

Rapport - 2023

Etude sur les facteurs de réussite et d'échec des étudiants en 1^{ère} année à l'université

Menée par l'Observatoire de la Transformation Pédagogique

<https://otp.edu.umontpellier.fr>

Résumé

Ce rapport présente une étude dont l'objectif est d'identifier les principaux facteurs qui influencent la réussite et l'échec des étudiants en première année à l'Université de Montpellier. Parmi les facteurs possibles, l'étude considère de multiples caractéristiques des étudiants (qu'elles soient socio-académiques, psycho- et socio-cognitives ou relevant de l'engagement dans les études) ainsi que la formation suivie. Elle se base sur une méthode statistique (*Random Forest*) qui permet de comparer directement l'influence relative de l'ensemble de ces facteurs (48 variables) sur la validation du semestre. En outre, cette étude invite à une réflexion sur la notion même de réussite à l'université. Elle propose de ne pas la réduire à la seule validation du semestre et d'envisager d'autres mesures de réussite, en lien avec la curiosité, l'autonomie, la réflexivité et l'esprit critique des étudiants. Pour ce faire, elle s'appuie sur la recherche en psychologie cognitive et sciences de l'éducation. Trois mesures de réussite sont ainsi considérées : la motivation (auto-déterminée), l'approche de l'apprentissage (en profondeur) et les croyances épistémiques (évaluatistes). De même, deux mesures d'échec sont examinées : les difficultés perçues et l'amotivation. Les facteurs de ces différentes mesures de réussite et d'échec ont été investigués sur la base de données portant sur 2047 étudiants dans 10 composantes. Il ressort notamment que :

- les principaux facteurs de réussite en termes de validation du semestre sont la composante de la formation suivie, la mention au Bac et la profession des parents ;
- quatre facteurs se révèlent très importants pour plusieurs formes de réussite et d'échec : la profession des parents, l'approche de l'apprentissage intégrative, la motivation intrinsèque à la connaissance et la composante de la formation suivie ;
- le soutien social perçu des enseignants et de la famille sont des facteurs relativement moins importants pour les différentes formes de réussite et d'échec ;
- l'engagement comportemental des étudiants (assiduité et travail personnel) ainsi que leurs données personnelles (âge, genre et nationalité) comptent parmi les facteurs les moins déterminants des différentes formes de réussite et d'échec.

Les résultats pointent également l'existence de liens étroits entre l'approche de l'apprentissage (en profondeur), la motivation (auto-déterminée) et les croyances épistémiques (évaluatistes). Afin de mieux comprendre ces liens, des analyses statistiques supplémentaires ont été réalisées. Celles-ci permettent de proposer un modèle de la manière dont les étudiants s'engagent dans leurs études.

1. Présentation de l'étude

1.1 Les objectifs de l'étude et son originalité

Ce rapport présente les résultats d'une étude menée par l'OTP dont l'objectif est d'identifier les principaux facteurs qui influencent la réussite ou l'échec des étudiants en première année à l'Université de Montpellier. À noter qu'une étude similaire sera réalisée en 2^e, puis en 3^e année. Parmi l'ensemble des facteurs possibles, on peut distinguer les caractéristiques des étudiants et les caractéristiques de la formation suivie. La présente étude se centre principalement sur les caractéristiques des étudiants et, de façon plus minimale, sur celles de la formation suivie. Elle vise à comparer le poids relatif de ces facteurs. En particulier, elle vise à comparer le poids relatif des origines sociales et du bagage scolaire des étudiants, de leurs caractéristiques socio- et psycho-cognitives et de la formation suivie.

Cette étude est originale à deux égards. Premièrement, elle se base sur une méthode statistique de type *machine learning*, à savoir l'algorithme *Random Forest* (RF), qui permet d'intégrer dans un même modèle un grand nombre de variables se présentant sous des formes différentes (quantitatives ou nominales). Ainsi, il est possible de comparer directement le poids relatif des différents facteurs pris en compte, lesquels sont habituellement étudiés de façon disjointe (Hellas et al., 2018).

Deuxièmement, cette étude ne réduit pas la réussite ou l'échec des étudiants à la validation ou non du semestre, autrement dit à leurs performances aux examens. Elle propose de considérer également d'autres mesures de réussite et d'échec en lien avec la curiosité, l'autonomie, la réflexivité et l'esprit critique des étudiants. Les mesures utilisées dans cette étude sont tirées de la recherche en psychologie cognitive et sciences de l'éducation. Il s'agit de la motivation (auto-déterminée), l'approche de l'apprentissage (en profondeur), les croyances épistémiques (évalutistes), les difficultés perçues et l'amotivation. Ces mesures correspondent à des concepts définis précisément et pouvant être étudiés au moyen d'un questionnaire administré aux étudiants. L'objectif est donc aussi d'identifier les principaux facteurs qui influencent la réussite ou l'échec dans ses formes complémentaires.

La motivation, l'approche de l'apprentissage et les croyances épistémiques sont des concepts qui décrivent différentes facettes socio et psycho-cognitives des étudiants. Or ces facettes ne sont pas indépendantes les unes des autres et entretiennent des relations complexes. Dans la dernière partie du rapport, nous présentons des analyses qui permettent de dégager un modèle pour mieux comprendre ces relations.

1.2 Le recueil des données

Les données ont été recueillies en 2020 et en 2021 auprès d'étudiants en 1^{ère} année du premier cycle, dans 10 composantes de l'Université, par le biais d'un questionnaire administré au moyen du logiciel *Le Sphinx*. Celui-ci a été soumis aux étudiants 2 à 3 mois après le début de leur formation, lors du 1^{er} semestre. Il comporte 80 questions relatives à un ensemble de variables présentées ci-après. Les étudiants ont pris en moyenne 22 minutes pour y répondre. En complément, des données issues d'Apogée ont été récupérées et intégrées à l'étude.

1.3 Les variables étudiées

Les variables considérées dans cette étude sont nombreuses (49 variables) et se présentent sous différentes formes (certaines étant numériques continues ou discrètes, d'autres nominales). Elles

décrivent des caractéristiques propres aux étudiants et, de façon minimale, des caractéristiques des formations suivies. Ces différentes caractéristiques relèvent de plusieurs dimensions :

- Les dimensions socio-académiques caractérisant l'étudiant à l'entrée en formation

Dimensions	Caractéristiques
Personnelle	Âge
	Genre
	Nationalité française ou étrangère
Académique	Type de Bac obtenu
	Mention au Bac
	Néobachelier ou non
	Redoublant ou non
Sociologique	Profession des parents
	Boursier ou non
	Job étudiant ou non

- Les dimensions psycho- et socio-cognitives de l'étudiant concernant ses apprentissages

Dimensions	Caractéristiques
Motivation	<i>Motivation autodéterminée, composée de :</i> Motivation intrinsèque à la connaissance Motivation intrinsèque à l'accomplissement Motivation intrinsèque à la stimulation Motivation extrinsèque identifiée
	<i>Motivation contrôlée, composée de :</i> Motivation extrinsèque externe Motivation extrinsèque introjectée
	Amotivation (ou résignation)
Approche de l'apprentissage	<i>Approche de l'apprentissage en profondeur, composée de :</i> Approche intégrative Approche réflexive Interactions avec les enseignants Interactions avec les étudiants
Croyances épistémiques	<i>Croyances épistémiques évaluatistes concernant :</i> Incertitudes des connaissances Complexité des connaissances Source des connaissances Justification des connaissances
Difficultés ressenties	<i>Difficultés concernant :</i> Orientation (choix du parcours ou des options) Comprendre les consignes Complexité des contenus Suivre le cours Saisir les enjeux du cours Méthodes d'apprentissage Organisation en dehors du cours
Soutien social perçu	<i>Source du soutien :</i> Parents Pairs Enseignants Services de l'université

	<i>Nature du soutien :</i> Financier et matériel Fonctionnement de l'université Apprentissages Confiance en soi Projet
Climat social perçu	<i>Emotionnel</i> Stressant versus rassurant
	<i>Autonomie</i> Cadré versus libre

- La dimension liée à l'engagement de l'étudiant dans ses études

Dimension	Caractéristiques
Engagement	Assiduité aux TD
	Temps alloué aux études en dehors des cours

- Les dimensions relatives à la formation suivie

Dimensions	Caractéristiques
Contexte de la formation suivie	Composante de la formation
	Type de diplôme
Formation souhaitée ou non	Formation correspondant au 1 ^{er} vœu sur Parcoursup ou non

1.4 La population des répondants et sa représentativité

L'échantillon de l'étude est constitué de 2168 étudiants (29,7 % de l'ensemble de la population interrogée). L'âge moyen est 18 ans et demi. Le pourcentage de filles parmi les répondants s'élève à 58,9%. Il est légèrement supérieur à celui de la cohorte étudiée (qui est de 50,6 %). Les filles se sont donc un peu plus mobilisées pour répondre. L'échantillon est constitué à 80,4 % de bacheliers généraux, 12,8 % de bacheliers technologiques et 1,3 % de bacheliers professionnels. 5.5 % ont des équivalents au Bac (diplômes étrangers, DAEU, ...). Cette répartition est bien représentative de celle de la population globale. En ce qui concerne les mentions au Bac, cet échantillon se compose davantage de bacheliers ayant eu la mention « assez bien » (30,4 %) ou « bien » (29,3 %) que de bacheliers « sans mention » (18,0 %) ou avec la mention « très bien » (16,7 %). En comparant la population des répondants à celle de la promotion, on observe que les bacheliers avec les mentions « bien » et « « très bien » ont relativement plus répondu. La population des répondants est donc globalement bien représentative de la promotion avec une légère surreprésentation des filles et des étudiants ayant les mentions les plus élevées. À noter des écarts parfois plus marqués lorsque les données sont examinées par composante. Les données globales et par composantes sont disponibles dans l'annexe A.

Au sein de l'échantillon des répondants, certaines données issues d'Apogée sont manquantes. Cela concerne 121 étudiants, lesquels ont été sortis de la base des données. Les analyses statistiques qui suivent portent donc sur les réponses de 2047 étudiants.

1.4 Les méthodes d'analyse

De nouvelles méthodes statistiques de type *machine learning* sont maintenant disponibles et l'une d'elles, parmi les plus performantes en termes de prédiction, permet de prendre en charge les variables dans leurs différentes formes. Il s'agit de l'algorithme *Random Forest* (RF), lequel se base sur

la méthode des arbres de décision binaires. Il fournit un modèle construit avec une partie des données et testé avec l'autre partie. On peut ainsi associer au modèle un niveau de performance prédictive (le taux de bonnes prédictions). Dans notre étude, nous avons choisi cette méthode des RF, d'une part, en raison de ses hautes performances prédictives, et d'autre part, parce qu'elle permet de déterminer et de comparer directement dans un même modèle l'influence sur la réussite ou l'échec d'un ensemble de variables qui autrement sont incommensurables. Des précisions sur cette méthode sont fournies dans l'annexe B.

2. Les mesures de la réussite et de l'échec

2.1 Comment mesurer la réussite et l'échec ?

Dans cette étude, nous proposons d'abord de suivre une approche de la réussite à l'université pouvant être qualifiée de « traditionnelle », puis, dans le prolongement de certains chercheurs (van der Zanden et al., 2019), d'opérer un changement de perspective en considérant d'autres formes de réussite et d'échec. Ce changement de perspective prend en compte des enjeux forts de l'université qui dépassent le seul objectif du diplôme validé et qui relèvent de l'épanouissement et du développement des étudiants. En effet, de nombreux enseignants à l'université tentent de développer chez leurs étudiants la curiosité, l'autonomie, la réflexivité et l'esprit critique, lesquels sont des compétences indispensables pour leur formation académique, pour leur future vie professionnelle, mais aussi pour leur engagement dans la société comme citoyens responsables et critiques. À plus ou moins long terme, ces compétences leur permettront de s'adapter à des situations changeantes ou nouvelles. Pour étudier ces compétences, différents concepts ont été définis par les chercheurs en psychologie cognitive et en sciences de l'éducation. Dans la présente étude, cinq concepts ont ainsi été retenus comme des mesures complémentaires de la réussite et de l'échec : la motivation (auto-déterminée), l'approche de l'apprentissage (en profondeur) et les croyances épistémiques (évaluatistes) comme des mesures de la réussite (et des mesures inverses de l'échec) ; les difficultés perçues et l'amotivation comme des mesures de l'échec (et des mesures inverses de la réussite). Ces concepts sont brièvement décrits dans ce qui suit. Des précisions sont apportées en annexe : sur la motivation et l'amotivation (annexe C), l'approche de l'apprentissage (annexe D) et les croyances épistémiques (annexe E).

La validation du semestre

Suivant les notes ministérielles (par exemple : Ménard, 2021), la réussite à l'université est associée à l'obtention du diplôme ou à la validation d'une année de la formation suivie. Autrement dit, elle est mesurée par la performance des étudiants aux examens. Les chiffres des taux de réussite et d'échec mesurés sur cette base sont relayés par les médias et constituent une référence pour les pouvoirs publics, les présidents d'université, les responsables des facultés, les enseignants, les étudiants et leurs parents. Dans la présente étude, nous mesurons la réussite et l'échec suivant cette approche traditionnelle en considérant la validation ou non du premier semestre en première année.

La motivation auto-déterminée

La motivation auto-déterminée (Ryan & Deci, 2000) correspond à une première mesure alternative de la réussite. Cette forme de motivation décrit une autonomie de l'étudiant dans la régulation de son comportement. Plus la motivation d'un.e étudiant.e est auto-déterminée, plus les raisons de s'engager dans une activité à l'université relèvent d'un choix libre, et moins elles traduisent un contrôle social extérieur (ex : travailler pour avoir une bonne note et faire ainsi plaisir à ses parents). Ce choix libre est associé à une curiosité à l'égard des connaissances et à une appétence pour les activités dans

lesquelles l'étudiant recherche la stimulation et le plaisir d'accomplir les tâches demandées. On parle à cet égard de motivation « intrinsèque ». Ce choix libre est aussi associé à la prise de conscience que certaines activités n'apportant aucun plaisir en l'état (ex : apprendre les verbes irréguliers en anglais) sont indispensables et ont du sens pour mener d'autres activités (parler anglais de manière fluide). On parle à cet égard de motivation « extrinsèque identifiée ». En bref, la motivation auto-déterminée traduit une forme de curiosité des étudiants à l'égard des connaissances et d'autonomie à l'égard des apprentissages.

L'approche de l'apprentissage en profondeur

L'approche de l'apprentissage en profondeur (Biggs & Tang, 2011) correspond à une seconde mesure alternative de la réussite. Elle décrit la manière dont les étudiants s'engagent dans les activités. Plus un étudiant suit une approche de l'apprentissage en profondeur, plus il cherche à mettre en lien les contenus enseignés, à réfléchir sur les consignes et ce qui est appris, et à interagir avec les étudiants et les enseignants pour mieux comprendre. Une telle approche contraste avec une approche de l'apprentissage en surface qui consiste essentiellement à mémoriser par cœur les contenus enseignés. En bref, l'approche de l'apprentissage en profondeur traduit une forme de réflexivité des étudiants sur les contenus, les activités et leurs propres apprentissages.

Les croyances épistémiques évaluatistes

Les croyances épistémiques évaluatistes (Hofer & Pintrich, 1997) correspondent à une troisième mesure alternative de la réussite. Le terme « croyance » est à entendre ici au sens philosophique de ce qui est tenu pour vrai, tandis que le terme « épistémique » renvoie à la notion de connaissance. Les croyances épistémiques décrivent ainsi les représentations des connaissances, en l'occurrence celles développées ou adoptées par les étudiants. Ces représentations peuvent être plus ou moins critiques. On distingue les représentations sur le statut des connaissances et les représentations sur les processus de construction et de validation des connaissances. Plus un étudiant a des représentations critiques sur le statut des connaissances, plus il reconnaît que les connaissances enseignées comportent des incertitudes, peuvent évoluer et sont complexes, et moins il suppose qu'elles sont absolues, figées et réduites à une simple accumulation de faits. Plus un étudiant a des représentations critiques sur les processus de construction et de validation des connaissances, plus il reconnaît que les sources des connaissances enseignées sont à évaluer et que les justifications qui les fondent sont à analyser, et moins il les prend pour argent comptant. Lorsqu'elles sont critiques, les croyances épistémiques sont appelées « évaluativistes ». En bref, celles-ci traduisent une forme d'esprit critique des étudiants concernant les connaissances enseignées.

Les difficultés perçues par les étudiants

Les difficultés perçues (Cameron & Rideout, 2020) correspondent à une première mesure alternative de l'échec. L'entrée à l'université implique des changements profonds pour les étudiants, à la fois sur le plan académique et le plan social. Ils doivent s'adapter à un nouvel environnement et peuvent se heurter à un ensemble de difficultés, lesquelles ont potentiellement un impact sur leurs études. Ainsi, ces difficultés constituent à la fois une mesure de l'adaptation des étudiants à cet environnement mais aussi un facteur de non adaptation. Dans la présente étude, nous considérons les difficultés telles qu'elles sont perçues par les étudiants et nous les analysons comme des révélateurs de l'échec. Plus précisément, nous distinguons les difficultés d'orientation (choix du parcours ou des options), les difficultés pour comprendre les consignes, les difficultés liées à la complexité des contenus, les

difficultés à suivre le cours, les difficultés à en saisir les enjeux, les difficultés en termes de méthodes d'apprentissage et les difficultés d'organisation du travail en dehors du cours.

L'amotivation

L'amotivation (Ryan & Deci, 2000) correspond à une seconde mesure alternative de l'échec. Elle signifie l'absence de motivation ou la résignation. Plus un.e étudiant.e est amotivé.e, plus elle.il a le sentiment de perdre son temps dans sa formation, plus elle.il est résigné.e, et moins les études font sens pour elle.lui. Cette situation que peut vivre un individu au cours d'une activité spécifique, ou d'une situation plus globale, ou parfois même sur l'ensemble de sa vie, est traditionnellement perçue comme un échec (je n'y arrive pas), voire comme un renoncement (j'en n'ai plus envie / ça suffit), elle peut aussi traduire l'absence de confiance en ses capacités (ce n'est pas fait pour moi).

3. Les résultats concernant les facteurs de réussite et d'échec

3.1 Les facteurs qui influencent la validation du premier semestre

La réussite est d'abord appréhendée *via* la validation du premier semestre (S1). Elle est déterminée par une note moyenne aux examens du S1 supérieure ou égale à 10/20. Dans notre étude, le taux des étudiants ayant eu une note supérieure ou égale à 10 est de 0.62, celui des étudiants ayant eu une note inférieure à 10 est de 0.38.

Un modèle de RF construit avec 48 variables prédit la validation du S1 avec un taux de bonnes prédictions de 0.77. Dans ce modèle, le pourcentage de contribution de chacune des variables est donné dans la figure 1 ci-après. On observe que les facteurs principaux de la validation du S1 (en jaune dans la figure 1) sont dans l'ordre décroissant d'importance :

- la composante de la formation ;
- la mention au Bac ;
- et la profession des parents.

Suivant la méthode du coude, on peut également distinguer un ensemble de facteurs secondaires de la validation du S1 (en bleu dans la figure 1) :

- l'amotivation ;
- l'approche de l'apprentissage intégrative ;
- l'approche de l'apprentissage interactions étudiantes ;
- le soutien des enseignants ;
- et la motivation intrinsèque à l'accomplissement.

À noter que les pourcentages de contribution de ces variables secondaires à la prédiction sont très proches, ce qui signifie que l'ordre de leur importance n'est pas significatif.

Pour éclairer les résultats de l'analyse par RF, nous présentons de quelle manière les trois facteurs principaux identifiés influencent la validation du S1, en considérant les modalités respectives de ces trois facteurs.

Le tableau 1 donne à voir de quelle manière la composante constitue un facteur prédictif de la validation du S1. Les différences entre composantes peuvent être dues à la sélection des étudiants à l'entrée de certaines formations, au nombre d'étudiants dans la promotion, au taux d'encadrement par étudiant, mais aussi aux méthodes d'enseignement.

Tableau 1. Validation du S1 en fonction de la composante de la formation

Validation du S1	Composante de la formation									
	FDE	FDS	FDSP	FE	IUTB	IUTMS	IUTN	MOMA	Polytech	STAPS
Validé (taux)	0.96	0.38	0.49	0.35	0.86	0.79	0.69	0.69	0.75	0.58
Non validé (taux)	0.04	0.62	0.51	0.65	0.14	0.21	0.31	0.31	0.25	0.42

FDE : Faculté d'éducation ; FDS : Faculté des Sciences ; FDSEP : Faculté d'économie et des sciences politiques ; FE : Faculté d'économie ; IUTB, MS, N : IUT de Béziers, de Montpellier-Sète et de Nîmes ; MOMA : Montpellier Management ; Polytech : Ecole d'ingénieur Polytech ; STAPS : Faculté des sciences et techniques des activités physiques et sportives

Le tableau 2 montre que plus la mention au Bac est élevée, plus le taux de réussite est élevé. Le point de flexion se situe au niveau des étudiants détenteurs d'une mention Assez Bien. Pour les étudiants sans mention, le taux de réussite se réduit à 0.35.

Tableau 2. Validation du S1 en fonction de la mention au Bac

Validation du S1	Mention au Bac			
	Sans mention	Assez bien	Bien	Très bien
Validé (taux)	0.35	0.53	0.72	0.86
Non validé (taux)	0.65	0.47	0.28	0.14

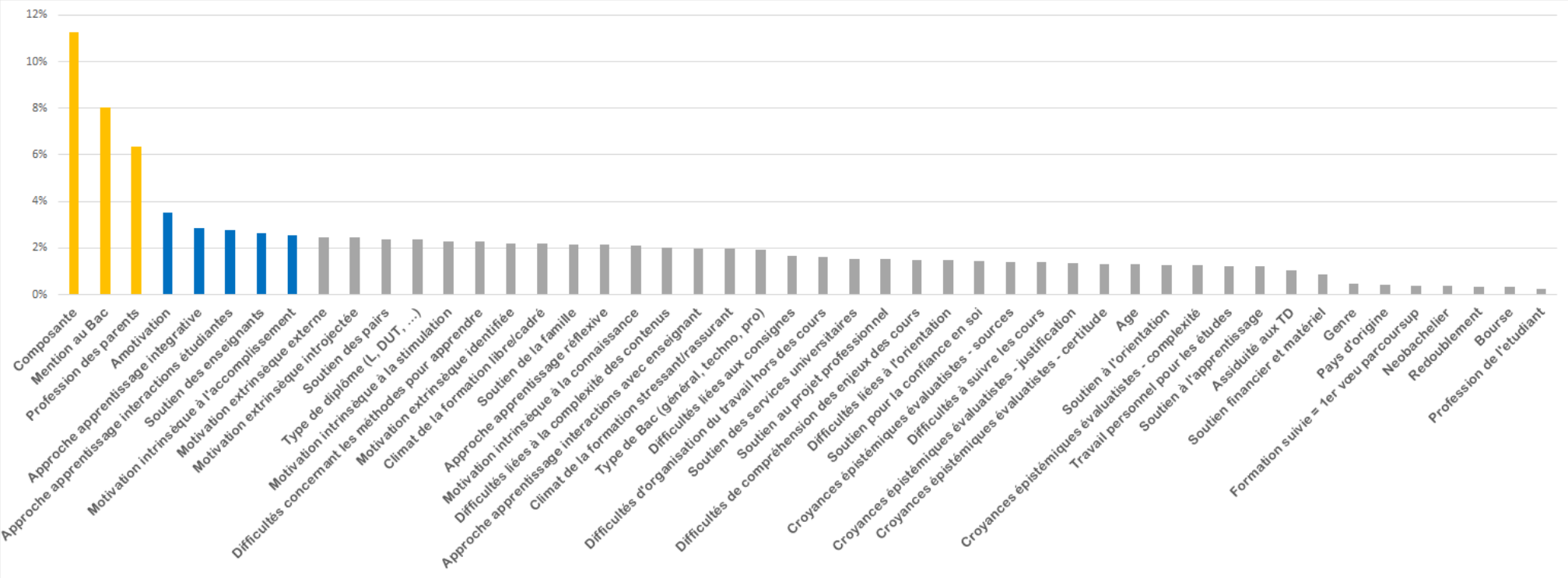
Dans le tableau 3, on observe que la proportion des étudiants qui valident le S1 est la plus forte pour ceux dont les deux parents appartiennent au groupe « cadre/profession intermédiaire »¹ (lequel inclut notamment les professions libérales, cadres, ingénieurs, médecins, enseignants, professions intermédiaires, techniciens et infirmiers). À l'opposé la proportion des étudiants qui ne valident pas le S1 est la plus élevée pour ceux dont un parent appartient au groupe « chef d'entreprise » (lequel inclut notamment les chefs d'entreprise, exploitants de l'agriculture, artisans et commerçants) ou est inactif.

Tableau 3. Validation du S1 en fonction de la profession des parents

Profession des parents	S1 validé oui	S1 validé non
1 cadre/intermédiaire	65,31%	34,69%
1 cadre/intermédiaire et 1 chef d'entreprise	58,77%	41,23%
1 cadre/intermédiaire et 1 employé/ouvrier	61,74%	38,26%
1 cadre/intermédiaire et 1 inactif	51,79%	48,21%
1 chef d'entreprise	33,33%	66,67%
1 chef d'entreprise et 1 employé/ouvrier	62,18%	37,82%
1 chef d'entreprise et 1 inactif	50,94%	49,06%
1 employé/ouvrier	53,57%	46,43%
1 employé/ouvrier et 1 inactif	51,35%	48,65%
1 inactif	36,00%	64,00%
2 cadres/intermédiaires	70,88%	29,12%
2 chefs d'entreprise	60,71%	39,29%
2 employés/ouvriers	59,57%	40,43%
2 inactifs	48,78%	51,22%
inconnu	48,21%	51,79%

¹ Pour simplifier les analyses (où sont à considérer les combinaisons de professions des deux parents), nous avons rassemblé certains « groupes socio-professionnels » tels que définis par l'INSEE en des groupes plus larges.

Figure 1. Contribution de 48 variables à la prédiction de la validation au S1 selon une analyse par RF (en pourcentages à partir de la normalisation de la décroissance moyenne de l'indice de Gini associée à chaque variable)



3.2 Les facteurs qui influencent la motivation autodéterminée

Une manière alternative de mesurer la réussite des étudiants consiste à déterminer le taux d'étudiants qui, dans leurs études, sont animés par une motivation autodéterminée. Pour mener une analyse par RF et ainsi identifier les facteurs qui influencent cette mesure, nous avons pris la moyenne de la motivation autodéterminée (moyenne de 4.94 sur une échelle de 1 à 7) comme valeur de séparation et distingué ainsi deux groupes d'étudiants : ceux pour lesquels la motivation autodéterminée est « présente » et ceux pour lesquels elle est « absente ».

Un modèle de RF construit avec 41 variables prédit la motivation autodéterminée avec un taux de bonnes prédictions de 0.74. Dans ce modèle, le pourcentage de contribution de chacune des variables est donné dans la figure 2 ci-après. On observe que les facteurs principaux de la motivation autodéterminée (en jaune dans la figure 2) sont dans l'ordre décroissant d'importance :

- l'approche de l'apprentissage intégrative ;
- et l'approche de l'apprentissage réflexive.

Suivant la méthode du coude, on peut également distinguer un ensemble de facteurs secondaires de la motivation auto-déterminée (en bleu dans la figure 2) :

- la profession des parents ;
- l'approche de l'apprentissage interactions avec l'enseignant ;
- l'approche de l'apprentissage interactions avec les étudiants ;
- la composante de la formation ;
- et les croyances épistémiques évaluatistes concernant la complexité des connaissances.

À noter que les pourcentages de contribution de ces variables secondaires à la prédiction sont très proches, ce qui signifie que l'ordre de leur importance n'est pas significatif.

Pour éclairer les résultats de l'analyse par RF, nous présentons de quelle manière les deux facteurs principaux identifiés influencent la motivation autodéterminée, en distinguant quatre groupes d'étudiants issus de la population des répondants associés à ces facteurs (quartiles).

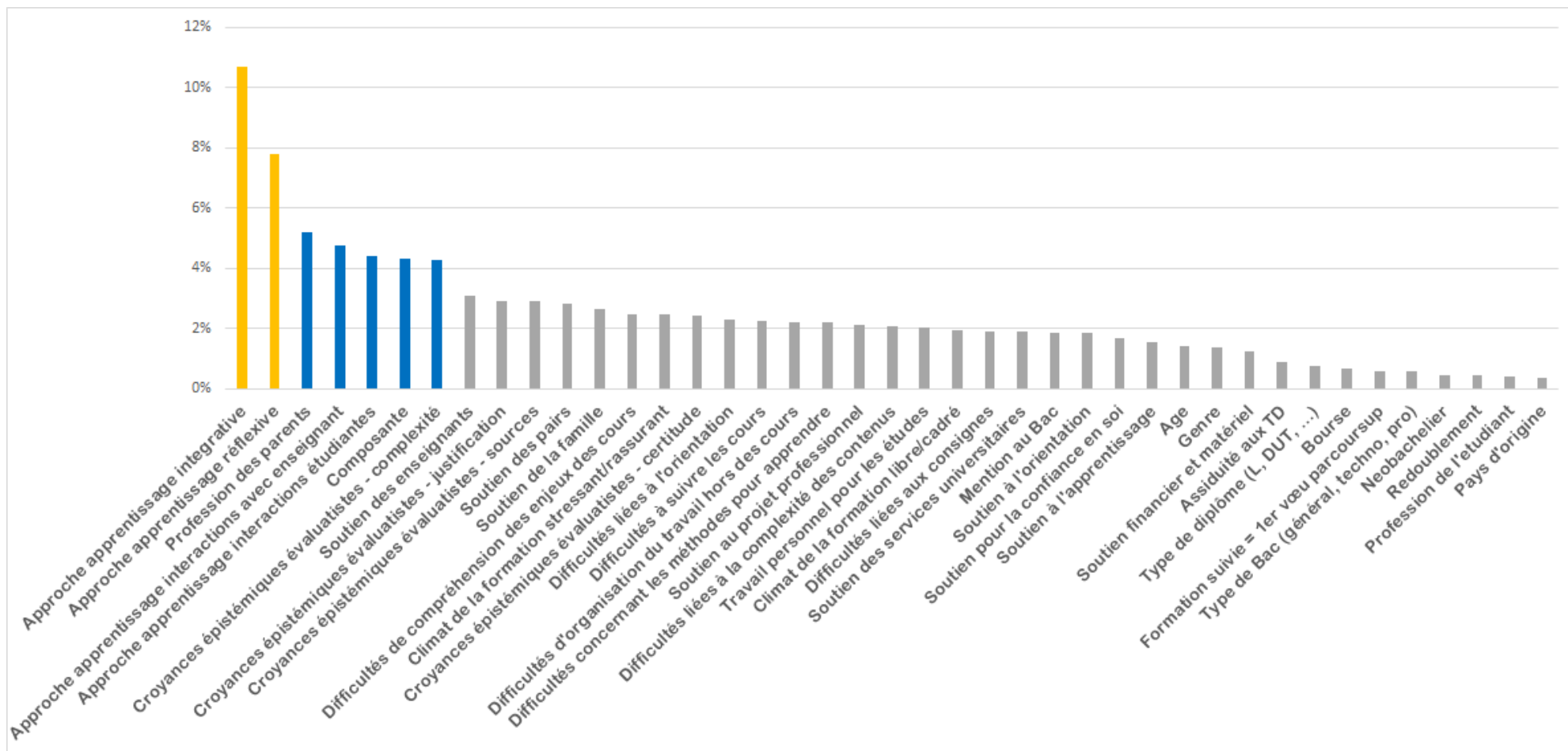
Tableau 4. Motivation autodéterminée en fonction de l'approche de l'apprentissage intégrative

Motivation autodéterminée	Approche de l'apprentissage intégrative			
	Faible 1 ^{er} quartile : ≤3.8	Plutôt faible 2 ^e quartile :]3.8 ; 4.6]	Plutôt forte 3 ^e quartile :]4.6 ; 5.4]	Forte 4 ^e quartile :]5.4 ; 7]
Présente (taux)	0.18	0.40	0.67	0.85
Absente (taux)	0.82	0.60	0.33	0.15

Tableau 5. Motivation autodéterminée en fonction de l'approche de l'apprentissage réflexive

Motivation autodéterminée	Approche de l'apprentissage réflexive			
	Faible 1 ^{er} quartile : ≤4.75	Plutôt faible 2 ^e quartile :]4.75 ; 5.5]	Plutôt forte 3 ^e quartile :]5.5 ; 6]	Forte 4 ^e quartile :]6 ; 7]
Présente (taux)	0.22	0.47	0.68	0.79
Absente (taux)	0.78	0.53	0.32	0.21

Figure 2. Contribution de 41 variables à la prédiction de la motivation autodéterminée selon une analyse par RF (en pourcentages à partir de la normalisation de la décroissance moyenne de l'indice de Gini associée à chaque variable)



3.3 Les facteurs qui influencent l'approche de l'apprentissage en profondeur

Une troisième manière de mesurer la réussite des étudiants consiste à déterminer le taux d'étudiants qui, dans leurs études, suivent une approche de l'apprentissage en profondeur. Pour mener une analyse par RF et ainsi identifier les facteurs qui influencent cette mesure, nous avons pris la moyenne de l'approche de l'apprentissage en profondeur (moyenne de 4,68 sur une échelle de 1 à 7) comme valeur de séparation et distingué ainsi deux groupes d'étudiants : ceux pour lesquels l'approche de l'apprentissage en profondeur est « présente » et ceux pour lesquels elle est « absente ».

Un modèle de RF construit avec 44 variables prédit l'approche de l'apprentissage en profondeur avec un taux de bonne prédiction de 0.78. Dans ce modèle, le pourcentage de contribution de chacune des variables est donné dans la figure 3 ci-après. On observe que les facteurs principaux de l'approche de l'apprentissage en profondeur (en jaune dans la figure 3) sont dans l'ordre décroissant d'importance :

- la motivation intrinsèque à la connaissance ;
- la motivation intrinsèque à la stimulation ;
- et la motivation intrinsèque à l'accomplissement.

Suivant la méthode du coude, on peut également distinguer un ensemble de facteurs secondaires de l'approche de l'apprentissage en profondeur (en bleu dans la figure 3) :

- la motivation extrinsèque identifiée ;
- et les croyances épistémiques évaluatistes concernant les sources des connaissances.

À noter que les pourcentages de contribution de ces variables secondaires à la prédiction sont très proches, ce qui signifie que l'ordre de leur importance n'est pas significatif.

Pour éclairer les résultats de l'analyse par RF, nous présentons de quelle manière les trois facteurs principaux identifiés influencent la motivation autodéterminée, en distinguant quatre groupes d'étudiants issus de la population des répondants associés à ces facteurs (quartiles).

Tableau 6. Approche de l'apprentissage en profondeur en fonction de la motivation intrinsèque à la connaissance

Approche de l'apprentissage en profondeur	Motivation intrinsèque à la connaissance			
	Faible 1 ^{er} quartile : ≤4.75	Plutôt faible 2 ^e quartile :]4.75 ; 5.75]	Plutôt forte 3 ^e quartile :]5.75 ; 6.5]	Forte 4 ^e quartile :]6.5 ; 7]
Présente (taux)	0.19	0.44	0.66	0.84
Absente (taux)	0.81	0.56	0.34	0.16

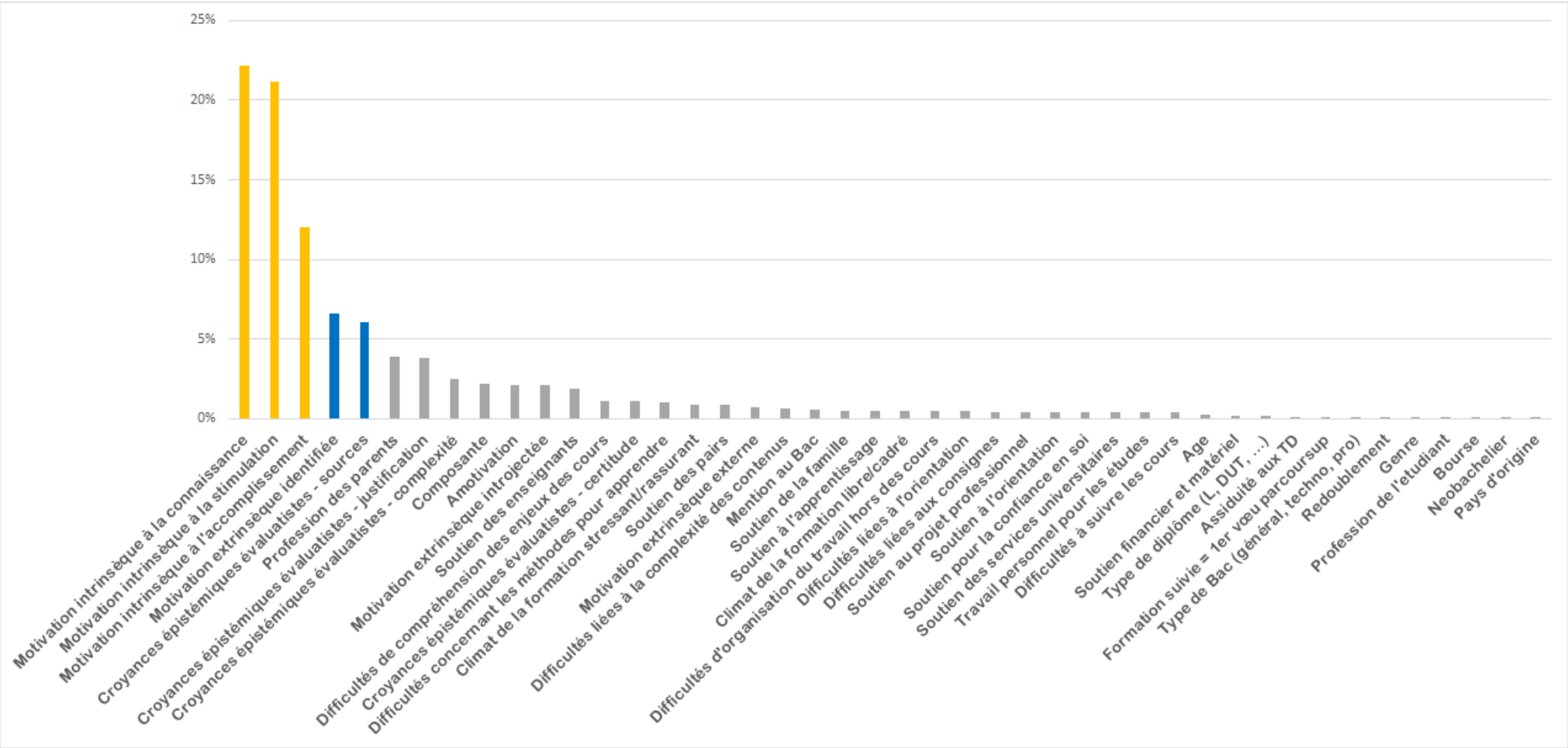
Tableau 7. Approche de l'apprentissage en profondeur en fonction de la motivation intrinsèque à la stimulation

Approche de l'apprentissage en profondeur	Motivation intrinsèque à la stimulation			
	Faible 1 ^{er} quartile : ≤2.75	Plutôt faible 2 ^e quartile :]2.75 ; 3.75]	Plutôt forte 3 ^e quartile :]3.75 ; 4.75]	Forte 4 ^e quartile :]4.75 ; 7]
Présente (taux)	0.22	0.41	0.66	0.84
Absente (taux)	0.78	0.59	0.34	0.16

Tableau 8. Approche de l'apprentissage en profondeur en fonction de la motivation intrinsèque à l'accomplissement

Approche de l'apprentissage en profondeur	Motivation intrinsèque à l'accomplissement			
	Faible 1 ^{er} quartile : ≤3.5	Plutôt faible 2 ^e quartile :]3.5 ; 4.75]	Plutôt forte 3 ^e quartile :]4.75 ; 5.75]	Forte 4 ^e quartile :]5.75 ; 7]
Présente (taux)	0.25	0.38	0.64	0.83
Absente (taux)	0.75	0.62	0.36	0.17

Figure 3. Contribution de 44 variables à la prédiction de l'approche de l'apprentissage en profondeur selon une analyse par RF (en pourcentages à partir de la normalisation de la décroissance moyenne de l'indice de Gini associée à chaque variable)



3.4 Les facteurs qui influencent les croyances épistémiques évaluateuses

Une quatrième manière de mesurer la réussite des étudiants consiste à déterminer le taux d'étudiants qui, dans leurs études, ont des croyances épistémiques évaluateuses. Pour mener une analyse par RF et ainsi identifier les facteurs qui influencent cette mesure, nous avons pris la moyenne des croyances épistémiques évaluateuses (moyenne de 5,13 sur une échelle de 1 à 7) comme valeur de séparation et distingué ainsi deux groupes d'étudiants : ceux pour lesquels les croyances épistémiques évaluateuses sont « présentes » et ceux pour lesquels elles sont « absentes ».

Un modèle de RF construit avec 44 variables prédit les croyances épistémiques évaluateuses avec un taux de bonne prédiction de 0.71. Dans ce modèle, le pourcentage de contribution de chacune des variables est donné dans la figure 4 ci-après. On observe que les facteurs principaux des croyances épistémiques évaluateuses (en jaune dans la figure 4) sont dans l'ordre décroissant d'importance :

- l'approche de l'apprentissage intégrative ;
- la motivation intrinsèque à la connaissance ;
- et la profession des parents.

Suivant la méthode du coude, on peut également distinguer un ensemble de facteurs secondaires des croyances épistémiques évaluateuses (en bleu dans la figure 4) :

- l'approche de l'apprentissage réflexive ;
- la composante de la formation ;
- la motivation extrinsèque identifiée ;
- et la motivation intrinsèque à la stimulation.

À noter que les pourcentages de contribution de ces variables secondaires à la prédiction sont très proches, ce qui signifie que l'ordre de leur importance n'est pas significatif.

Pour éclairer les résultats de l'analyse par RF, nous présentons de quelle manière les trois facteurs principaux identifiés influencent les croyances épistémiques évaluateuses, en distinguant quatre groupes d'étudiants issus de la population des répondants associés à ces facteurs (quartiles).

Tableau 9. Croyances épistémiques évaluateuses en fonction de l'approche de l'apprentissage intégrative

Croyances épistémiques évaluateuses	Approche de l'apprentissage intégrative			
	Faible 1 ^{er} quartile : ≤3.8	Plutôt faible 2 ^e quartile :]3.8 ; 4.6]	Plutôt forte 3 ^e quartile :]4.6 ; 5.4]	Forte 4 ^e quartile :]5.4 ; 7]
Présente (taux)	0.27	0.42	0.61	0.79
Absente (taux)	0.73	0.58	0.39	0.21

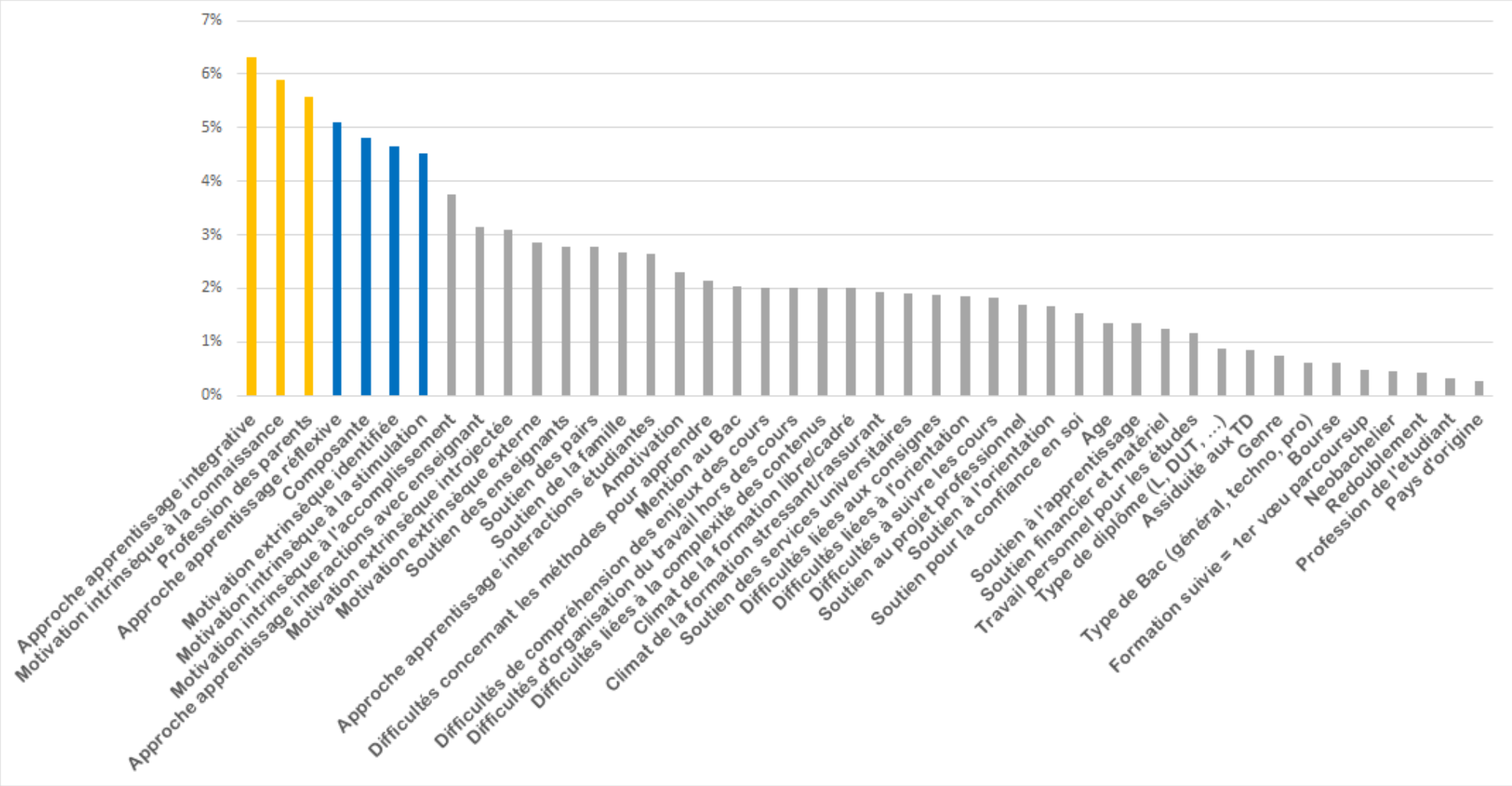
Tableau 10. Croyances épistémiques évaluateuses en fonction de la motivation intrinsèque à la connaissance

Croyances épistémiques évaluateuses	Motivation intrinsèque à la connaissance			
	Faible 1 ^{er} quartile : ≤4.75	Plutôt faible 2 ^e quartile :]4.75 ; 5.75]	Plutôt forte 3 ^e quartile :]5.75 ; 6.5]	Forte 4 ^e quartile :]6.5 ; 7]
Présente (taux)	0.30	0.43	0.63	0.76
Absente (taux)	0.70	0.57	0.37	0.24

Tableau 11. Croyances épistémiques évaluateurs en fonction de la profession des parents

Profession des parents	Évaluateur oui	Évaluateur non
1 cadre/intermédiaire	42,86%	57,14%
1 cadre/intermédiaire et 1 chef d'entreprise	54,39%	45,61%
1 cadre/intermédiaire et 1 employé/ouvrier	48,02%	51,98%
1 cadre/intermédiaire et 1 inactif	55,36%	44,64%
1 chef d'entreprise	75,00%	25,00%
1 chef d'entreprise et 1 employé/ouvrier	45,38%	54,62%
1 chef d'entreprise et 1 inactif	50,94%	49,06%
1 employé/ouvrier	35,71%	64,29%
1 employé/ouvrier et 1 inactif	56,76%	43,24%
1 inactif	40,00%	60,00%
2 cadres/intermédiaires	53,92%	46,08%
2 chefs d'entreprise	67,86%	32,14%
2 employés/ouvriers	47,39%	52,61%
2 inactifs	58,54%	41,46%
inconnu	44,64%	55,36%

Figure 4. Contribution de 44 variables à la prédiction des croyances épistémiques évaluatistes selon une analyse par RF (en pourcentages à partir de la normalisation de la décroissance moyenne de l'indice de Gini associée à chaque variable)



3.5 Les facteurs qui influencent les difficultés perçues

Une manière alternative de mesurer l'échec des étudiants consiste à déterminer le taux d'étudiants qui, dans leurs études, estiment rencontrer des difficultés. Pour mener une analyse par RF et ainsi identifier les facteurs qui influencent cette mesure, nous avons pris la moyenne des difficultés perçues (moyenne 3.33 de sur une échelle de 1 à 7) comme valeur de séparation et distingué ainsi deux groupes d'étudiants : ceux pour lesquels les difficultés perçues sont « présentes » et ceux pour lesquels elles sont « absentes ».

Un modèle de RF construit avec 44 variables permet de prédire les difficultés perçues avec un taux de bonne prédiction de 0.67. Dans ce modèle, le pourcentage de contribution de chacune des variables est donné dans la figure 5 ci-après. On observe que le facteur principal des difficultés perçues (en jaune dans la figure 5) est :

- l'amotivation.

Suivant la méthode du coude, on peut également distinguer un ensemble de facteurs secondaires des difficultés perçues (en bleu dans la figure 5) :

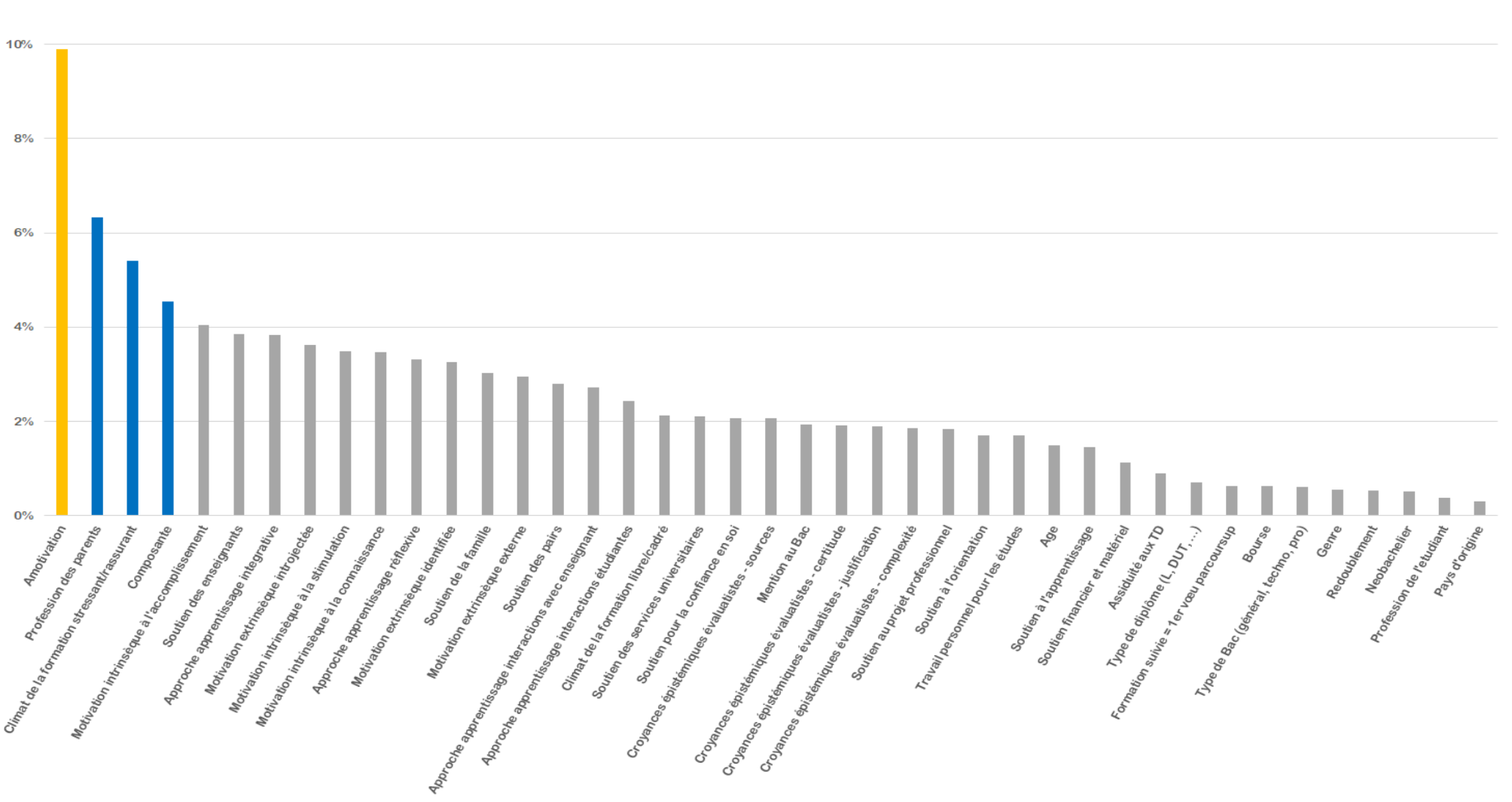
- la profession des parents ;
- le climat de la formation perçu comme plutôt stressant ou rassurant ;
- et la composante de la formation suivie.

Pour éclairer les résultats de l'analyse par RF, nous présentons de quelle manière le facteur principal influence les difficultés perçues, en distinguant quatre groupes d'étudiants issus de la population des répondants associés à ces facteurs (quartiles).

Tableau 12. Difficultés perçues en fonction de l'amotivation

Difficultés perçues	Amotivation			
	Faible 1 ^{er} quartile : ≤1	Plutôt faible 2 ^e quartile :]1 ; 1.25]	Plutôt forte 3 ^e quartile :]1.25 ; 2.25]	Forte 4 ^e quartile :]2.25 ; 7]
Présente (taux)	0.25	0.46	0.57	0.80
Absente (taux)	0.75	0.54	0.43	0.20

Figure 5. Contribution de 41 variables à la prédiction des difficultés perçues selon une analyse par RF (en pourcentages à partir de la normalisation de la décroissance moyenne de l'indice de Gini associée à chaque variable)



3.6 Les facteurs qui influencent l'amotivation

Une autre possibilité de mesurer l'échec des étudiants consiste à déterminer le taux d'étudiants qui, dans leurs études, sont amotivés. Pour mener une analyse par RF et ainsi identifier les facteurs qui influencent cette mesure, nous avons pris la moyenne de l'amotivation (moyenne 1.83 de sur une échelle de 1 à 7) comme valeur de séparation et distingué ainsi deux groupes d'étudiants : ceux pour lesquels l'amotivation est « absente » et ceux pour lesquels elle est « présente ».

Un modèle de RF construit avec 44 variables prédit l'amotivation avec un taux de bonne prédiction de 0.78. Dans ce modèle, le pourcentage de contribution de chacune des variables est donné dans la figure 6 ci-après. On observe que les facteurs principaux de l'amotivation (en jaune dans la figure 6) sont dans l'ordre décroissant d'importance :

- les difficultés perçues liées à l'orientation (choix du parcours ou des options) ;
- et la profession des parents.

Suivant la méthode du coude, on peut également distinguer un ensemble de facteurs secondaires de l'amotivation (en bleu dans la figure 6) :

- la composante de la formation suivie ;
- l'approche de l'apprentissage réflexive ;
- l'approche de l'apprentissage intégrative ;
- et les difficultés perçues liées à la complexité des contenus.

Pour éclairer les résultats de l'analyse par RF, nous présentons de quelle manière ces trois variables influencent l'amotivation, en distinguant quatre groupes d'étudiants issus de la population des répondants associés à ces variables (quartiles).

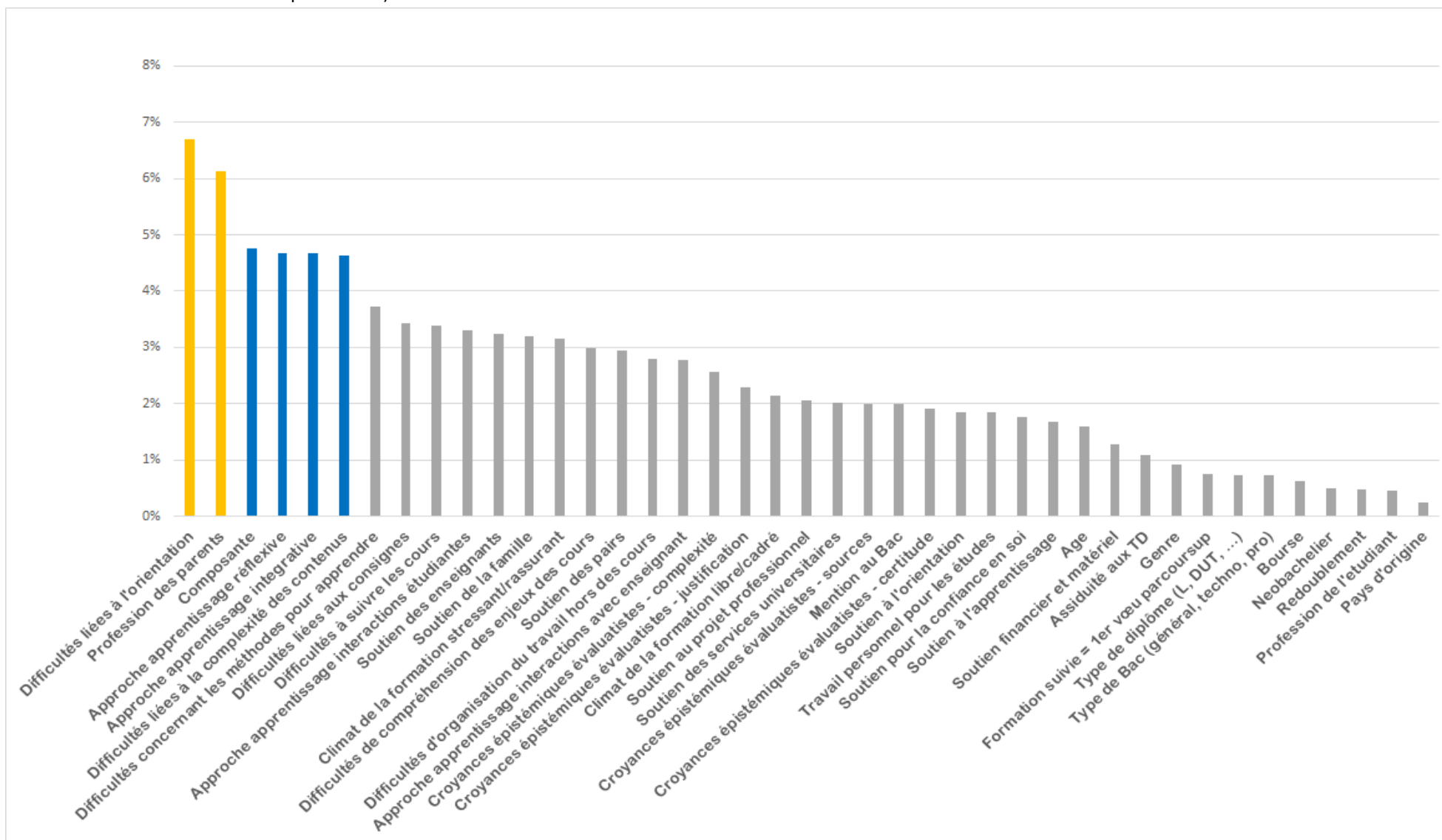
Tableau 13. L'amotivation en fonction des difficultés à l'orientation

Amotivation	Difficultés à l'orientation			
	Faible 1 ^{er} quartile : ≤1	Plutôt faible 2 ^e quartile :]1 ; 2]	Plutôt forte 3 ^e quartile :]2 ; 4]	Forte 4 ^e quartile :]4 ; 7]
Présente (taux)	0.14	0.24	0.46	0.59
Absente (taux)	0.86	0.76	0.54	0.41

Tableau 14. L'amotivation en fonction de la profession des parents

Profession des parents	Amotivé oui	Amotivé non
1 cadre/intermédiaire	38,78%	61,22%
1 cadre/intermédiaire et 1 chef d'entreprise	32,46%	67,54%
1 cadre/intermédiaire et 1 employé/ouvrier	30,87%	69,13%
1 cadre/intermédiaire et 1 inactif	36,61%	63,39%
1 chef d'entreprise	41,67%	58,33%
1 chef d'entreprise et 1 employé/ouvrier	36,13%	63,87%
1 chef d'entreprise et 1 inactif	35,85%	64,15%
1 employé/ouvrier	42,86%	57,14%
1 employé/ouvrier et 1 inactif	30,41%	69,59%
1 inactif	48,00%	52,00%
2 cadres/intermédiaires	30,24%	69,76%
2 chefs d'entreprise	39,29%	60,71%
2 employés/ouvriers	31,30%	68,70%
2 inactifs	39,02%	60,98%
inconnu	33,93%	66,07%

Figure 6. Contribution de 41 variables à la prédiction de l'amotivation selon une analyse par RF (en pourcentages à partir de la normalisation de la décroissance moyenne de l'indice de Gini associée à chaque variable)



4. Discussion concernant les facteurs de réussite et d'échec

4.1 La précision de la prédiction de la réussite et de l'échec

Nous constatons d'abord que les différentes formes de réussite et d'échec ont pu être prédites avec une précision relativement élevée, comparativement à de précédentes études (Cassidy, 2012; Sæle et al., 2017; Pinxten et al., 2015). Cette performance prédictive peut s'expliquer par le fait que les modèles développés ont intégré une multiplicité de facteurs : socio-académiques, psycho- et socio-cognitifs, relevant de l'engagement dans les études ainsi que de la formation suivie. Ces facteurs, bien qu'ils soient représentés par des variables de différents types (quantitatives et nominales), ont pu être pris en compte conjointement dans ces modèles grâce au recours à des RF. À noter cependant que l'une des formes d'échec, à savoir les difficultés perçues, a été prédite avec une précision moindre. Le modèle développé pour prédire cette mesure d'échec est donc moins complet. Des recherches supplémentaires méritent d'être réalisées pour mieux identifier les origines des difficultés perçues par les étudiants.

4.2 Les facteurs de réussite en termes de validation du 1^{er} semestre

Considérons ensuite la réussite au sens traditionnel, celle associée aux performances académiques et ici mesurée par la validation du 1^{er} semestre. Les résultats montrent que le facteur principal de cette forme de réussite est la composante de la formation suivie. Ce facteur pourrait être décomposé en un ensemble de facteurs plus spécifiques caractérisant la formation : non seulement la discipline, dont le rôle sur les performances a déjà été mis en évidence par plusieurs études (Clerici et al., 2015 ; Fokkens-Bruinsma et al., 2021), mais aussi le processus de sélection des étudiants à l'entrée de la formation, le taux d'encadrement des étudiants, le nombre d'étudiants dans la promotion, ainsi que les méthodes pédagogiques. Cette première étude n'a pas considéré ces caractéristiques individuellement. Il serait intéressant de mener de nouvelles études pour examiner leur poids respectif dans la prédiction des performances académiques.

Les résultats confirment de précédentes études montrant que les performances antérieures (au lycée), traduites ici par la mention au Bac, sont également un prédicteur important des performances académiques à l'université (Clerici et al., 2015 ; Diseth et al., 2010 ; Sothan, 2019). On peut remarquer que les formations qui sélectionnent les étudiants sur la base des performances au lycée augmentent ainsi mécaniquement la probabilité d'obtenir un taux de réussite élevé en termes de validation du semestre ou de l'année.

Le niveau socio-culturel mesuré par la profession des parents ressort comme un troisième facteur important, résultat en accord avec plusieurs études précédentes (Brinbaum et al., 2018 ; Pinxten et al., 2015 ; Sothan, 2019). Des étudiants dont les parents sont cadres ont ainsi deux fois plus de chance de valider le semestre que des étudiants dont un parent est inactif.

En outre, les résultats montrent que l'approche de l'apprentissage en profondeur est un facteur secondaire de la validation du semestre, mais non un facteur principal. Cela signifie que les évaluations ne récompensent pas toujours les étudiants qui adoptent une approche en profondeur, comme cela a déjà été pointé antérieurement (Bruinsma, 2004).

Les évaluations récompensent encore bien moins l'engagement comportemental des étudiants (mesuré par l'assiduité aux TD et le temps alloué aux études en dehors des cours), lequel se révèle être un facteur relativement faible, ce qui s'accorde également avec une précédente recherche (Fokkens-Bruinsma et al., 2021).

D'après notre étude, les données personnelles (âge, genre et nationalité) n'ont également qu'une faible influence sur la validation au semestre. Concernant la nationalité, il convient de noter toutefois

que la population des étudiants ne comportait qu'une faible proportion d'étrangers (6%), ce qui incite à prendre ce résultat avec prudence.

4.3 Comparaison des facteurs des différentes formes de réussite et d'échec

À présent, considérons également les autres formes de réussite et d'échec, celles utilisées pour mesurer la curiosité, l'autonomie, la réflexivité et l'esprit critique des étudiants. Comparons l'influence de l'ensemble des facteurs pris en compte dans l'étude. Les résultats montrent que les facteurs principaux et secondaires ne sont pas identiques pour toutes les formes de réussite et d'échec (un tableau synthétique se trouve dans l'annexe F).

Cependant, certains facteurs ressortent comme plus déterminants que d'autres. Ainsi, la profession des parents est l'un des facteurs principaux de trois formes de réussite et d'échec (la validation du semestre, les croyances épistémiques évaluatistes et l'amotivation) et l'un des facteurs secondaires de deux autres formes de réussite et d'échec (la motivation auto-déterminée et les difficultés perçues). Ce résultat met en évidence le déterminisme socio-culturel relativement fort qui s'exerce en première année d'université.

L'approche de l'apprentissage intégrative est l'un des facteurs principaux de deux formes de réussite (la motivation auto-déterminée et les croyances épistémiques évaluatistes) et l'un des facteurs secondaires de deux autres formes de réussite et d'échec (la validation du semestre et l'amotivation). Rappelons que les étudiants qui adoptent cette approche cherchent à mettre en lien les contenus des cours, à les explorer davantage et à les appliquer dans de nouvelles situations. Contrairement à la profession des parents, l'approche intégrative constitue un levier potentiel pour les enseignants : favoriser une telle approche de l'apprentissage est un moyen de mettre les étudiants sur le chemin de différentes formes de réussite.

La motivation intrinsèque à la connaissance est l'un des facteurs principaux de deux formes de réussite (l'approche de l'apprentissage en profondeur et les croyances épistémiques). Les étudiants qui adoptent ce type d'attitude à l'égard des connaissances révèlent un rapport particulier avec les connaissances autant dans le plaisir qu'elles procurent que dans les questionnements qu'elles stimulent. Ce rapport aux connaissances peut être associé à la curiosité évoquée plus haut.

Comme souligné précédemment, la composante de la formation suivie est l'un des facteurs principaux d'une forme de réussite (la validation du semestre), mais aussi l'un des facteurs secondaires de quatre autres formes de réussite et d'échec (la motivation auto-déterminée, les croyances épistémiques évaluatistes, les difficultés perçues et l'amotivation). Il se peut que les formations influencent plus ou moins le développement des caractéristiques psycho-cognitives que recouvrent ces quatre formes de réussite et d'échec. Toutefois, les données de l'étude ont été recueillies deux mois seulement après le début des formations, de sorte qu'on ne peut pas exclure que ces différences entre formations soient liées au processus de sélection des étudiants à l'entrée dans la formation. Le rôle des formations pourra être déterminé de façon plus précise à travers la poursuite de cette étude en 2^e et 3^e années du premier cycle universitaire.

Les résultats montrent aussi que certains facteurs sont moins déterminants que d'autres. Ainsi, les performances antérieures au lycée, mesurées par la mention au Bac, ne sont un facteur important que de la validation du semestre. Ce facteur apparaît beaucoup moins déterminant pour toutes les autres formes de réussite et d'échec. Cela signifie que valider le semestre n'est pas un gage de réussite en termes de développement de la curiosité, de l'autonomie, de la réflexivité et de l'esprit critique.

Parmi les facteurs moins déterminants, on compte également les différentes natures et sources de soutien social. Ce résultat semble ainsi remettre en question certaines études qui pointent

l'importance du soutien (DeBerard et al., 2004), en particulier celui des enseignants et de la famille (Dupont et al., 2015). Soulignons toutefois que les RF de notre étude situent le soutien des enseignants dans la première moitié du classement des facteurs en fonction de leur poids prédictif, et ce, pour toutes les formes de réussite et d'échec. Autrement dit, le soutien des enseignants ne constitue pas un levier majeur pour favoriser les différentes formes de réussite, mais ne doit pas être considéré comme un facteur négligeable pour autant.

L'engagement comportemental des étudiants apparaît comme un facteur encore moins déterminant, quelle que soit la forme de réussite ou d'échec. Ce résultat se trouve en cohérence avec une étude (Fokkens-Bruinsma et al., 2021), mais en désaccord avec plusieurs autres études (Dupont et al., 2015 ; Sothan, 2019 ; Torenbeek et al., 2013). On peut remarquer que chaque étude est réalisée dans le contexte d'une université différente. Les paramètres de ces différents contextes qui font de l'engagement un facteur déterminant ou non de la réussite serait donc à investiguer.

Les données personnelles, à savoir l'âge, le genre et la nationalité, comptent également parmi les facteurs les moins déterminants des différentes formes de réussite et d'échec. De même que pour l'engagement, ce résultat s'accorde avec certaines études et en contredit d'autres (Bruinsma, 2004 ; Clerici et al., 2015 ; Craft, 2019).

Deux autres résultats ressortent de la comparaison des facteurs les plus prédictifs des différentes formes de réussite. Premièrement, la motivation autodéterminée et l'approche en profondeur sont toutes deux des facteurs majeurs l'une pour l'autre. Dans une moindre mesure, cette interdépendance inclut également les croyances épistémiques évaluatistes. Les résultats de cette étude confirment ainsi les liens étroits entre ces trois dimensions psycho-cognitives qui ont été pointées dans de précédentes études (Biggs, 1987 ; Entwistle & Peterson, 2004 ; Muis, 2007). Deuxièmement, les difficultés perçues (en particulier celles liées à l'orientation) et l'amotivation sont également des facteurs majeurs l'une pour l'autre. Les liens étroits entre ces deux dimensions psycho-cognitives mériteraient d'être étudiées davantage.

4. Un modèle de la manière dont les étudiants s'engagent dans leurs études

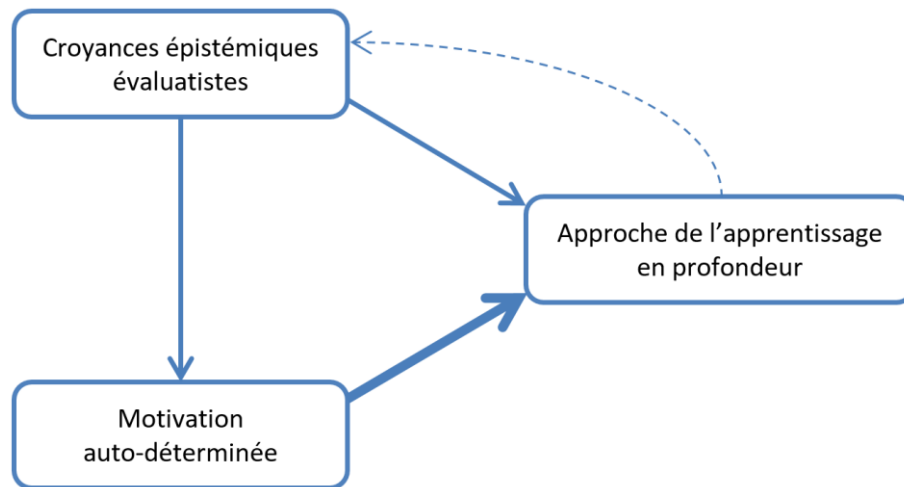
Les résultats présentés ci-dessus pointent notamment l'existence de liens complexes entre l'approche de l'apprentissage (plus ou moins en profondeur), la motivation (plus ou moins auto-déterminée) et les croyances épistémiques (plus ou moins évaluatistes). Afin de mieux comprendre ces liens, d'autres analyses statistiques ont été réalisées, à savoir des analyses en médiation. Celles-ci permettent de proposer un modèle de la manière dont les étudiants s'engagent dans leurs études.

Rappelons la signification des trois variables :

- approche de l'apprentissage en profondeur : engagement dans les apprentissages avec une réflexivité sur les contenus, les activités et ses propres apprentissages ;
- motivation auto-déterminée : comportement animé par une curiosité à l'égard des connaissances et un désir d'autonomie dans les apprentissages ;
- croyances épistémiques évaluatistes : représentations des connaissances faisant preuve d'esprit critique.

Une analyse en médiation intégrant les trois variables montre l'existence d'un lien direct significatif entre les croyances épistémiques et l'approche de l'apprentissage des étudiants, ainsi que d'un lien indirect significatif entre ces deux variables qui est médiatisé par la motivation auto-déterminée. Cette analyse conduit à proposer le modèle présenté dans la figure 7.

Figure 7. Modèle des liens entre l'approche de l'apprentissage en profondeur, la motivation auto-déterminée et les croyances épistémiques évaluatistes basé sur des analyses en médiation



Ce modèle permet d'interpréter les données de la manière suivante :

- Plus les étudiants sont motivés de façon auto-déterminée, plus ils ont une approche de l'apprentissage en profondeur. Voici une illustration possible de ce lien : plus un étudiant est curieux à propos des connaissances enseignées, plus il cherche à réfléchir sur les connaissances enseignées pour mieux les comprendre. D'après les analyses, ce lien direct est relativement fort (coefficient standardisé $\beta = 0.54$).
- Plus les étudiants sont évaluatistes, plus ils ont une approche de l'apprentissage en profondeur. Voici une illustration possible de ce lien : plus un étudiant est conscient que les connaissances enseignées sont complexes et susceptibles d'évoluer, plus il cherche à réfléchir sur celles-ci pour mieux les comprendre. D'après les analyses, ce lien direct est modéré ($\beta = 0.23$).
- De plus, la motivation auto-déterminée joue le rôle de médiateur entre les croyances épistémiques évaluatistes et l'approche de l'apprentissage en profondeur. Voici une illustration possible de ce processus de médiation : plus un étudiant est conscient que les connaissances enseignées sont complexes et susceptibles d'évoluer, plus il est curieux à propos de ces connaissances et plus il cherche à réfléchir sur celles-ci pour mieux les comprendre. D'après les analyses, ce lien indirect est modéré ($\beta = 0.25$).

Ces analyses statistiques mettent ainsi en lumière deux chemins psycho-cognitifs que les étudiants peuvent suivre lorsqu'ils s'engagent dans leurs apprentissages en première année d'université.

Soulignons que l'analyse en médiation n'exclut pas d'autres interprétations des mêmes données. En particulier, il se peut que les corrélations entre les croyances épistémiques évaluatistes et l'approche de l'apprentissage en profondeur expriment une relation d'influence s'exerçant aussi dans l'autre sens (ce qui est représenté dans la figure 7 par la flèche en traits tirés). Voici une illustration possible de cette influence inverse : plus un étudiant cherche à réfléchir sur les connaissances enseignées pour mieux les comprendre, plus il est conscient qu'elles sont complexes et susceptibles d'évoluer.

Les relations entre les trois variables ici considérées montrent que si des actions sont menées pour faire évoluer positivement l'une de ces variables, cela aura aussi possiblement des effets sur les deux autres. Par exemple, si les enseignants mettent en œuvre des activités sur la nature des connaissances et les processus de leur construction et de leur validation, celles-ci pourraient non seulement rendre

les étudiants plus critiques sur les connaissances enseignées, mais aussi accroître leur motivation à les explorer et les inciter à les explorer plus en profondeur.

Références

- Biggs, J. (1987). *Study process questionnaire manual*. Australian Council for Educational Research.
- Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university*. Open University Press.
- Brinbaum, Y., Hugrée, C., & Poullaouec, T. (2018). 50% to the bachelor's degree... but how? Young people from working-class families at university in France. *Economics and Statistics*, 499, 79–105.
- Bruinsma, M. (2004). Motivation, cognitive processing and achievement in higher education. *Learning and Instruction*, 14(6), 549–568.
- Cameron, R., & Rideout, C. A. (2020). 'It's been a challenge finding new ways to learn': first-year students' perceptions of adapting to learning in a university environment. *Studies in Higher Education*, 47(3), 668–682.
- Cassidy, S. (2012). Exploring individual differences as determining factors in student academic achievement in higher education. *Studies in higher education*, 37(7), 793–810.
- Clerici, R., Giraldo, A., & Meggiolaro, S. (2015). The determinants of academic outcomes in a competing risks approach: evidence from Italy. *Studies in Higher Education*, 40(9), 1535–1549.
- Craft, J., (2019). Academic performance of low SES students at an Australian university satellite campus. *Studies in Higher Education*, 44(8), 1372–1385.
- DeBerard, M., Spielmans, G., & Julka, D. (2004). Predictors of academic achievement and retention among college freshman: A longitudinal study. *College Student Journal*, 38(1), 66–80.
- Diseth, A., Pallesen, S., Brunborg, G., & Larsen, S. (2010). Academic achievement among first semester undergraduate psychology students: The role of course experience, effort, motives and learning strategies. *Higher Education*, 59, 335–352.
- Dupont, S. Galand, B., & Nils, F. (2015). The impact of different sources of social support on academic performance: Intervening factors and mediated pathways in the case of master's thesis. *European Review of Applied Psychology*, 65(5), 227–237.
- Entwistle, N., & Peterson, E. (2004). Conceptions of learning and knowledge in education: Relationships with study behaviour and influences of learning environment. *International Journal of Educational Research*, 41, 407-428.
- Fokkens-Bruinsma, M., Vermue, C., Deinum, J-F., & Rooij, E. V. (2021). First-year academic achievement: the role of academic self-efficacy, self-regulated learning and beyond classroom engagement. *Assessment & Evaluation in Higher Education*, 46(7), 1115–1126.
- Hellas, A., Ihtantola, P., Petersen, A., Ajanovski, V., Gutica, M., Hynninen, T., Knutas, A., Leinonen, J., Messom, C., & Nam Liao, S. (2018). Predicting academic performance: A systematic literature review. In Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education (pp. 175–199).
- Hofer, B., & Pintrich, P. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140.
- Ménard, B. (2021). Parcours et réussite en licence: Les résultats de la session 2020. *Note Flash du SIES*, n°24.
- Muis, K. (2007). The role of epistemic beliefs in self-regulated learning. *Educational Psychologist*, 42(3), 173-190.
- Pinxten, M., De Fraine, B., Van Den Noortgate, W., Damne, J., Boonen, T., & Vanlaar, G. (2015). 'I choose so I am': a logistic analysis of major selection in university and successful completion of the first year. *Studies in higher education*, 40(10), 1919–1946.
- Ryan, R., & Deci, E. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, Article 101860.
- Sæle, R., Dahl, T., Sorlie, T., & Friborg, O. (2017). Relationships between learning approach, procrastination and academic achievement amongst first-year university students. *Higher Education*, 74, 757–774.
- Sothan, S. (2019). The determinants of academic performance: Evidence from a Cambodian University. *Studies in higher education*, 44(11), 2096–2111.
- Torenbeek, M., Jansen, E., & Suhre, C. (2013). Predicting undergraduates' academic achievement: The role of the curriculum, time investment and self-related learning. *Studies in higher education*, 38(9), 1393–1406.
- Van der Zanden, P., Denessen, A., Cillessen, A. & Meijer, P. (2019). Patterns of success: First-year student success in multiple domains. *Studies in Higher Education*, 44(11), 2081–2095.

Annexe A – Données sur la représentativité des étudiants ayant répondu

Les trois tableaux suivants présentent des données permettant d'apprécier la représentativité de la population des étudiants ayant répondu en la comparant à celle de la promotion. Les données concernent la population globale, c'est-à-dire réunissant les 10 composantes de l'étude, et les populations relatives à chaque composante.

Tableau A1. Effectifs et pourcentages des répondants

	Global	FDE	FDS	FDSP	FE	IUTB	IUTMS	IUTN	MCMA	Polytech	STAPS
Effectif répondants	2168	172	638	296	59	65	441	148	139	107	103
Effectif promo	7301	291	1638	1785	551	203	914	495	573	169	682
% répondants	29,7%	59,1%	38,9%	16,6%	10,7%	32,0%	48,2%	29,9%	24,3%	63,3%	15,1%

Tableau A2. Age moyen au moment de l'enquête

		Global	FDE	FDS	FDSP	FE	IUTB	IUTMS	IUTN	MCMA	Polytech	STAPS
répondants	Age moyen	18,5	18,6	19,0	18,5	18,8	18,4	18,2	18,5	18,4	17,8	18,4
	Ecart-type	1,3	1,27	2,43	1,70	1,51	1,16	0,80	1,26	1,20	0,53	0,93
promo	Age moyen	18,7	18,5	19,1	18,8	19,1	18,3	18,3	18,5	18,4	17,8	18,5
	Ecart-type	1,64	1,13	2,19	1,79	1,95	1,01	1,14	1,30	1,04	0,51	1,00

Tableau A3. Genre, type de Bac et mention au Bac

		Global	FDE	FDS	FDSP	FE	IUTB	IUTMS	IUTN	MCMA	Polytech	STAPS	
Filles	répondants	58,9%	94,2%	55,2%	78,0%	47,5%	38,5%	56,7%	31,1%	67,6%	47,7%	34,9%	
	promo	50,6%	92,4%	45,6%	73,4%	31,6%	36,0%	46,9%	23,2%	55,1%	49,7%	26,1%	
Bac gén	répondants	80,4%	82,0%	79,8%	86,5%	79,7%	49,2%	75,5%	73,6%	88,5%	100,0%	79,6%	
	promo	79,3%	79,4%	78,3%	84,1%	68,1%	53,7%	74,2%	69,1%	90,8%	100,0%	85,2%	
Bac techno	répondants	12,8%	12,8%	7,1%	5,7%	10,2%	41,5%	22,9%	25,7%	3,6%	0,0%	16,5%	
	promo	12,9%	15,8%	7,9%	7,8%	13,8%	39,4%	24,2%	29,7%	4,5%	0,0%	10,9%	
Bac pro	répondants	1,3%	4,7%	0,8%	2,0%	3,4%	4,6%	0,2%	0,0%	2,2%	0,0%	1,0%	
	promo	2,1%	4,5%	1,3%	2,9%	5,4%	3,0%	0,1%	0,2%	1,6%	0,0%	2,9%	
Autres	répondants	5,5%	0,6%	12,4%	5,7%	6,8%	4,6%	1,1%	0,7%	5,8%	0,0%	1,0%	
	promo	5,7%	0,3%	12,4%	5,1%	12,7%	3,9%	1,5%	1,0%	4,5%	0,0%	1,0%	
Mention	Sans	répondants	18,0%	22,7%	27,0%	15,5%	23,7%	18,5%	6,3%	23,6%	15,1%	0,0%	23,3%
	promo	24,8%	22,3%	31,5%	22,9%	36,7%	16,7%	9,2%	27,9%	24,4%	0,0%	32,4%	
AB	répondants	30,4%	37,8%	28,4%	33,1%	37,3%	40,0%	27,2%	43,2%	33,8%	1,9%	32,0%	
	promo	32,9%	42,6%	28,3%	32,3%	31,4%	45,8%	30,1%	43,8%	37,5%	1,2%	38,7%	
B	répondants	29,3%	30,8%	18,7%	27,7%	18,6%	29,2%	45,6%	29,1%	28,1%	31,8%	31,1%	
	promo	24,7%	25,4%	16,2%	24,9%	16,2%	29,1%	43,8%	23,2%	24,3%	39,1%	21,4%	
TB	répondants	16,7%	7,0%	13,3%	20,3%	8,5%	10,8%	19,3%	3,4%	15,8%	62,6%	10,7%	
	promo	10,5%	6,9%	9,0%	13,8%	3,8%	6,4%	14,4%	3,8%	7,2%	58,6%	4,5%	

Annexe B – Méthode statistique par *Random Forest*

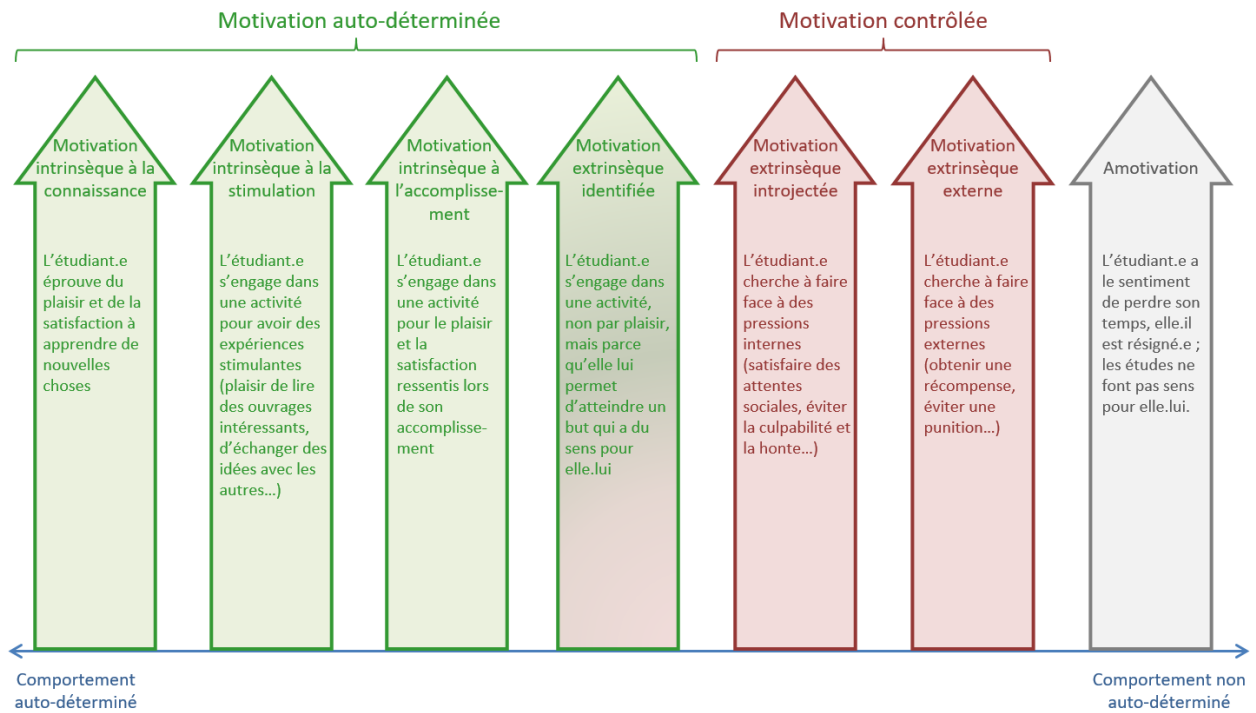
Dans cette étude, nous avons utilisé la méthode statistique par *Random Forest* (RF). Voici des éléments clés à son propos :

- RF est une méthode d'analyse statistique de classification, c'est-à-dire permettant d'établir les relations entre une variable de sortie, qui est binaire, et un ensemble de variables d'entrée (dont les valeurs et les distributions sont quelconques).
- C'est un algorithme de type *machine learning*, c'est-à-dire qui implique une phase d'apprentissage avec une partie des observations (dans notre étude, des étudiants) sélectionnée de façon aléatoire (méthode du *bagging* ; dans notre étude, 2/3 des étudiants) et une phase de prédiction avec le reste des observations (échantillon *out-of-bag* ; dans notre étude, 1/3 des étudiants).
- Il fournit donc un modèle qui est testé (lors de la phase de prédiction) et auquel on peut associer un niveau de performance prédictive (le taux de bonnes prédictions) et un taux d'erreurs (le taux de mauvaises prédictions).
- Il se base sur la méthode des arbres de décision binaires (CART), laquelle consiste à diviser l'ensemble des observations par des questions binaires successives. Chaque question porte sur une variable d'entrée (par exemple, « L'étudiant a-t-il un niveau de motivation autodéterminée au-dessus de 4,1 ou non ? »). L'ordre des questions est déterminé en fonction de leur pouvoir de discrimination relativement à la variable de sortie, la première question étant la plus discriminante. Les questions successives permettent de former des sous-ensembles d'observations qui sont de plus en plus homogènes relativement à la variable de sortie (par exemple, des sous-ensembles d'étudiants qui, pour la plupart, ont validé le S1).
- Un modèle RF combine un grand nombre d'arbres de décision (dans notre étude, 300), ce qui justifie le terme *forest*. Chaque arbre de décision porte sur un échantillon d'observations sélectionnées de façon aléatoire parmi l'ensemble des observations disponibles (les observations faisant partie du *bag*). De plus, pour chaque division dans un arbre donné, un sous-ensemble de variables (dans notre étude, 5 variables) est sélectionné de façon aléatoire parmi l'ensemble des variables d'entrée. Ces deux procédures aléatoires justifient l'adjectif *random*. Pour une nouvelle observation, le modèle établit une prédiction concernant la variable de sortie (ex : la validation au S1) en prenant en compte les prédictions de tous les arbres de décision et en retenant la prédiction la plus fréquente (dite « majoritaire » ; ex : validation au S1 ou non validation au S1).
- Pour tester le modèle RF, les observations de l'ensemble *out-of-bag* sont utilisées.
- Le recours à un grand nombre d'arbres de décision différents, basés sur des ensembles d'observations et de variables d'entrée différents, minimise les biais liés aux arbres individuels et augmente les performances prédictives comparativement à un arbre individuel.
- Un indice, dit « indice de Gini », est associé à chaque variable permettant de mesurer sa contribution à la prédiction. C'est cet indice que nous utilisons dans la présentation de nos résultats en termes de pourcentages de contribution à la prédiction.

Principal avantage des RF pour notre étude :

Les RF permettent d'intégrer dans le modèle des variables hétérogènes dans leur forme, c'est-à-dire des variables numériques, continues ou discrètes, et des variables nominales. Dans notre étude, les RF permettent donc de déterminer et de comparer directement dans un même modèle l'influence sur la réussite d'un ensemble de variables de formes variées.

Annexe C – La motivation et l'amotivation

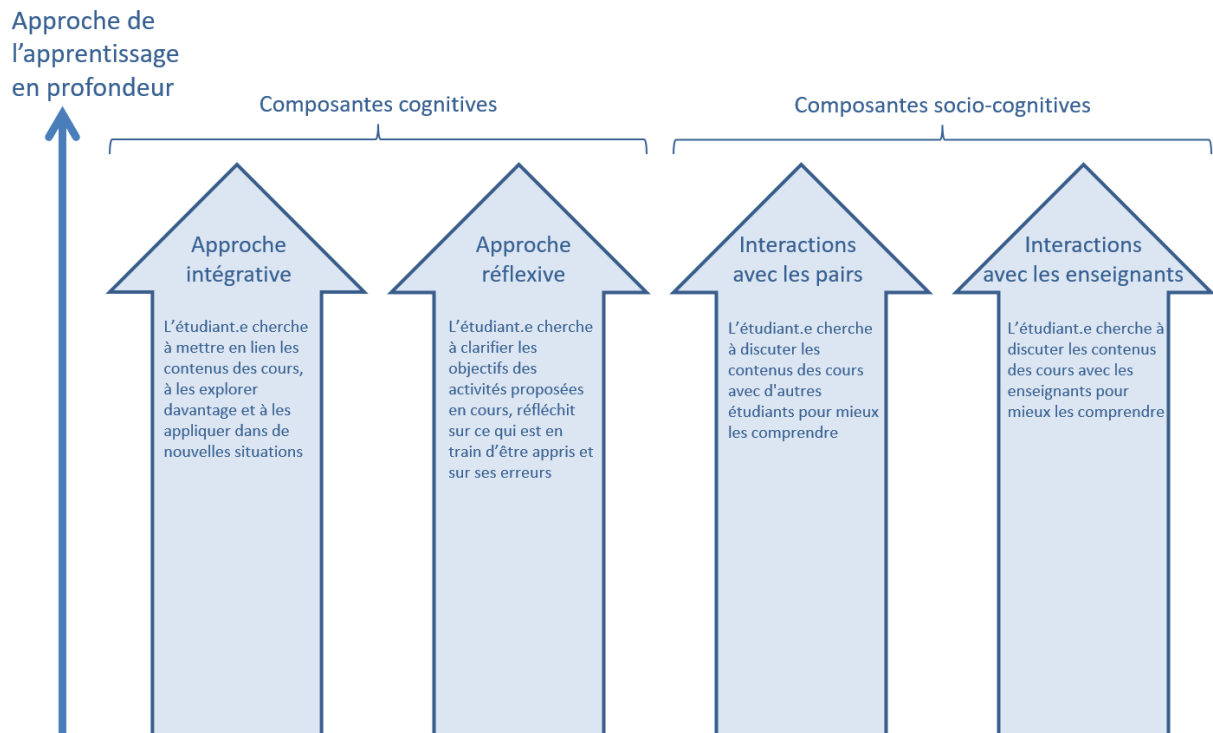


Précisions théoriques : la motivation est « intrinsèque » si l'activité elle-même procure du plaisir et de la satisfaction. La motivation est « extrinsèque » si l'activité est réalisée pour des raisons extérieures à celle-ci. La motivation est « auto-déterminée » si les sources de la motivation sont intériorisées par l'étudiant.e. Les différentes formes de la motivation peuvent être reliées à trois besoins que les individus cherchent à satisfaire, à savoir les besoins d'autonomie, d'affiliation et de compétence.

Précisions méthodologiques : les différentes dimensions sont mesurées au moyen d'un questionnaire (échelle de Likert de 1 à 7, avec 1 = « non, pas du tout d'accord » et 7 = « oui, tout à fait d'accord »). Pour chacune des quatre composantes de la motivation auto-déterminée, le niveau des étudiants est mesuré en prenant la moyenne de leurs réponses à plusieurs questions. La motivation auto-déterminée est calculée à partir de la moyenne de ces quatre mesures. Il en va de même pour la motivation contrôlée et ses deux composantes. Le niveau d'amotivation des étudiants est directement mesuré en prenant la moyenne de leurs réponses à plusieurs questions.

Sources : Ryan & Deci (2000), Vallerand et al. (1989, 1992)

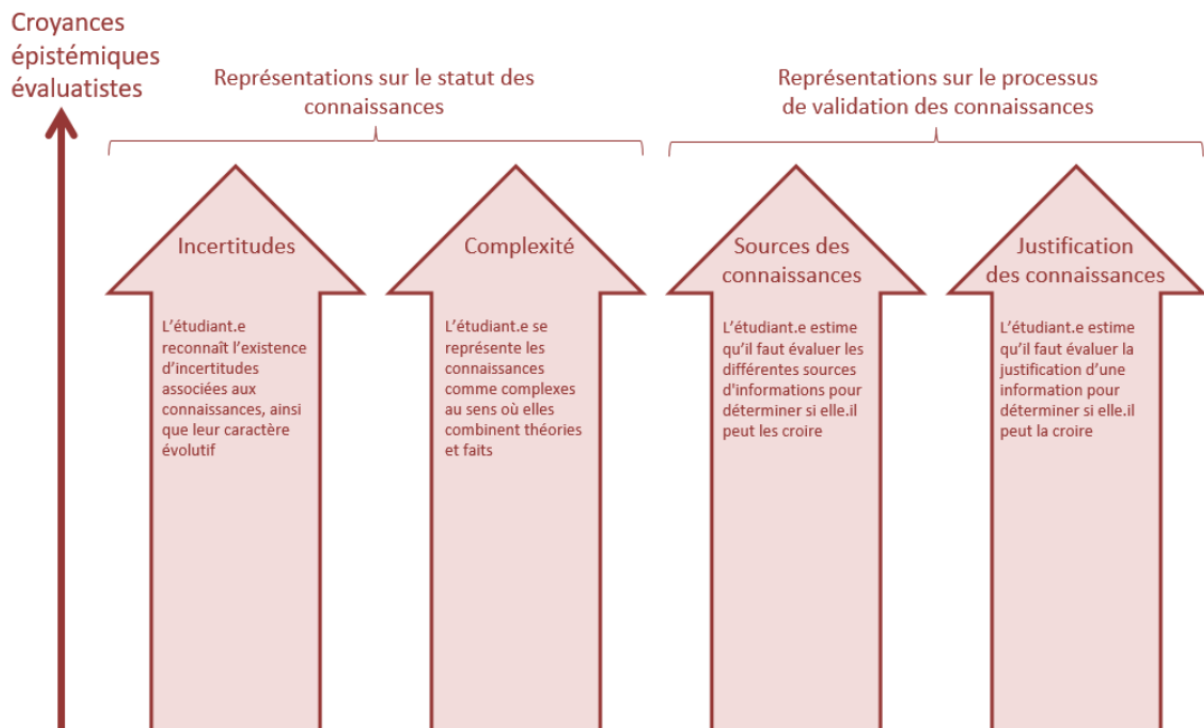
Annexe D – L’approche de l’apprentissage



Précisions méthodologiques : les différentes dimensions sont mesurées au moyen d'un questionnaire (échelle de Likert de 1 à 7, avec 1 = « non, pas du tout d'accord » et 7 = « oui, tout à fait d'accord »). Pour chacune des quatre composantes de l'approche de l'apprentissage en profondeur, le niveau des étudiants est mesuré en prenant la moyenne de leurs réponses à plusieurs questions. L'approche de l'apprentissage en profondeur est calculée à partir de la moyenne de ces quatre mesures. À noter que les deux composantes cognitives sont issues de la littérature, tandis que les deux composantes socio-cognitives ont été ajoutées dans le cadre des études de l'OTP.

Sources : Biggs et al. (2001), Entwistle et al. (2013)

Annexe E – Les croyances épistémiques



Précision méthodologique : le niveau des croyances épistémiques des étudiants est mesuré au moyen d'un questionnaire (échelle de Likert de 1 à 7, avec 1 = « non, pas du tout d'accord » et 7 = « oui, tout à fait d'accord »).

Sources : Hofer & Pintrich (1997) pour la catégorisation des croyances épistémiques ; items originaux créés par l'équipe de l'OTP

Annexe F – Facteurs principaux et secondaires de la réussite et de l'échec

Synthèse des facteurs principaux et secondaires des différentes mesures de réussite et d'échec, suivant les modèles de RF, avec mise en évidence (par un code couleur) des récurrences des facteurs

Mesures de réussite ou d'échec	Facteurs principaux	Facteurs secondaires
Validation du semestre	Composante de la formation suivie	Amotivation
	Mention au Bac	Approche de l'apprentissage intégrative
	Profession des parents	Approche de l'apprentissage interactions étudiantes Soutien des enseignants Motivation intrinsèque à l'accomplissement
Motivation auto-déterminée	Approche de l'apprentissage intégrative	Profession des parents
	Approche de l'apprentissage réflexive	Approche de l'apprentissage interactions avec l'enseignant Approche de l'apprentissage interactions avec les étudiants Composante de la formation suivie Croyances épistémiques évaluateurs - complexité des connaissances
Approche de l'apprentissage en profondeur	Motivation intrinsèque à la connaissance	Motivation extrinsèque identifiée
	Motivation intrinsèque à la stimulation	Croyances épistémiques évaluateurs - sources des connaissances
	Motivation intrinsèque à l'accomplissement	
Croyances épistémiques évaluateurs	Approche de l'apprentissage intégrative	Approche de l'apprentissage réflexive
	Motivation intrinsèque à la connaissance	Composante de la formation suivie
	Profession des parents	Motivation extrinsèque identifiée Motivation intrinsèque à la stimulation
Difficultés	Amotivation	Profession des parents
		Climat stressant ou rassurant
		Composante de la formation suivie
Amotivation	Difficultés liées à l'orientation	Composante de la formation suivie
	Profession des parents	Approche de l'apprentissage réflexive
		Approche de l'apprentissage intégrative
		Difficultés liées à la complexité des contenus