The foundations of school effectiveness*. Oxford: Pergamon Press.
- SEFRI (2017). *Modèle de référence pour le plan de formation, modèle axé sur les compétences opérationnelles* (du 31 août 2012, état le 31 janvier 2018). Berne : SEFRI.
- Slavin, R., & Cooper, R. (1999). Improving intergroup relations: lessons learned from cooperative learning programs. *Journal of social issues*, 55, 647-663.

- Stallings, J. (1985). Effective elementary classroom practices. In M.J. Kyle (Ed.), *Reaching for excellence: an effective sourcebook* (pp. 14-42). Washington, DC: GOP.
- Steele, C.M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist*, 52(6), 613-629.
- Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in education: principles, policy and practice*, 18(4), 433-449.
- Talbot, L. (2012). Les recherches sur les pratiques enseignantes efficaces. Synthèse, limites et perspectives. *Questions vives*, 6(18), 129-140.
- Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research*. London: Falmer Press.
- Van Damme, J., Opdenakker, M., Van Landeghem, G., De Fraine, B., Pustjens, H., & Van de Gaer, E. (2009). Fondements et principaux résultats de recherche sur l'efficacité dans l'enseignement. Dans X. Dumay (Éd.), *L'efficacité dans l'enseignement : promesses et zones d'ombre* (pp. 17-34). Louvain-la-Neuve : De Boeck Supérieur.
- Van Houtte, M. (2004). Tracking effects on school achievement: a quantitative explanation in terms of the academic culture of school staff. *American journal of education*, 110, 354-388.
- Verner, M., & Helbling, L. (2019a). *Sampling ÜGK 2016. Technischer Bericht zu Stichprobendesign, Gewichtung und Varianzschätzung bei der Überprüfung des Erreichens der Grundkompetenzen 2016*. Zürich : IBE.
- Verner, M., & Helbling, L. (2019b). *Sampling ÜGK 2017. Technischer Bericht zu Stichprobendesign, Gewichtung und Varianzschätzung bei der Überprüfung des Erreichens der Grundkompetenzen 2017*. Zürich : IBE.
- Watanabe, M. (2008). Tracking in the era of high-stakes state accountability reform: case studies of classroom instruction in North Carolina. *Teachers college record*, 110, 489-533.
- Wiliam, D. (2011). What is assessment for learning? *Studies in educational evaluation*, 37(1), 3-14.
- Yerli, G. (2017). Les raisons du faible usage des résultats d'évaluation externe par les enseignants. Étude croisée dans trois contextes éducatifs. *Contextes et didactiques*, 9, 60-71.

*Éclairages constitue une série de publications du SRED, transversale à différentes recherches, évaluations et prestations unies par une problématique commune. Il s'agit de faire dialoguer les connaissances produites dans des contextes singuliers, avec des entrées plurielles et par différentes méthodes pour proposer une lecture nouvelle et synthétique de cette question. Les Éclairages proposent aussi des pistes de réflexion pour la pratique et un questionnement prospectif sur de possibles recherches à venir.*

Lien vers le document en pdf :  
<https://www.ge.ch/document/35538/telecharger>

