

VADÉMÉCUM L'ÉVALUATION EN COMMUN

Berne, 09.11.2020

Table des matières

1.	L'«ÉVALUATION EN COMMUN» DANS LE CONTEXTE SUISSE	3
	La notion d'«évaluation en commun»	3
	Recommandation de la Conférence des directeurs de l'instruction publique.....	3
	Directives et recommandations cantonales	4
2.	FORMES DE COOPÉRATION.....	4
3.	MODÈLES D'HARMONISATION DE L'ÉVALUATION EN COMMUN	5
	Degrés d'harmonisation selon Franz Baeriswyl	5
	Le modèle circulaire «Évaluation en commun» selon Katharina Maag Merki	6
	Défis	8
4.	OPPORTUNITÉS OFFERTES PAR L' «ÉVALUATION EN COMMUN».....	9
	Pour les élèves.....	9
	Pour les enseignants.....	9
5.	INFORMATIONS GÉNÉRALES SUR LE THÈME «ÉVALUATION»	9
	L'harmonisation des objectifs d'apprentissage et des types d'exercices lors des contrôles des connaissances	9
	Règles internes pour l'évaluation des performances et le système de notation.....	11
	Check-list sur le thème «Évaluer»	12
6.	GLOSSAIRE.....	13
7.	BIBLIOGRAPHIE	18

1. L'«ÉVALUATION EN COMMUN» DANS LE CONTEXTE SUISSE

Dans le cadre de l'évaluation de la maturité («EVAMAR II») conformément au [règlement sur la reconnaissance des certificats de maturité gymnasiale RRM 95](#), il a été constaté en 2008 qu'il existait, au sein même des gymnases et entre les écoles, de grandes disparités concernant les exigences de performances posées aux élèves. Le constat selon lequel les aptitudes des gymnasiennes et gymnasiens ne satisfont pas toujours aux exigences des universités dans les disciplines «langue première» et «mathématiques» a suscité l'inquiétude (Eberle/Brüggenbrock 2013: 53). La comparabilité et la transparence des évaluations des performances gymnasiales ont donné lieu à des discussions.

En conséquence, la Conférence des directeurs suisses de l'instruction publique (CDIP) a commandé en 2012 la réalisation de cinq sous-projets pour garantir à long terme l'accès sans examen aux hautes écoles universitaires pour les titulaires d'une maturité gymnasiale, un «cas particulier» suisse» (Eberle/Brüggenbrock 2013: 23/38). Le sous-projet 2, consacré au thème de l'«évaluation en commun», était l'un d'entre eux.

La notion d'«évaluation en commun»

Dans le rapport final sur le sous-projet 2, l'«évaluation en commun» désigne le processus d'évaluation des performances des élèves par un groupe d'enseignants. Il peut s'agir de parties d'un groupe de discipline ou du groupe de discipline complet d'une école. En général, le processus d'évaluation inclut plusieurs étapes, de la conception commune des épreuves à leur passage et leur correction. Dans le cadre de l'«évaluation en commun», toutes ou certaines étapes de ce processus font l'objet d'une responsabilité commune. Le niveau d'exigences, les domaines de compétences et/ou les domaines de contenus sont discutés. Les épreuves élaborées en commun répondent en outre aux critères de qualité que sont l'objectivité et l'équité. Dans la pratique, cela signifie que l'on veille tout spécialement à garantir une «préparation similaire durant les cours et une procédure identique lors du passage de l'épreuve» (Joachim/Heller-Andrist 2015: 1).

Recommandation de la Conférence des directeurs de l'instruction publique

Dans la décision du 17 mars 2016, la CDIP formule cinq recommandations pour garantir à long terme un accès sans examen aux hautes écoles universitaires pour les titulaires d'une maturité gymnasiale. Dans la recommandation 2 «Soutien des écoles dans l'élaboration de l'évaluation en commun», il est recommandé aux cantons de définir des conditions cadres pour des épreuves de maturité harmonisées (ce qui ne veut pas dire «uniformes») dans le canton et de soutenir les écoles dans la réalisation de projets de développement scolaire dédiés à l'«évaluation en commun».

La notion d'«évaluation en commun» signifie que les enseignants élaborent et utilisent en commun les épreuves et qu'ils les corrigent et les notent selon des critères définis en commun. L'évaluation en commun ne consiste donc pas à instituer un service central qui définirait les épreuves. (CDIP 2016: 5)

Directives et recommandations cantonales

Une harmonisation des épreuves écrites de maturité par établissement et par discipline est également la norme en Suisse. Certains cantons ont formulé des directives explicites; dans d'autres, la «maturité interne» harmonisée existe depuis de longue date, même si elle n'est pas formellement réglementée. Toutefois, les attentes et la pratique courante varient d'un canton à l'autre. Plusieurs cantons imposent des épreuves absolument identiques par discipline et par école; d'autres exigent seulement leur «comparabilité». Dans certains cantons, les commissions cantonales de maturité vérifient la comparabilité des épreuves de maturité élaborées au niveau de l'école.

Différents cantons ont entre-temps élaboré des concepts relatifs à l'«évaluation en commun» qui sont mis en œuvre localement sous forme de processus de développement scolaire. Les groupes de discipline sont tenus de réaliser des projets d'«évaluation en commun» (travaux de comparaison p. ex.). Une grande marge de manœuvre est toutefois accordée aux groupes de discipline dans la conception. Un tableau synoptique présentant la pratique et les instructions disponibles est mis en ligne sur le site Internet du ZEM CES ([ZEM CES: évaluation en commun](#)).

2. FORMES DE COOPÉRATION

La collaboration au sein d'un groupe de discipline représente un défi de l'«évaluation en commun». Sur la base de la psychologie organisationnelle, Cornelia Gräsel et al. (2006: 205-219) ont décrit, au niveau de l'élaboration des cours, les degrés de coopération suivants qui s'appliquent également lors de l'«évaluation en commun»:

Degré/désignation	Description	Conditions et conclusion
Échange collégial	S'informer réciproquement des contenus et situations professionnels et échanger des supports de cours.	<ul style="list-style-type: none"> ▪ Travail indépendant des personnes ▪ Différents supports ▪ Échange réciproque <p><i>Conclusion: suggestions non contraignantes</i></p>
Coopération fondée sur la division du travail	Division du travail entre des individus pour parvenir à un résultat planifié en commun et relevant de la responsabilité commune.	<ul style="list-style-type: none"> ▪ Travail indépendant des personnes ▪ Définition précise d'objectifs portés par tous ▪ Confiance (et partenaires dignes de confiance) <p><i>Conclusion: coopération qui contribue à améliorer l'efficacité</i></p>
Co-construction	Les partenaires de collaboration travaillent ensemble sur un exercice et mettent en relation leurs connaissances individuelles les unes avec les autres de sorte que les protagonistes acquièrent des connaissances et élaborent en commun des solutions aux exercices et aux problèmes.	<ul style="list-style-type: none"> ▪ Caractère obligatoire et engagement à travailler en commun ▪ Harmonisation concernant le processus de travail ▪ Confiance (et partenaires dignes de confiance) <p><i>Conclusion: coopération qui contribue à améliorer l'efficacité</i></p>

3. MODÈLES D'HARMONISATION DE L'ÉVALUATION EN COMMUN

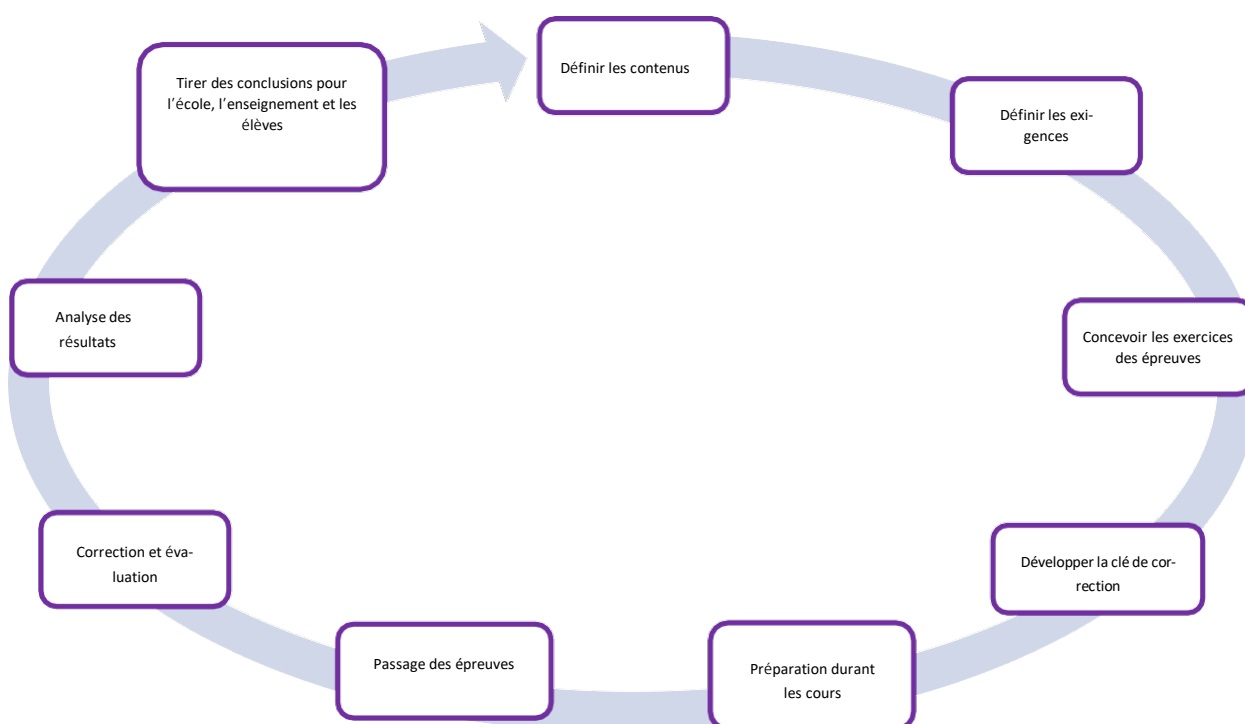
Degrés d'harmonisation selon Franz Baeriswyl

Le rapport final relatif au sous-projet 2 dans le cadre du projet de la CDIP «Garantie de l'accès sans examen aux hautes écoles universitaires» reprend les six degrés d'harmonisation selon Franz Baeriswyl (Joachim 2016: 4-6) ; ces derniers peuvent plus ou moins être librement combinés:

1 Divulgation réciproque des modèles de performances	<p>La divulgation réciproque des exercices parmi les examinateurs</p> <ul style="list-style-type: none"> ▪ permet de sensibiliser à la qualité des épreuves, ▪ est une mesure qui renforce la confiance, ▪ permet d'échanger des points de vue personnels.
2 Harmonisation des exigences liées à la procédure	<p>Les examinateurs se mettent d'accord sur les variables suivantes:</p> <ul style="list-style-type: none"> ▪ outils autorisés, ▪ modalité de l'épreuve (orale/écrite) et durée de l'épreuve, ▪ modalités de préparation, ▪ type d'exercice (p. ex. questions à choix multiples, questions semi-ouvertes ou ouvertes, tâches de projet), ▪ modalités de correction et d'évaluation (p. ex. pour les langues étrangères: est-ce les erreurs ou les réponses correctes qui sont comptées?). <p>Ne sont pas définis:</p> <ul style="list-style-type: none"> ▪ domaines de contenus, compétences testées, niveaux d'exigences des différents exercices.
3 Harmonisation des domaines de contenus	<p>Cette harmonisation repose sur les objectifs indicatifs des plans d'études cantonaux:</p> <ul style="list-style-type: none"> ▪ les enseignants peuvent définir eux-mêmes des thèmes prioritaires, ▪ les expériences positives issues de la collaboration peuvent améliorer cette dernière en permanence.
4 Harmonisation des domaines de compétences	<p>Élaboration de modèles de compétences spécifiques à chaque discipline:</p> <ul style="list-style-type: none"> ▪ les contenus peuvent être librement choisis, élaboration de compétences transversales, ▪ travail d'harmonisation intensif. Élaboration des domaines de compétences centraux à partir des plans d'études.
5 Harmonisation des domaines de contenus et des domaines de compétences ainsi que des niveaux d'exigences	<p>Chaque exercice doit être analysé selon au moins trois dimensions principales (analyse triangulaire):</p> <ul style="list-style-type: none"> ▪ attribution à un domaine de contenus dans la discipline, ▪ attribution à un domaine de compétences (un domaine de compétences peut souvent être testé au moyen de différents contenus), ▪ définition du niveau d'exigences.
6 Dispositions générales centrales et épreuves internes aux établissements (maturité interne p. ex.)	<p>Directives cadres cantonales avec certaines conditions, p. ex.:</p> <ul style="list-style-type: none"> ▪ pour les disciplines de base, une épreuve de maturité par établissement, ▪ épreuve portant sur des compétences spécifiques et transversales, ▪ les disciplines principales et secondaires doivent être uniformes en termes de compétences et de niveau d'exigences (différents contenus possibles).

Le modèle circulaire «Évaluation en commun» selon Katharina Maag Merki

Tandis que le modèle de Franz Baeriswyl concerne le degré d’harmonisation de l’épreuve finie, le modèle circulaire de Katharina Maag Merki se penche sur l’harmonisation du processus. Maag Merki distingue différentes étapes dans le processus de conception, de création, de passage et d’évaluation d’une épreuve.¹ En principe, ces étapes sont exécutées par chaque enseignant, même si celui-ci travaille seul. Dans le cadre de l’«évaluation en commun», plusieurs étapes peuvent faire l’objet d’une concertation. L’analyse de la pratique montre que les étapes permettent de définir des contenus et des exigences, de concevoir des exercices d’épreuve et de développer des clés de correction. Elle montre aussi que la réalisation des épreuves est souvent une responsabilité commune, que les corrections sont parfois effectuées en commun, mais qu’une analyse en commun permettant de tirer des conclusions pour l’enseignement est rarement effectuée (Holmeier et al. 2017). Le modèle de Maag Merki montre aussi que l’achèvement d’une épreuve représente en même temps le point de départ pour mener une réflexion sur l’enseignement (Joachim 2016: 5).



Source: (Holmeier et al. 2017)

¹ On parle généralement d’«évaluation en commun» si au minimum le degré «Développer la clé de correction» est harmonisé.

Dans la pratique, l'«évaluation en commun» peut être effectuée sous différentes formes et avec une périodicité différente:

- sous forme d'épreuve annuelle au début ou à la fin d'une année scolaire,
- sous forme de travaux comparatifs pendant l'année scolaire,
- épreuve orale ou écrite,
- épreuve individuelle ou en tandem,
- lors de l'évaluation de journaux d'apprentissage, de portefeuilles (hiver 2009) ou de podcasts réalisés par exemple dans des semaines de projet ou des phases d'apprentissage autonome,
- sous forme d'épreuves de maturité harmonisées (la «maturité interne» conformément aux directives cantonales).

Dans le cadre des projets de développement scolaire, les directions des écoles formulent, en collaboration avec l'équipe enseignante, les conditions cadres et la mise en œuvre organisationnelle de l'«évaluation en commun». La fréquence et le degré de coordination des contenus peuvent varier. Plus l'évaluation en commun est fréquente, plus la collaboration sera intensive.

Explications et remarques concernant les étapes du modèle circulaire:

Définir les contenus	Le groupe de discipline s'entend sur les contenus prioritaires en fonction du plan d'études; l'équilibre entre les préférences communes et les préférences personnelles doit être négocié. L'«évaluation en commun» repose sur des étapes d'apprentissage terminées, mais peut aussi être plus complète et se rapporter par exemple au contenu d'une année scolaire complète.
Définir les exigences	Le groupe de discipline formule des objectifs d'apprentissage orientés vers les compétences. Ces objectifs sont communiqués aux élèves. Ce qui est important doit être bien clair pour les élèves.
Concevoir les exercices des épreuves	Le groupe de discipline définit les compétences, élabore - ensemble ou selon le principe de division des tâches - des exercices pour les épreuves. Ces derniers sont appropriés et clairement formulés. Le groupe de discipline les valide et s'assure qu'ils sont cohérents avec les cours et qu'ils sont orientés vers les compétences. En fonction du degré d'harmonisation du contenu des cours, une partie des exercices peut être différente ou les points communs de l'épreuve peuvent concerner surtout sa structure, les types d'exercices et le niveau d'exigences.
Développer la clé de correction	Au besoin, le groupe de discipline élabore des modèles de solutions avec des variantes, détermine la pondération des exercices, définit les critères d'évaluation, conçoit, en particulier pour les exercices complexes et les épreuves orales, des grilles de critères qui doivent être communiquées aux élèves avec les objectifs d'apprentissage. Une attention particulière doit être accordée aux échelles de mesure de réussite: qu'est-ce qui est suffisant ? Qu'est-ce qui est excellent? Il faut aussi déterminer la pondération appliquée à l'épreuve dans le bulletin de notes.
Préparation durant les cours	Dans un esprit d'équité, le groupe de discipline s'entend sur la forme et sur l'intensité de la préparation de l'épreuve et utilise des documents communs (p. ex. dossier d'exercices pour réviser, en d'exercices sur un cloud, outils en ligne pour s'exercer). Une méthode non autorisée d'enseignement, uniquement ciblée sur l'examen, («teaching to the test») peut ainsi

être évitée.

Passage de l'épreuve	Le groupe de discipline s'entend sur les outils autorisés et sur la durée de l'épreuve et règle les aspects organisationnels.
Correction et évaluation	Le groupe de discipline se met d'accord sur des formes de collaboration adaptées lors de la correction. L'équilibre collégial de la pratique de notation, en utilisant des critères d'évaluation communs, par exemple avec la contre-correction de dissertations, ou l'enregistrement des épreuves orales, favorise la comparabilité des exigences et confère une sécurité aux enseignants.
Analyse des résultats	Un aperçu du nombre de points obtenus dans les différents exercices permet à l'équipe enseignante de discuter des anomalies dans le cadre d'une discussion collégiale: rapport entre les performances fournies et les attentes de l'enseignant, degré de difficulté des exercices, signes d'une sollicitation excessive ou insuffisante des élèves ou de classes entières, notation, notamment vérification de la limite satisfaisante dans le système de notation. Les données appartiennent en principe aux enseignants.
Tirer des conclusions	Le processus de développement du modèle circulaire est ancré à ce niveau. Pour les élèves, les épreuves sommatives au cours de l'année scolaire contiennent une composante formative, par exemple lorsque l'élève reprend son travail individuellement dans le cadre d'améliorations, de journaux d'apprentissage ou d'apprentissages en tandem. La discussion au sein de l'équipe enseignante conduit si nécessaire au perfectionnement du répertoire méthodologique et didactique si des thèmes sont de nouveau abordés en cours. Le cas échéant, une équipe décide de procéder à des ajustements dans les contenus prioritaires et la progression. Si nécessaire, le concept d'évaluation en commun ainsi que le type d'exercices peuvent aussi être repensés.

Défis

Dans l'étude «Évaluation en commun. Une analyse de cas dans les gymnases de Suisse» (Holmeier et al. 2017: 241-261), huit champs de tension récurrents sont mis en évidence dans l'analyse portant sur les exemples de mise en œuvre. Ces champs de tension doivent être pris en compte dans les groupes de discipline lors de la planification de l'«évaluation en commun»:

1. Décision d'une «évaluation en commun»: caractère facultatif ou obligatoire?
2. Fonction de l'épreuve: contrôle ou encouragement? L'épreuve a-t-elle une finalité sommative ou formative?
3. Création des exercices: qui crée et connaît les exercices?
4. Préparation des élèves: en commun ou individuelle?
5. Elaboration de bons exercices: «Plus petit dénominateur commun» ou haute qualité des exercices?
6. Correction: en commun, individuelle, anonyme?
7. Note: pondération et importance,
8. Utilisation des résultats: oublier très vite les résultats ou déduire ce qu'il faut faire?

4. OPPORTUNITÉS OFFERTES PAR L' «ÉVALUATION EN COMMUN»

Pour les élèves

- Ils apprennent à adapter les techniques d'apprentissage et de travail à des épreuves parfois globales.
- Le niveau d'exigences, les objectifs d'apprentissage, les critères de notation et les conditions de réussite sont comparables et ne sont pas dépendants de l'enseignant.
Cela crée de la confiance!

Pour les enseignants

- Une équipe se penche sur le plan d'études et définit une base commune. Une marge de manœuvre suffisante est toutefois laissée lors de la mise en œuvre et pour les préférences.
- La collaboration d'égal à égal crée de la confiance et des automatismes.
- Les enseignants réfléchissent ensemble sur la manière dont ils perçoivent l'épreuve.
- Ils développent une pratique comparable en matière d'exigences et de notation.
- La répartition des tâches au sein de l'équipe et la collecte en commun des supports permettent un gain de temps.
- Les évaluations en commun permettent aux enseignants de constater s'ils sont à peu près sur la même voie concernant la progression, les exigences et la pratique de notation.
Cela crée de la sécurité!

5. INFORMATIONS GÉNÉRALES SUR LE THÈME «ÉVALUATION»

L'harmonisation des objectifs d'apprentissage et des types d'exercices lors des contrôles des connaissances²

La taxinomie de Metzger (1993) peut être un instrument utile lors de la discussion concernant les objectifs d'apprentissage et l'élaboration des épreuves. Elle offre un système de référence, une sorte de langage commun quand il s'agit d'aborder les objectifs d'apprentissage, la mise en place des cours et l'évaluation des performances.

Échelons de processus selon Anderson	Caractéristiques	Exemples
1. Se souvenir de connaissances	Se souvenir	
1.1. Reconnaître	Reconnaître des informations apprises dans un environnement inchangé.	Reconnaître des concepts, définitions, faits, noms, données, processus, règles, etc. présentés dans les exercices.

² Repris de Heller-Andrist/Joachim (2015): 4-6 dans une forme abrégée et un peu simplifiée.

1.2. Restituer, faire appel à des connaissances	Reproduire de manière inchangée des informations apprises (citer, énumérer, expliquer, etc.).	Restituer de manière inchangée des concepts, définitions, faits, noms, données, processus, règles, théories, etc.; qui ont été appris et exécuter de manière routinière des processus assimilés.
2. Traiter les connaissances	Comprendre et appliquer	
2.1. Comprendre	Saisir le sens, «représenter» par analogie les informations (les décrire, les expliquer, les interpréter, les justifier, les comprendre, etc.).	Expliquer avec ses propres mots des concepts/formules/faits; résumer un texte; interpréter des graphiques; traduire des textes; trouver des exemples se rapportant à un fait.
2.2. Appliquer	Transposer la structure apprise sur un contenu nouveau sur le plan linguistique, identique sur le plan du contenu.	Transposer des règles/lois/méthodes/formules/procédures, etc. sur de nouveaux contenus; appliquer ce qui a été appris dans des situations pratiques; élaborer ses propres exemples par rapport à des structures existantes.
3. Créer des connaissances	Traiter des problèmes de manière globale	
3.1. Analyser	Examiner un fait de manière complète et systématique; dans ce cadre, la structure de critères nécessaire à cet effet doit être recréée.	Comparer des faits (données, avis, affirmations, etc.) au moyen de critères élaborés soi-même; diviser les faits/systèmes dans des parties logiques.
3.2. Évaluer	Évaluer un fait de manière complète et systématique; dans ce cadre, la structure de critères nécessaire à cet effet doit être recréée.	Comparer, évaluer et sélectionner des alternatives; formuler un jugement personnel fondé; évaluer des faits au moyen de critères adaptés et cohérents.
3.3. Créer	Relier des informations distinctes pour former un nouvel ensemble.	Ebaucher de nouveaux plans/structures/schémas/théories; saisir de nouveaux textes; formuler des propositions pour résoudre des problématiques concrètes.

Les différents échelons de processus montrent comment les connaissances sont traitées. La première catégorie (se souvenir de connaissances) se réfère au rappel et à la reproduction (tâches de reproduction), tandis que les deux autres catégories (traiter et créer des connaissances) impliquent le transfert des connaissances apprises (tâches de transfert).

Pour définir le degré de difficulté d'un exercice dans l'épreuve, il est donc essentiel de prendre en compte le cours précédent. En effet, un exercice difficile à première vue, dans lequel il s'agit de résoudre un problème, n'est qu'une simple tâche de reproduction s'il a déjà été traité en cours car l'élève, pendant l'épreuve, doit seulement faire appel au processus de résolution de problème.

En plus de son rôle de système de référence lorsque l'on aborde les niveaux d'exigences, cette taxinomie peut remplir d'autres fonctions dans la pratique:

1. Elle peut faire apparaître le degré de difficulté différent des objectifs d'apprentissage et donner des indications sur le niveau d'exigences: ainsi, une situation dans laquelle tous les objectifs d'apprentissage se situent sur les mêmes échelons de processus, ou dans laquelle d'autres échelons de

processus ne sont pas du tout couverts, ne passera pas inaperçue.

2. Elle peut être une aide pour la planification du cours grâce au classement des objectifs d'apprentissage.
3. Elle peut être une aide lorsque l'on réfléchit comment s'assurer que les objectifs d'apprentissage sont atteints et quels sont les exercices appropriés à cet effet.
4. Elle peut être une aide en vue de la coordination des objectifs d'apprentissage, de la conception des cours et de la vérification des objectifs d'apprentissage.

Règles internes pour l'évaluation des performances et le système de notation

Dans la publication «Des notes, comment faire autrement», Anton Strittmatter formule seize «règles didactiques et d'équité» que ne se réfèrent pas seulement à l'«évaluation en commun», mais qu'il convient d'examiner lors de sa planification et sa mise en œuvre (Fischer et al. 2009: 71-74). Voici une synthèse des recommandations. Il faut ...

1. que la finalité de l'épreuve soit claire pour soi-même et pour les élèves (diagnostic d'apprentissage personnel/objectif de développement; finalité sommative/sélective; motivation);
2. proposer des exercices qui permettent de tester les compétences et les principaux objectifs d'apprentissage conformément au plan d'études (validité du contenu);
3. formuler les exercices de manière claire, univoque et compréhensible;
4. expliquer clairement aux élèves la finalité des connaissances testées en vue de la suite de l'apprentissage et montrer quels sont les objectifs d'apprentissage (aptitudes, sujet) à atteindre;
5. laisser aux élèves une marge de manœuvre suffisante pour apprendre ce qui sera évalué;
6. satisfaire à l'exigence d'objectivité: tous les examinateurs doivent évaluer et noter de la manière la plus semblable possible les exercices;
7. évaluer selon des critères;
8. pondérer les critères/éléments et communiquer à l'avance la pondération;
9. définir ensuite les échelles de mesure de réussite et les communiquer;
10. faire en sorte que les corrections et les notes soient compréhensibles: lien entre auto-évaluation et évaluation par un tiers;
11. baser sur plusieurs épreuves les appréciations ayant un impact sur la promotion;
12. donner un feedback en temps utile et de façon nuancée;
13. donner des possibilités de correction;
14. si, contre toute attente, les épreuves se sont avérées trop faciles/trop difficiles: ne pas les prendre en compte pour la promotion;
15. se contrôler soi-même en ce qui concerne des «distorsions» inconscientes, comme la première impression, les normes comparatives implicites, les attentes implicites.

Check-list sur le thème «Évaluer»

Un document axé sur la pratique, consacré au thème de l'évaluation en général est la «check-list comme aide» publiée en 2016 par l'Office de l'enseignement secondaire du 2^e degré et de la formation professionnelle du canton de Zurich. Ce document est disponible au format PDF (en allemand) sur <https://www.zemces.ch/de/kompetenzzentrum-mittelschulen/themen/gemeinsames-pruefen>, «Conditions cadres et recommandations des cantons».

6. GLOSSAIRE

Cadre de référence	<p>Le cadre de référence décrit précisément quelles capacités les élèves devraient avoir dans un domaine et quelles situations-problèmes devraient être résolues (formulations «sait faire»).</p> <p>Un cadre de référence forme la base pour la planification et le développement de programmes d'apprentissage des langues et de diplômes de langue. Les compétences et le niveau requis sont définis à l'aide de descripteurs s'appuyant sur des bases empiriques.</p> <p>=> <i>Cadre européen commun de référence CECR</i></p>
Cadre européen commun de référence pour les langues (CECR)	=> <i>Cadre de référence</i>
Caractéristiques de la procédure	<p>Caractéristiques de la procédure lors de l'élaboration, du passage et de l'évaluation d'épreuves communes (p. ex. degré d'obligation, moment du passage de l'épreuve, champ de l'épreuve, etc.).</p> <p>=> <i>Exigences liées à la procédure</i></p>
Certification	<p>Une épreuve standardisée au niveau national ou international et dont les résultats sont précisément comparables. À titre d'exemples, on peut citer le diplôme linguistique DELF pour le français, les épreuves Cambridge pour l'anglais et les épreuves PLIDA pour l'italien.</p>
Compétence diagnostique	<p>Capacité qu'a l'enseignant à préparer et à effectuer une évaluation des performances permettant de tirer les bonnes conclusions en ce qui concerne les processus d'enseignement et d'apprentissage.</p>
Compétences	<p>Capacités et aptitudes cognitives innées ou acquises permettant à l'individu de résoudre certains problèmes; capacités et aptitudes motivationnelles, volitives et sociales qui en découlent et permettant d'utiliser avec succès et de manière responsable ces solutions dans différentes situations (d'après Weinert 2001). «Avoir des compétences dans un domaine particulier signifie que l'on peut agir avec succès dans ce domaine.» (Klieme et al. 2007)</p>
Compétences disciplinaires de base (CDIP 2016b)	<p>Compétences indispensables pour les «études générales» et qui ne peuvent être compensées.</p>
Conditions cadres	<p>Directives pour l'élaboration et la réalisation de l'épreuve, p. ex. définition des domaines de contenus et de compétences, des outils autorisés, etc.</p>
Confidentialité	<p>L'une des => <i>exigences liées à la procédure</i>. La correction de l'épreuve est-elle effectuée de manière anonyme et l'utilisation des données est-elle conforme aux règles de confidentialité?</p>
Critères d'évaluation	<p>Les critères qui permettront d'évaluer une épreuve peuvent être fixés à l'avance, en se référant par exemple à des réponses types.</p>
Critères de qualité	<p>Aussi appelés => <i>Critères de qualité des tests</i>: => <i>Objectivité</i>, => <i>Validité</i>, => <i>Fiabilité</i> sont les 3 principaux critères de qualité. Critères supplémentaires: => <i>Egalité des chances</i>, => <i>Rendement</i> et => <i>Équité</i>.</p>

Critères de qualité des tests	Également appelés critères de qualité => <i>Critères de qualité, => Théorie des tests</i>
Développement individuel	L'=> <i>évaluation formative des performances</i> cible le développement individuel de l'élève.
Directives de correction	Directives élaborées ou prescrites en commun et indiquant comment une épreuve doit être corrigée et évaluée.
Domaine d'exigences	Précise le type et la pondération de la performance attendue à l'épreuve. Exemples de domaines d'exigences (d'après Bloom): <ul style="list-style-type: none"> - Connaître, savoir et comprendre: exécuter des procédures courantes suivant des directives, appliquer des formules et des règles élémentaires (niveau I) - Appliquer, utiliser: choisir et appliquer des stratégies simples pour résoudre les problèmes, travailler avec des modèles (niveau II) - Jugement scientifique, créativité: concevoir des modèles pour des situations complexes et les utiliser dans de nouvelles situations (niveau III) Chaque domaine d'exigences peut compter différents niveaux d'exigences.
Domaine de compétences	Formulation des exigences que les élèves doivent satisfaire dans un domaine d'une discipline (cf. modèle d'harmonisation 4 selon Baeriswyl). On distingue d'une part la détermination des domaines de compétences dans une discipline et d'autre part celle des domaines de contenus. => <i>Domaine de contenus, => Modèle d'harmonisation</i>
Domaine de contenus	Les domaines de contenus sont fixés d'après les directives des plans d'études cantonaux en fonction des performances à évaluer. On distingue d'une part les domaines de contenus (définis par un niveau élevé d'harmonisation et une liberté restreinte de l'enseignant lors de l'établissement de son programme) et d'autre part les => <i>domaines de compétences</i> .
Egalité des chances	Un => <i>critère de qualité des tests</i> . Une épreuve doit mesurer une performance indépendamment de l'école, de la classe et du contexte socio-économique. => <i>Critères de qualité, théorie des tests</i>
Epreuve «de promotion»	L'épreuve est-elle prise en compte pour la promotion au degré supérieur ou non? => <i>Évaluation sommative des performances</i>
Epreuve commune	Epreuve en tant que telle, résultat du processus propre au groupe de discipline => <i>Évaluation en commun</i> .
Epreuve comparable	Epreuve qui, grâce à son niveau d'harmonisation par rapport à une autre épreuve, permet une comparaison des performances obtenues. => <i>Transparence</i>
Epreuve identique	Le sujet, les exercices, la pondération, les conditions cadres et les exigences liées à la procédure sont les mêmes. => <i>Epreuve identique</i>
Epreuve identique	Une épreuve identique correspond au degré maximal de l'=> <i>harmonisation</i> .
Epreuve transversale	Terme utilisé comme synonyme d'une => <i>évaluation en commun</i> qui comprend p. ex. toute une volée d'élèves (dans une ou plusieurs écoles).

Équité	Un => <i>critère de qualité des tests</i> . L'épreuve doit être claire: ses objectifs, sa forme et son sujet doivent être communiqués aux élèves de manière transparente.
Évaluation en commun	Processus de discussion (au sein d'un groupe de discipline) pour élaborer une épreuve. «Évaluer en commun» signifie formuler des attentes communes relatives à un => <i>domaine d'exigences</i> .
Évaluation formative des performances	Évaluation visant l'amélioration et le développement individuel des élèves pendant leur formation. Analogie au club de sport: l'entraîneur conseille les joueurs pour que chacun s'entraîne de manière optimale et puisse s'améliorer. => <i>Développement individuel</i> , => <i>Évaluation pronostique des performances</i> , => <i>Évaluation sommative des performances</i>
Évaluation pronostique des performances	Évaluation concernant une performance future. Analogie au club de sport: l'entraîneur choisit ses meilleurs sportifs pour le prochain championnat. => <i>évaluation sommative des performances et évaluation formative des performances</i>
Évaluation sommative des performances	Évaluation finale visant à contrôler le taux d'atteinte d'un objectif d'apprentissage. Analogie au club de sport: l'entraîneur veut savoir à quel niveau de performances ses joueurs se situent et définit un critère de temps (limite satisfaisante). => <i>Évaluation pronostique des performances, évaluation formative des performances</i>
Exigences liées à la procédure	Exigences concernant l'élaboration, le passage et l'évaluation => <i>Epreuves en commun</i> . Ces exigences sont développées sur la base des => <i>caractéristiques de la procédure</i> .
Fiabilité	Un => <i>critère de qualité des tests</i> . La fiabilité est l'un des trois principaux critères de qualité des tests tirés de la théorie des tests: le test a une cohérence intrinsèque. La fiabilité d'un test peut être améliorée par: <ul style="list-style-type: none"> - Sa forme homogène, - La => <i>Sélectivité</i> de chaque exercice, - L'indépendance des exercices les uns par rapport aux autres, - L'abandon des «speed-tests» (connaissances combinées à la rapidité de réponse). - L'augmentation du nombre d'items (exercices), - L'augmentation du nombre d'exercices de niveau de difficulté moyen, => <i>Critères de qualité</i>
Groupe de référence	Groupe dans lequel, resp. par rapport auquel une performance est mesurée (p. ex. une classe, une volée): Dans le cas d'une => <i>épreuve reposant sur la norme sociale</i> , le groupe de référence est défini à l'avance.
Harmonisation superficielle	Niveau d'harmonisation le plus faible. On définit p. ex. les outils autorisés, la durée de l'épreuve et le type de questions. => <i>Harmonisation</i>
Harmonisation/ Modèles d'harmonisation	Processus visant à rendre comparables les épreuves communes. Peut aller d'un degré harmonisation minimal (outils de travail autorisés, durée, forme des questions) à un degré d'harmonisation maximal, c'est-à-dire à une=> <i>épreuve identique</i> . Voir les 6 modèles d'harmonisation selon Baeriswyl. => <i>Harmonisation superficielle</i>

Liberté pédagogique	Liberté qu'ont les enseignants de choisir leur programme et leur méthodologie (à ne pas confondre avec la «liberté de l'enseignant»).
Maturité interne	Epreuve de maturité interne à l'établissement et identique pour l'ensemble des élèves d'une discipline. => <i>Epreuve identique</i>
Modalités de passage de l'épreuve	Comment l'épreuve se déroule-t-elle (lieu, moment, outils autorisés)? Ce type de dispositions correspond à une harmonisation superficielle et constitue un exemple d'exigence liée à la procédure. => <i>Harmonisation / Modèle d'harmonisation</i>
Modèle de compétences	Ensemble de compétences à tester dans une discipline.
Niveau d'exigences	Taux de réussite attendu par exercice. Le niveau d'exigences peut reposer sur un => <i>modèle de compétences</i> ou sur un => <i>plan d'études</i> . Il peut être normatif si la procédure prévoit une sélection. S'il est de caractère formatif, il permet de poser un diagnostic pour formuler d'autres exercices d'apprentissage. Il est important que le niveau d'exigences soit «correct» et judicieux pour l'ensemble des exercices et que l'on ne teste par exemple pas seulement les connaissances et les savoirs. => <i>Domaine d'exigences</i>
Niveau de performances	Qualité des performances par rapport à un groupe ou à un niveau d'exigences donné.
Norme basée sur une liste de critères	Le résultat indique si l'épreuve – sans tenir compte du groupe – est réussie ou non. => <i>Norme sociale</i>
Norme sociale	Le résultat de l'épreuve indique le niveau de performance par rapport à un groupe de référence. => <i>Norme basée sur une liste de critères</i>
Objectivité	Un => <i>critère de qualité des tests</i> . L'objectivité est l'un des trois principaux critères de qualité tirés de la théorie des tests. Les résultats des tests sont valables indépendamment de la personne chargée du test et de la situation d'examen si l'objectivité est garantie lors du passage, de l'évaluation et de l'interprétation de l'épreuve. => <i>Critères de qualité</i>
Plan d'études	L'un des outils à disposition pour l'harmonisation des épreuves: p. ex. plan d'études cadre pour les écoles de maturité, principaux objectifs de l'école, objectifs par discipline, etc.
Pondération	Lors de l'évaluation, on définit l'importance de chaque exercice par rapport à l'ensemble de l'épreuve. On tient compte du programme/sujet et du => <i>niveau d'exigences</i> , => <i>Répartition des exercices</i>
Rendement	Un => <i>critère de qualité des tests</i> . La gestion du temps doit être optimale lors de l'élaboration d'épreuves communes; il doit donc y avoir un juste équilibre entre effort investi et bénéfice. => <i>Critères de qualité</i>

Répartition des exercices	L'une des => <i>exigences liées à la procédure</i> , également nommée représentativité. Comment les différents exercices sont-ils répartis par sujet? Sont-ils représentatifs des objectifs d'apprentissage à tester et du sujet? => <i>Pondération</i>
Sélectivité	Lien entre l'exercice (item) et le résultat global de l'épreuve.
Sujet	Domaines de contenus testés dans une discipline.
Taux de réussite	Le taux de réussite concerne le => <i>niveau d'exigences d'une épreuve</i> et s'exprime en pourcent. Ainsi, ce taux doit être de 50% pour qu'un examen de diplôme soit considéré comme réussi, alors qu'il doit souvent être de 66% à 75 % pour obtenir la mention «suffisant» à une épreuve gymnasiale.
Teaching to the test	Préparation ciblée sur une épreuve («entraînement pour une épreuve»). Le «teaching to the test» est l'un des inconvénients pouvant se produire dans le cadre de l'«évaluation en commun».
Test	Ici synonyme d'«épreuve».
Théorie des tests	Théorie qui permet de définir la qualité des épreuves. => <i>Critères de qualité</i>
Transparence	Une épreuve est transparente si elle est comparable à d'autres épreuves, par exemple lorsque les => <i>conditions cadres</i> , => <i>exigences liées à la procédure</i> , => <i>compétences</i> ou => <i>sujet</i> sont définis. => <i>Epreuves comparables</i>
Type d'exercice	On opère une distinction entre les exercices «ouverts» et «fermés». Les exercices «ouverts» exigent des critères d'évaluation et des directives de corrections détaillées. Les exercices «fermés», tels que les questionnaires à choix multiples, sont en revanche compliqués à créer mais plus simples à évaluer.
Validité	Un => <i>critère de qualité des tests</i> . La validité est l'un des trois principaux critères tirés de la théorie des tests: la mesure est valide lorsqu'elle mesure ce qui doit être mesuré: <ul style="list-style-type: none"> - Validité du contenu: l'objet de la mesure peut clairement être décrit. - Validité des critères: il y a un lien entre la performance et le critère avec lequel le test est corrélé. - Validité de construction : le test mesure l'aptitude qu'il doit évaluer et pas une autre. - Validité apparente: un profane voit le lien entre les exercices et l'objet évalué. => <i>Critères de qualité</i>

7. BIBLIOGRAPHIE

Eberle, Franz/Brüggenbrock, Christel (2013): Bildung am Gymnasium. Bern: Schweizerische Konferenz der kantonalen Erziehungsdirektoren.

Eberle, Franz et al. (2008): Evaluation der Maturitätsreform 1995 (EVAMAR). Schlussbericht zur Phase II. Bern: Staatssekretariat für Bildung und Forschung SBF. <https://www.sbf.admin.ch/sbf/de/home/bildung/maturitaet/gymnasiale-maturitaet/evamar.html> (24.09.2020).

EDK (2016a): Gymnasiale Maturität – Langfristige Sicherung des prüfungsfreien Hochschulzugangs: Entscheidung über den Abschluss der Teilprojekte - Folgemaassnahmen. Beschluss Plenarversammlung vom 17.3.2016.

EDK (2016b): Anhang zum Rahmenlehrplan für die Maturitätsschulen vom 9. Juni 1994 : Basale fachliche Kompetenzen für allgemeine Studierfähigkeit in Erstsprache und Mathematik vom 17. März 2016. <https://edudoc.ch/record/121436?ln=de> (24.09.2020).

EDK (2016c): Gymnasiale Maturität. Abschluss der Projekte in Folge von EVAMAR II. EDK-Newsletter Nr. 1, 3-5. <https://www.edk.ch/dyn/17833.php> (24.09.2020).

Gräsel, Cornelia/Fussangel, Kathrin/Pröbstel Christian (2006): Lehrkräfte zur Kooperation anregen - eine Aufgabe für Sisyphos? In: Zeitschrift für Pädagogik 52 (2): 205–219.

Heller-Andrist, Simone/ Joachim; Karin (2015): Évaluation en commun – einige Grundlagen. <https://www.zemces.ch/de/kompetenzzentrum-mittelschulen/themen/gemeinsames-pruefen> (24.09.2020).

Klieme, Eckhard/ Maag Merki, Katharina/ Hartig, Johannes (2007): Kompetenzbegriff und Bedeutung von Kompetenzen im Bildungswesen. In: Hartig, Johannes/Klieme, Eckhard (Hrsg.): Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik. Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung. Bonn u. a.: BMBF, 5-15.

Holmeier, Monika/Maag Merki Katharina/Hirt, Carmen (2017): Évaluation en commun. Eine Fallanalyse in Gymnasien der Schweiz. Wiesbaden: Springer Verlag.

Mero, Romano (2011): Évaluation en commun: Alphabetisches Glossar. <https://www.zemces.ch/de/kompetenzzentrum-mittelschulen/themen/gemeinsames-pruefen> (24.09.2020).

Mero, Romano (2011): Évaluation en commun: Thematisches Glossar. <https://www.zemces.ch/de/kompetenzzentrum-mittelschulen/themen/gemeinsames-pruefen> (24.09.2020).

Mero, Romano (2013): Bericht zum Teilprojekt 2, Phase 1: Unterstützungsangebote zum Gemeinsamen Prüfen. EDK-Projekt «Gymnasiale Maturität – langfristige Sicherung des prüfungsfreien Hochschulzugangs».

<https://www.zemces.ch/de/kompetenzzentrum-mittelschulen/themen/gemeinsames-pruefen> (24.09.2020).

Mero, Romano (2014): Bericht zum Teilprojekt 2, Phase 2: Umsetzungsarbeiten zum Gemeinsamen Prüfen. EDK-Projekt «Gymnasiale Maturität – langfristige Sicherung des prüfungsfreien Hochschulzugangs».

<https://www.zemces.ch/de/kompetenzzentrum-mittelschulen/themen/gemeinsames-pruefen> (24.09.2020).

Metzger, Christoph et al. (1993): Niveau d'exigences von Lernzielen und Prüfungen im kognitiven Bereich: Inhaltsniveaus, Prozeptionsniveaus, Reproduktion und Transfer. St. Gallen: Institut für Wirtschaftspädagogik an der Hochschule St. Gallen.

Strittmatter, Anton (2009): Hausregeln für Leistungsbewertung und Notengebung. In: Fischer, Doris/Strittmatter, Anton/Vögeli-Mantovani, Urs (Hrsg.): Noten, was denn sonst?! Zürich: Verlag LCH, 71- 75.

Weinert, Franz Emanuel (Hrsg.) (2001): Leistungsmessung in Schulen. Weinheim und Basel: Beltz.

Winter, Felix (2009): Das Portfolio: Impulse für das Lernen und die Leistungsbeurteilung. In: Fischer, Doris/Strittmatter, Anton/Vögeli-Mantovani, Urs (Hrsg.): Noten, was denn sonst?! Zürich: Verlag LCH, 53- 60.