

**PERTINENCE ET VALIDITÉ DES SITUATIONS D'ÉVALUATION DES  
COMPÉTENCES DE L'ENSEIGNEMENT MOYEN GÉNÉRAL DU SÉNÉGAL : CAS  
DES ÉPREUVES STANDARDISÉES DES SVT ADMINISTRÉES, À L'INSPECTION  
D'ACADÉMIE (IA) DE DAKAR, DE 2016 À 2019**

**Pierre Baligue DIOUF<sup>1</sup>**

Université Laval, Canada

[pierrebalique.diouf@ucad.edu.sn](mailto:pierrebalique.diouf@ucad.edu.sn) / [pierre-baligue.diouf.1@ulaval.ca](mailto:pierre-baligue.diouf.1@ulaval.ca)

**Résumé :** Depuis 2014, c'est le décret présidentiel n° 2014-633 du 07 mai 2014 qui fixe et organise les modalités d'évaluation de contrôle continu et compositions dans l'enseignement moyen secondaire sénégalais. Celles-ci sont mises en œuvre à travers les PHARES (progressions harmonisées et évaluations standardisées). Ainsi, depuis la rentrée scolaire d'octobre 2014-2015, les inspections donnent des instructions, aux chefs d'établissements et aux différentes cellules pédagogiques, pour l'élaboration de propositions d'épreuves, parmi lesquelles elles vont choisir et de valider des épreuves dites standardisées. Dans cette recherche, l'objectif consiste à déterminer le niveau de pertinence et de validité des situations d'évaluation de compétences de l'enseignement moyen général sénégalais, en partant du cas des épreuves standardisées des SVT administrées, dans l'IA de Dakar, de 2016 à 2019. Grâce aux directives officielles, en matière d'évaluation des apprentissages, selon l'Approche Par Compétences (APC), mais aussi aux apports théoriques sur les caractéristiques d'une situation d'évaluation selon l'APC, nous avons élaboré un modèle d'analyses, sous forme d'une grille, comportant quatre (04) niveaux (A, B, C, D) : **A- bien pertinente et bien valide (exemplaire) ; B- pertinente et valide ; C- peu pertinente et peu valide ; et D- non pertinente et non valide.** Ainsi, cette étude a révélé que ces situations d'évaluations standardisées, sont presque à **49% pertinentes et valides ; les 10% sont exemplaires, c'est-à-dire bien pertinentes et bien valides.** En revanche, les **41%** sont des situations **peu pertinentes et peu valides (jusqu'à 30% des cas) ou non pertinentes et non valides (pour les 10 à 11%).**

**Mots clés :** Evaluation des apprentissages ; Epreuves standardisées ; Pertinence ; Situation d'évaluation ; Validité.

**RELEVANCE AND VALIDITY OF COMPETENCE ASSESSMENT SITUATIONS  
GENERAL MIDDLE EDUCATION IN SENEGAL : CASE OF STANDARDIZED  
TESTS OF SVT ADMINISTERED AT THE ACADEMY INSPECTION (IA) IN DAKAR,  
FROM 2016 TO 2019**

**Abstract :** Since 2014, it is the presidential decree n° 2014-633 of May 07, 2014, which fixes and organizes the methods of evaluation of continuous control and compositions in Senegalese secondary education. These are implemented through the PHARES (harmonized progressions and standardized assessments). Thus, since the start of the school year in October 2014-2015, the inspectorates give instructions to the heads of establishments and to the various educational units, for the development of test proposals, from which they will choose and validate tests say standardized. In this research, we have set ourselves the objective of determining the level of relevance and validity of skills assessment situations in Senegalese general middle education, starting from the case of the standardized tests of SVTs administered, in the academic inspection of the Dakar, from 2016 to 2019. Thanks to the official directives, in terms of evaluation of learning, according to the Competency-Based Approach (APC), but also to theoretical contributions on the characteristics of an evaluation situation according to the APC, we have developed a model of analyses, in the form of a grid, comprising four (04) levels (A, B, C, D): **A- relevant and valid (exemplary); B- relevant and valid; C- irrelevant and invalid; and D-**

<sup>1</sup> Institut des sciences de l'éducation - Chaire Unesco en sciences de l'éducation (ISE-CUSE)

irrelevant and invalid. Thus, this study revealed that these standardized evaluation situations are almost 49% relevant and valid; the 10% are exemplary, that is to say very relevant and very valid. On the other hand, the 41% are irrelevant and invalid situations (up to 30% of cases) or irrelevant and invalid situations (for 10 to 11%).

**Keywords:** Evaluation of learning; Standardized tests; Relevance; Evaluation situation; Validity.

## Introduction

Dans les programmes de SVT de l'enseignement moyen secondaire général du Sénégal, comme tout programme scolaire, l'évaluation prend différentes formes d'évaluation : l'évaluation diagnostique, l'évaluation formative, l'évaluation sommative et celle certificative. Toutefois, il importe de rappeler que les modèles théoriques de référence et les approches pédagogiques préconisées entraînent des pratiques d'évaluation à visées différentes (De Ketele & Dufays, 2003 cités par De Ketele & Gérard, 2005), et favorisent certains outils au détriment d'autres (Diouf, 2020, 2021). Dans un programme axé sur le développement des compétences, par exemple, Tardif (2006) proposera le dossier d'apprentissages et le portfolio comme outils.

Dans les textes introductifs des programmes sénégalais de SVT (mai 2008 pour l'enseignement moyen, septembre 2015 pour l'enseignement secondaire), on peut lire les indications suivantes, de la commission nationale des programmes. « *Compte tenu des caractéristiques de l'approche par compétences, le professeur de SVT fera preuve d'initiative en utilisant à bon escient les formes d'évaluation les plus courantes, notamment, l'évaluation diagnostique, l'évaluation formative et l'évaluation sommative* » (MEEPEN, 2008, p. 12; MEN, 2015, p. 18). C'est cette évaluation sommative qui fait l'objet de notre recherche. Pour élaborer des épreuves d'évaluation sommative, la commission nationale des programmes recommande aux professeurs le référentiel de « **définition des épreuves de SVT au BFEM et au Bac** » qui se trouve dans les documents officiels desdits programmes.

Au plan pilotage et institutionnel, depuis 2014, c'est le décret n° 2014-633 du 07 mai 2014 portant modalités d'évaluation des apprentissages dans les cycles moyen et secondaire de l'enseignement général, qui fixe et organise les modalités d'évaluation de contrôle continu et compositions au Sénégal. Celles-ci sont mises en œuvre à travers les PHARES (progressions harmonisées et évaluations standardisées). Ainsi, à partir de la rentrée scolaire d'octobre 2014-2015, les inspections donnent des instructions, aux chefs d'établissements et aux différentes cellules pédagogiques, pour l'élaboration de propositions d'épreuves, parmi lesquelles elles vont choisir et valider des épreuves dites standardisées.

Dans cette recherche, nous sommes engagés à déterminer le niveau de pertinence et de validité des situations d'évaluation de compétences de l'enseignement moyen général sénégalais, en partant des épreuves standardisées des SVT administrées, dans les quatre Inspections de l'Éducation et de la Formation (IEF) de l'IA de Dakar, durant les années 2016 à 2019.

Pour y arriver, nous avons dans un premier temps, abordé quelques éléments conceptuels de notre cadre de référence. Par la suite, nous avons élaboré un modèle d'analyse, sous forme d'une grille, comportant quatre (04) niveaux de pertinence et de validité, grâce aux prescriptions officielles, en matière d'évaluation des apprentissages, selon l'Approche Par

Compétences (APC), mais aussi en s'inspirant des apports de la littérature sur les caractéristiques d'une situation d'évaluation selon l'APC. Dans un second temps, nous avons mis en exergue les éléments de notre dispositif méthodologique. Et dans un troisième moment, nous avons présenté les résultats de la recherche, interprétés puis discutés, avant la conclusion.

## 1. Cadre de référence

### 1.1. Quelques éléments conceptuels

#### 1.1.1. Situation d'évaluation et typologie des situations

Pendant longtemps, les concepts **question** et **problème** ont prévalu. Le premier est ou était utilisé pour désigner l'unité de base d'un examen ou d'un test (item dans les écrits anglo-saxons) et le second pour rendre compte de la particularité de certains examens, notamment en mathématique et en géométrie ; certains examens pouvant comporter à la fois des deux (Scallon, 2004). Selon l'auteur, une telle terminologie est incomplète et réductrice, pour ne pas dire caduque, car elle exclut un certain nombre de productions complexes (ce qui est attendu dans le cas de l'APC). A titre d'exemples, « rédiger un conte ou monter une pièce de théâtre » (en littérature), élaborer un projet d'élevage ou préparer une préparation et un montage microscopique (en sciences) ; « confectionner une pièce de vêtement » (en couture) « ne peuvent pas être considérés comme une **réponse** apportée à une question, ni comme une **solution** donnée à un problème : mais dans tous ces cas, il s'agit plutôt de **tâches** (Scallon, 2004, p. 33). Donc, dans le cas d'une évaluation des compétences, l'expression **situation d'évaluation** serait mieux adaptée, afin de désigner à la fois les **questions**, les **problèmes** et les **tâches complexes** (Scallon, 2004).

Une **situation d'évaluation** de l'apprentissage, dite « *assessment to learning* » dans les écrits anglo-saxons, est une situation qui offre l'occasion, à l'évalué, de vérifier son degré de maîtrise des apprentissages par leur intégration réfléchie et leur réinvestissement (Durand et Chouinard, 2012, cités par Sakho, 2017). Cependant, elle n'est pas sans ambiguïté, car elle pourrait désigner une séance d'examen ou la correction de travaux par une équipe de professeurs (Scallon, 2004). Mais comme signalé précédemment, nous utiliserons ce terme, pour désigner à la fois les **questions**, les **problèmes** et les **tâches complexes**, proposés dans les épreuves des SVT, pour **évaluer les apprentissages** des élèves.

Une **typologie des situations d'évaluation**, qui est leur hiérarchisation en différents types suivant leurs degrés de complexité, leurs niveaux d'ouverture et taxonomiques, nous permet « de distinguer les situations de compétences des autres types de situations » (Scallon, 2004, cité par Ducharme, Durand & Sakho, 2014, p. 173). Selon ces auteurs, nous avons trois types : les **situations de connaissances**, les **situations d'habiletés** et les **situations de compétences** (Durand, Sakho, & Ducharme, 2014). Toutefois, en plus de ces trois types de situations, Scallon (2004) évoque la **situation de stratégie** qui précède celle de compétences. Roegiers (2003) évoque trois types de **problèmes** (**fermés**, **semi-ouverts** et **ouverts**). Quant à Rey et al. (2006), cités par Ducharme et al. (2014), ils préfèrent les termes de **compétences de premier, deuxième et troisième degré**.

Chaque situation d'évaluation permet de voir quel est le niveau taxonomique de Bloom (1956) qui est visé. Ainsi, Tarillon (2006) nous propose de résumer les situations d'évaluations de Scallon (2004) comme suit.

❖ **Situation d'évaluation de connaissances [problèmes fermés]**

C'est une question faisant clairement référence à une information supposée disponible, que l'apprenant a déjà mémorisée et qu'il devra restituer ou réutiliser (Tarillon, 2006). Dans une telle situation, il y a une seule possibilité de solution, et elle se limite à évaluer le premier niveau taxonomique de Bloom, la restitution de mémorisation, ce sont des **questions** de **type I**.

❖ **Situation d'évaluation d'habileté [problèmes semi-ouverts]**

C'est une question où l'élève devra lui-même trouver la connaissance nécessaire à la construction de sa réponse (Tarillon, 2006). Toutes les informations ne sont pas fournies dans la situation, une partie est révélée mais pas toute (d'où le terme semi-ouvert). Il y a également une seule possibilité de solution. Elle vise les niveaux de compréhension et d'application de la taxonomie de Bloom, par des **questions** de **type II** pouvant exiger de l'apprenant des réponses élaborées ou construites.

❖ **Situation d'évaluation de stratégies [problèmes ouverts]**

La **situation d'évaluation de stratégies** est un problème ouvert car la question suppose que l'élève choisisse un ensemble de procédures parmi plusieurs possibles dans le but de réaliser une tâche (Scallon, 2004; Tarillon, 2006). Toutes les informations ne sont pas fournies et il y a plusieurs possibilités de solutions. Selon Scallon (2004), cité par Tarillon (2006), la stratégie a plusieurs propriétés : elle est consciente (le sujet peut la verbaliser, l'expliciter) ; elle se distingue de l'habileté par le fait qu'elle n'est pas automatisable, elle ne s'applique pas de façon identique à toute situation. Pour eux, c'est un « savoir comment » couplé à un « savoir quand ». Bien qu'elle comporte des **questions** de **type III**, permettant de viser des niveaux taxonomiques supérieurs (analyse, synthèse et évaluation), par le fait qu'elle soit ouverte, elle ne permet pas encore d'inférer la compétence, sinon elle devrait être **complexe**.

❖ **Situation d'évaluation d'une compétence [problèmes ouverts et complexes]**

En plus d'être **ouverte**, la **situation d'évaluation d'une compétence** est un problème **complexe**, sans être forcément compliqué (Gérard, 2008).

Dans l'optique d'évaluation, nous avons retenu la conception selon laquelle la compétence serait un ensemble de **connaissances stabilisées**, **d'habiletés** et **d'attitudes** nécessaires pour assurer des tâches et des rôles dans une situation. Ainsi, pour construire une situation d'évaluation d'une compétence, il faut être sûr que les tâches permettront bien d'inférer la compétence. Il faut aussi les répéter pour une même famille, au moins trois fois, et étayer de moins en moins ou accroître le niveau de difficulté (Scallon, 2004; Tarillon, 2006). A ce niveau, on a des **questions** de **type III et IV**<sup>2</sup>, questions à développement long, à réponse construite et élaborée, exigeant un niveau de réflexion plus élevé de la part de l'apprenant (Louis, 1999, cité par Durand & Chouinard, 2012).

### 1.1.2. Pertinence

La pertinence d'une épreuve ou d'une **situation d'évaluation** c'est son degré de « compatibilité » avec les autres éléments du système auquel elle appartient (Raynal & Rieunier, 1997). Autrement dit, c'est le caractère plus ou moins approprié de la **situation**,

---

<sup>2</sup> Nous avons utilisé cette appellation pour désigner les questions exigeant des productions complexes de la part de l'élève, à travers des situations ouvertes et complexes

selon qu'elle s'inscrit dans la ligne des objectifs visés » (De Ketele et *al.*, 1989 ; De Ketele & Gérard, 2005).

Beaucoup d'auteurs, ayant écrit sur l'évaluation des compétences, (Roegiers, 2000, 2001 ; Scallon, 2004, 2005, 2007 ; Durand, Sakho, & Ducharme, 2014) remplacent le terme de pertinence par celui d'« **authenticité** ». Selon eux, une situation d'évaluation doit être contextualisée, authentique, nouvelle, et signifiante pour l'élève. Durand et Ducharme (2014) ont proposé une grille d'analyse de la pertinence d'une situation de compétence qu'ils échelonnent, à l'aide de plusieurs critères (complexité, contextualité, conformité, efficacité de l'instrumentation et précision des documents), en quatre niveaux : **niveau A** (**exemplaire**, situation d'évaluation dont la pertinence est plus parfaite) ; **niveau B** (situation pertinente) ; **niveau C** (situation à retravailler, peu pertinente) et **niveau D** (situation à refaire, non pertinente).

Ce concept de pertinence est souvent accompagné de concepts de validité et de fidélité dont il se distingue.

### 1.1.3. Validité

En mesure, De Ketele, Gérard et Roegiers (1997) diront que des « informations **valides** » signifient que celles effectivement récoltées correspondent aux informations recherchées et implique donc que l'évaluateur choisisse la bonne stratégie (méthodes, techniques, outils) à mettre en œuvre pour recueillir les informations qu'il souhaite obtenir.

Par exemple, si un enseignant veut vérifier si ses élèves sont capables de mesurer un courant à l'aide d'un multimètre, il ne doit pas demander aux élèves de décrire par écrit les opérations à effectuer, mais bien les mettre en situation pratique de mesurer un courant à l'aide d'un multimètre (De Ketele, Gérard, & Roegiers, 1997, p. 5).

Donc, on ne peut pas déclarer vouloir évaluer une compétence et proposer une situation qui se limite à évaluer des connaissances et habiletés, et inversement. « *La validité est le degré d'adéquation entre ce que l'on déclare faire (évaluer telle ou telle dimension) et ce que l'on fait réellement, entre ce que l'outil mesure et ce qu'il prétend mesurer* » (Laveault & Grégoire, 1997, 2002, cités par De Ketele & Gérard, 2005, p. 2). Pour évaluer une compétence, « (...) il convient de s'assurer que la situation dans laquelle un élève est placé pour démontrer telle ou telle compétence « soit complexe, structurée et l'incite vraiment à cette mobilisation [sinon une telle situation n'est pas valide] » (Scallon, 2004, p. 262).

*NB : pour les deux concepts de pertinence et de validité, nous emploierons le terme générique **alignement**.*

Enfin, arrivons à la question problème générale (QPG) de recherche suivante.

**Quel est le niveau de pertinence et de validité des situations d'évaluation de compétences au sein des épreuves standardisées des SVT administrées, à l'Inspection d'Académie (IA) de Dakar, de 2016 à 2019 ?**

Pour répondre à une telle QPG, il nous faut un modèle théorique d'analyse.

## 1.2. Modèle d'analyse

Compte tenu de cette recherche qualitative, et de notre QPG, nous avons opté d'utiliser un modèle d'analyse, sous forme d'une grille ou tableau, conçu à partir de l'inspiration et de « *la triangulation théorique* » (Savoie-Zajc, 2011, cité par Ducharme, 2017), de quatre références théoriques. Il s'agit : (i) du **référentiel de la définition des épreuves** des SVT dans le programme officiel ; (ii) du **modèle d'analyse** de la **pertinence** et **validité des situations d'évaluation** selon Durand et Ducharme (2014) ; (iii) des travaux sur

**l'élaboration et la validation des épreuves d'évaluation**, selon l'APC, de De Ketele et Gérard (2005) ; et (iv) du **modèle d'évaluation** des compétences basé sur **la complexité des situations d'évaluation** (Gérard, 2005). Les deux autres dernières références, qui sont de fructueux travaux de recherche, nous permettent de répondre à la question suivante : qu'est-ce qui atteste la validité d'une situation d'évaluation ou d'une épreuve ?

Selon De Ketele et Gérard (2005), et Gérard (2005), la situation d'évaluation de compétences ou **situation-problème** devra, en plus d'être ouverte, être complexe.

Tableau : modèle d'analyse de la pertinence et la validité des situations d'évaluation de compétences des épreuves standardisées de SVT de l'enseignement moyen général sénégalais. (Inspiré et adapté de Durand & Ducharme, 2014)

Niveaux de pertinence et de validité Critères	A- bien pertinent et bien valide	B-pertinent et valide	C- peu pertinent et peu valide	D- non pertinent et non valide
<i>Complexité</i>	La situation propose un défi cognitif sous la forme d'un problème ouvert et complexe exigeant la mobilisation de nombreuses et diverses ressources (interne et externes) : savoirs, habiletés et des savoir-faire déjà acquis.	La situation propose un problème semi-ouvert sous la forme d'une ou de plusieurs tâches qui exigent la mobilisation de ressources (internes ou externes) : des savoirs et des savoir-faire	La situation propose un ou des problèmes fermés qui exigent l'application de savoirs et des savoir-faire déjà acquis.	La situation propose une série de tâches qui exige l'application de savoirs déjà acquis.
<i>Contextualité ou signifiante</i>	La situation propose un contexte authentique qui suscitera l'intérêt de la clientèle visée.	La situation propose un contexte réaliste, susceptible d'attirer l'intérêt de la clientèle visée.	La situation propose un contexte réaliste, peu susceptible de motiver la clientèle visée ou porteur d'un biais culturel.	La situation est décontextualisée.
<i>Conformité</i>	La situation d'évaluation comporte des questions de compréhension, d'interprétation, d'analyse, d'appréciation, d'évaluation, et de création.	La situation d'évaluation comporte des questions de compréhension, d'interprétation, et d'appréciation	La situation d'évaluation comporte des questions de compréhension, et d'interprétation seulement	La situation d'évaluation ne comporte que des questions de compréhension.
	La situation comporte des tâches complexes, des questions de synthèse ou d'intégration d'une ou des compétences à évaluer	La situation comporte des tâches complexes, des questions de synthèse ou d'intégration en cohésion avec une ou des compétences à évaluer.	La situation comporte des questions d'interprétation en relation avec les éléments d'une ou des compétences à évaluer	La situation d'évaluation est en lien avec des éléments de la ou les compétences à évaluer et ne prévoit dans réalisation de la tâche aucune démarche particulière.

<i>Efficacité de l'instrumentation</i>	La situation prévoit plusieurs outils d'évaluation descriptifs, par approche analytique ou globale, pour la ou les compétences ciblées.	La situation prévoit des outils d'évaluation selon une échelle uniforme pour la ou les compétences ciblées.	La situation prévoit un corrigé.	La situation ne prévoit pas d'outils d'évaluation.
	Les critères d'évaluation ciblés sont reformulés en éléments observables et comprennent une dimension et un objet à évaluer.	Les critères d'évaluation ciblés sont reformulés en éléments observables et comprennent parfois deux dimensions.	Les critères d'évaluation sont trop ou pas assez nombreux et leur reformulation omet parfois une dimension.	Les critères d'évaluation ne sont pas reformulés et omettent la dimension ou l'objet à évaluer.
	La situation prévoit un barème de notation ainsi que des indications pour l'interprétation des résultats sous la forme de copies-types.	La situation prévoit un barème de notation ainsi que quelques indications pour l'interprétation des résultats.	La situation prévoit un barème de notation	La situation prévoit un barème de notation confus ou incomplet.
<i>Présence de documents supports</i>	Chaque exercice présente un ou des documents supports différents	Chaque exercice présente un document support	Un seul exercice présente un document support	Aucun exercice ne présente de document support
<i>Qualité des documents supports</i>	La présentation est bien aérée, cohérente et facile à utiliser.	La présentation est cohérente et son utilisation pourrait être plus simple.	La mise en page ou la police sont inadéquats ce qui rend son utilisation difficile.	L'organisation de l'information est déficiente.
	Les consignes données à l'élève sont claires, complètes, concises et appropriées pour le niveau de l'élève.	Les consignes données à l'élève sont complètes toutefois certaines pourraient être mieux formulées afin d'être plus appropriées pour le niveau de l'élève.	Les consignes données à l'élève sont trop descriptives ce qui nuit à leur compréhension.	Les consignes données à l'élève sont équivoques.

Source : (Diouf, 2021)

La première ligne du tableau présente les quatre niveaux de pertinence et de validité retenus (Diouf, 2021). Le niveau (A), que nous désignons par « **bien pertinent et valide** », est celui recommandé, correspond au niveau **exemplaire** de Ducharme et Durand (2014). Le niveau (B) que nous désignons par « **pertinent et valide** » correspond au niveau « **pertinent** » de Ducharme et Durand (2014) et au niveau « **valide** » de De Ketele et Gérard (2005) ; et de Gérard (2005). Autrement dit, une situation **pertinente et valide** c'est celle qui est **pertinente et valide**, au regard à la fois des exigences officielles du programme de SVT et des apports théoriques des références précitées. Le niveau (C) « **peu pertinent et peu valide** » correspond au niveau « **à retravailler** » de Ducharme et Durand (2014), et le dernier

niveau (D) « **non pertinent et valide** » correspond à celui désigné par « **à refaire** » chez Ducharme et Durand (2014), c'est-à-dire **ni pertinent ni valide**, au sens de De Ketele et Gérard (2005). Donc, une épreuve **non pertinent et non valide** ne satisfait aucun de nos critères indiqués dans le tableau.

## 2. Dispositif méthodologique

Cette présente section contient les éléments de cadrage méthodologique de recherche : matériel et population cible ; échantillonnage ; méthodes et outils de recueil d'informations ; et ceux du traitement des données obtenues.

### 2.1. Matériel et Population cible

Les questions de recherche et les objectifs que nous nous sommes assignés, dans cette étude, nous ont amenés à cibler des épreuves de contrôle continu et de compositions de SVT et le document de programme de SVT (mai 2008) de l'enseignement moyen général du Sénégal, qui nous servira de référence pour les directives ou prescriptions en matière d'évaluation.

### 2.2. Echantillonnage

Pour l'essentiel, nous avons fait recours à un échantillonnage typique non probabiliste, par choix raisonné (Quivy et Van Campenhoudt, 1995) ; c'est un échantillonnage ciblé ou dirigé (Mboup, 2003). Ensuite, l'échantillonnage est stratifié, suivant les quatre niveaux d'enseignement et les types d'épreuves, devoirs et compositions.

Pour les épreuves, nous avons, dans un premier temps, réparti les épreuves collectées par niveau : 6<sup>ème</sup>, 5<sup>ème</sup>, 4<sup>ème</sup> et 3<sup>ème</sup> ; et pour plus de fiabilité et de représentativité des échantillons, nous avons sélectionné quatre (04) épreuves par niveau d'études, soit seize (16) épreuves : deux épreuves proposées comme devoirs (contrôles continus) et deux autres comme compositions (ou BFEM blanc pour le 2<sup>nd</sup> semestre en 3<sup>ème</sup>). Ces épreuves sont toutes issues des évaluations standardisées, de 2016 à 2019, dans les quatre IEF de l'IA de Dakar.

### 2.3. Recueil d'informations et traitement des données

Pour rester conforme à notre démarche qualitative, nous avons retenu deux méthodes et outils de recueil d'informations tels que : **l'étude de documents** et **l'analyse de contenu** des documents, épreuves d'évaluation sommative de SVT. L'exploitation du document de programme nous a permis d'identifier toutes les prescriptions relatives à l'évaluation des apprentissages et la conception des épreuves d'évaluation sommative. Pour le traitement des données, nous avons fait recours à **Microsoft Excel**. Ce dernier nous a facilité l'établissement de tableaux de données et l'interprétation graphique de ces données.

## 3. Résultats

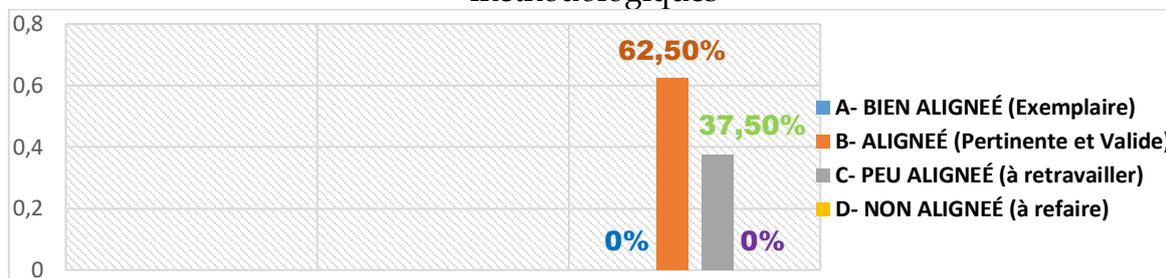
### 3.1. Présentation et interprétation des résultats

#### 3.1.1. Complexité des situations d'évaluation de compétences

D'après l'analyse des résultats, le niveau **B- pertinent et valide** représente le mode, avec près de **63%** des épreuves qui ont des **situations** qui proposent un problème semi-ouvert sous la forme d'une ou de plusieurs tâche(s) qui exige(nt) la mobilisation de

ressources (internes ou externes) : des savoirs et des savoir-faire. Le graphique 1, ci-après, permet d'en faire l'interprétation graphique.

Graphique 1 1: niveaux de complexité des situations d'évaluation de compétences méthodologiques

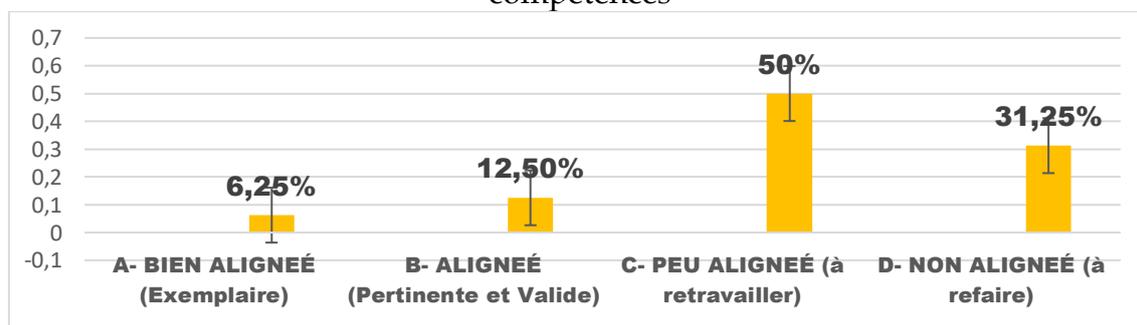


Aucune des épreuves (pour tout le cycle) n'a de situation d'évaluation de compétences qui propose un défi cognitif sous la forme d'un problème ouvert et complexe exigeant la mobilisation de nombreuses et diverses ressources (internes et externes). Par contre, les 37,50% des épreuves sont composés de situations proposant un ou des problème(s) fermé(s) qui exige(nt) l'application de savoirs et des savoir-faire déjà acquis. Toutefois, aucune des épreuves n'a de situation d'évaluation de compétences qui propose, uniquement, une série de tâches qui n'exige que l'application de savoirs déjà acquis.

### 3.1.2. Contextualité ou signifiante des situations d'évaluation de compétences

Les résultats révèlent qu'il y a 50% des situations proposant un contexte réaliste, porteur d'un biais culturel, peu susceptible de motiver les élèves évalués. Les 31,25% sont des situations décontextualisées, porteuses d'aucun biais culturel ou contexte authentique pouvant susciter l'intérêt des apprenants. Il n'y a que 6,25% qui contiennent des situations-problèmes proposant un contexte authentique qui pourrait susciter l'intérêt des enfants ; les 12,50% autres renferment des situations qui proposent un contexte réaliste, susceptible d'attirer l'intérêt des apprenants.

Graphique 2: niveaux de contextualité ou signifiante des situations d'évaluation de compétences

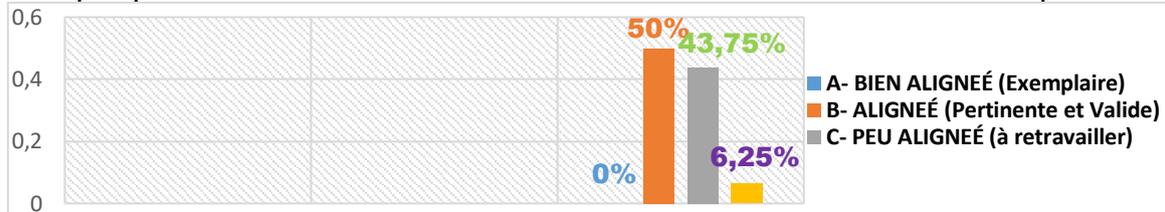


### 3.1.3. Conformité des situations d'évaluation de compétences

D'après notre modèle, les 50%, des situations d'évaluation comportent des questions de compréhension, d'interprétation, et d'appréciation avec des tâches complexes, des questions de synthèse ou d'intégration en cohésion avec une ou des compétence(s) méthodologique(s) à évaluer. Sinon les 43,75% des situations d'évaluation comportent des questions de compréhension, et d'interprétation seulement, en relation avec les éléments

d'une ou des compétence(s) à évaluer. Pour le reste, les **6,25%** ne comportent que des questions de compréhension, en lien avec des éléments de la ou des compétence(s) à évaluer et ne prévoient, dans la réalisation de la tâche, aucune démarche particulière. Le graphique 3 nous présente les résultats par rapport aux niveaux d'alignement de la conformité des situations d'évaluation de compétences.

Graphique 3: niveaux de conformité des situations d'évaluation de compétences

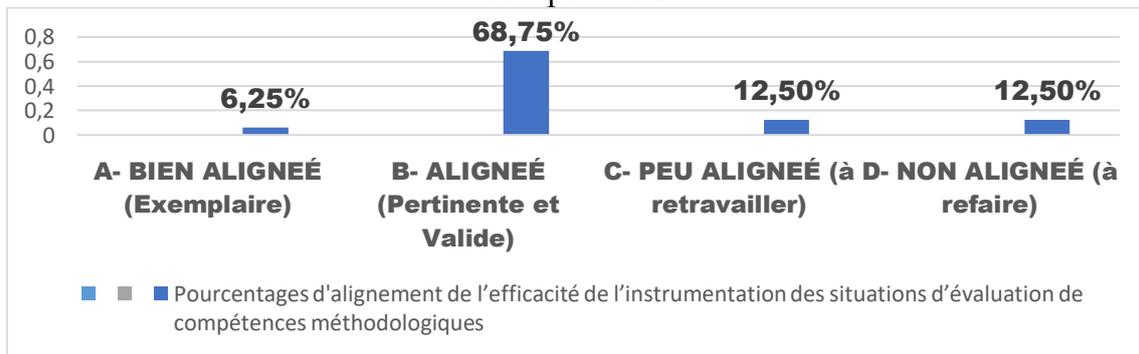


Enfin, pour tout le cycle, aucune épreuve n'a des situations d'évaluation comportant, à la fois, des questions de compréhension, d'interprétation, d'analyse, d'appréciation ou d'évaluation, et de création ; il y a absence totale de situation qui comporte des tâches complexes, des questions de synthèse ou d'intégration d'une ou des compétence(s) méthodologique(s) à évaluer.

#### 3.1.4. Efficacité de l'instrumentation des situations d'évaluation de compétences

L'analyse des données démontre qu'il y a près de **69%** des situations prévoyant des outils d'évaluation selon une échelle uniforme pour la ou les compétence(s) méthodologique(s) ciblée(s). Autrement dit, en SVT, ces **69%** des épreuves ont des situations qui prévoient un barème de notation mais pas d'autres indications pour l'interprétation des résultats. Il y a, d'une part, **12,5%** des situations qui prévoient juste un barème de notation, sans précisions sur les critères de réussite et donc d'évaluation ; et d'autre part, **12,5%** autres situations ont un barème de notation incomplet (ou non indiqué) ou même confus (voir graphique 4).

Graphique 4: niveaux d'efficacité de l'instrumentation des situations d'évaluation de compétences

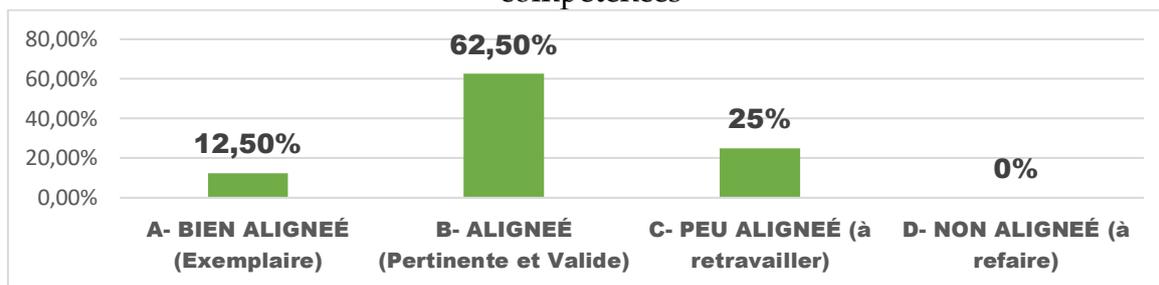


Avant d'arriver à la discussion des résultats, nous pouvons, déjà, dire que cette situation peut trouver son explication dans les directives du programme. Car, le référentiel de la définition des épreuves, ne donne pas d'indications, à ce sujet, pouvant aider les concepteurs des épreuves à proposer des situations dont l'instrumentation tient compte de cela.

### 3.1.5. Présence de documents supports des situations d'évaluation

En termes d'interprétation, les résultats montrent que ce sont les **62,50%** des épreuves qui ont des situations dont chacune présente un document support. Les **25%** des épreuves ont une seule situation qui présente un document support ; il n'y a que **12,50%** des épreuves qui ont des exercices dont chacun présente un ou des document(s) support(s) différent(s). Toutefois, aucune épreuve ne renferme des situations où il y a absence totale de document support (graphique 5).

Graphique 5: niveaux de présence de documents supports des situations d'évaluation de compétences

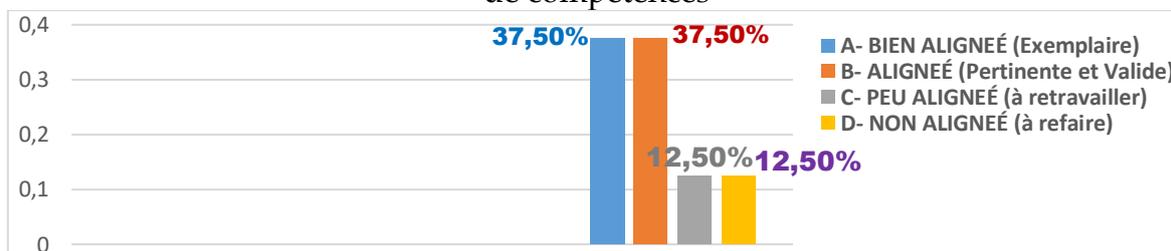


Mais qu'en sera-t-il de la qualité des documents supports.

### 3.1.6. Qualité des documents supports des situations d'évaluation

D'après notre modèle d'analyse, il y a **75%** (deux fois **37,5%**) des situations qui ont des documents supports dont la présentation est bien aérée, cohérente et facile à utiliser, mais dont leur utilisation pourrait être plus simple. Les consignes données à l'élève sont claires, complètes, concises et appropriées pour le niveau de l'élève ou, dans une moindre mesure, elles sont complètes mais certaines pourraient être mieux reformulées afin d'être plus appropriées par les élèves. Pour les **25%** restants, la mise en page et/ou la police sont inadéquates, ce qui rend difficile l'utilisation, ou encore l'organisation de l'information qui est déficiente. Pour les consignes données à l'élève, elles sont parfois soit trop descriptives soit équivoques, ce qui nuit à leur compréhension. Le graphique 6 présente ces résultats.

Graphique 6: niveaux de qualité des documents supports dans les situations d'évaluation de compétences



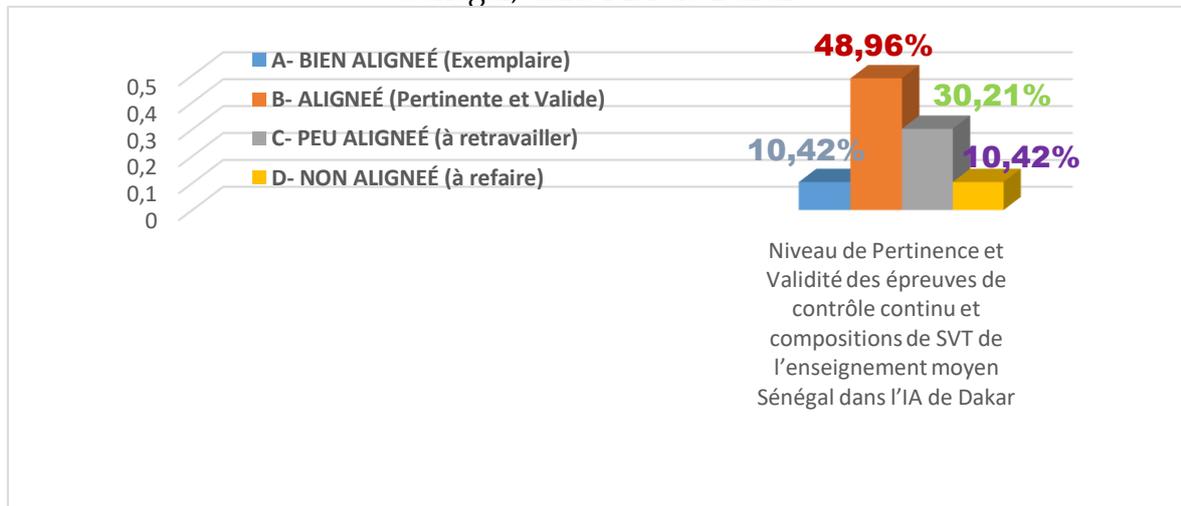
A partir de ce stade, il faut envisager la synthèse de ces résultats issus de l'analyse des critères de pertinence et de validité, afin d'apporter des réponses à la QPG.

### 3.1.7. Synthèse des résultats obtenus par rapport au niveau de pertinence et de validité des situations d'évaluation standardisées

En vue d'apporter des éléments de réponse à la QPG, nous avons, ci-dessous, la présentation graphique portant sur la distribution des six critères de pertinence et de

**validité** des situations de compétences des épreuves standardisées de SVT de l'enseignement moyen sénégalais, dans l'IA de Dakar.

Graphique 7: niveaux de pertinence et validité des situations d'évaluation des compétences des épreuves standardisées de SVT de l'enseignement moyen général du Sénégal, dans l'IA de Dakar



L'analyse des résultats cumulés révèle qu'il y a qu'environ **49%** des épreuves qui sont **pertinentes** et **valides**, et qu'il n'y a que **10%** qui sont **exemplaires**, autrement dit bien pertinentes et bien valides. Pour le reste, c'est-à-dire les **41%**, ce sont des épreuves à retravailler (jusqu'à **30%** des cas) ou à refaire carrément (pour les **10 à 11%** des situations).

### 3.2. Discussion des résultats

La synthèse et l'interprétation des résultats, pour toutes les épreuves et tous critères confondus, nous ont permis de déterminer le niveau d'alignement (pertinence et validité) des situations d'évaluation de compétences des épreuves standardisées, conçues par les professeurs de SVT de l'IA de Dakar, de 2016 à 2019. Selon notre modèle d'analyse, il y a qu'environ **49%** des épreuves qui sont **pertinentes** et **valides**, et **10%** qui sont exemplaires, bien **pertinentes** et **bien valides**. Le reste, c'est-à-dire **41%**, constitue des situations **peu ou non pertinentes** et **peu ou non valides**, ne respectant pas les exigences d'une évaluation d'un programme selon le développement des compétences. Ces épreuves sont à retravailler et/ou à refaire totalement (Durand, Sakho, & Ducharme, 2014). Car, on ne peut pas évaluer et/ou inférer des compétences avec des épreuves peu ou non pertinentes et valides (De Ketele & Roegiers, 1993; De Ketele & Gérard, 2004; Scallon, 2004).

Toutefois, les critères de pertinence et validité à retravailler sont essentiellement la **contextualité** ou **signifiante** (pour laquelle, **une seule situation** était **bien pertinente et valide** et **deux pertinentes et valides**), la **qualité des documents support**, la **conformité** des situations et leur **complexité**. Et dans une moindre, **l'efficacité de l'instrumentation** et la **présence de documents supports** dans les situations d'évaluations des compétences méthodologiques.

Nous savons que l'une des caractéristiques de l'Approche Par Compétences, c'est le fait de donner du **sens aux apprentissages** (Roegiers, 2000, 2001; Scallon, 2004). En plus, pour le développement et l'évaluation des compétences, il faut des situations **complexes** donnant l'opportunité, aux apprenants, de construire des connaissances utiles et des compétences

pour une meilleure prise en charge de leur environnement socioéconomique et culturel (Machado, 2008). Donc, les situations proposées, dans les épreuves de SVT, devraient être **signifiantes**, pour les élèves, en les offrant « *un accès à une multitude de situations, celles de leur vie quotidienne, familiale, communautaire, scolaire ou autres, comme à celles de leur vie personnelle et professionnelle* » (Machado, 2008, p. 13).

Malheureusement, les situations proposées dans les épreuves de contrôle continu et compositions de SVT, par les professeurs de SVT de l'IA de Dakar, sont **décontextualisées**. Pourtant, la **nature des contenus disciplinaires prescrits**, qui portent essentiellement sur la **vie** et la **terre**, presque sur le vécu quotidien des élèves, offrent une **large gamme d'éléments** permettant de construire, non seulement, des **situation-problèmes contextualisées, signifiantes**, mais aussi **complexes** (par la **diversité des thèmes et leçons étudiés**) ; et **conformes** (permettant d'avoir un panache de **questions de tous les types : I, II, III et IV**, et permettant de viser ainsi des **niveaux taxonomiques élevés**).

La **complexité** est un caractère essentiel des **situation-problèmes** devant permettre de **développer les compétences** mais aussi de les **évaluer** (De Ketele & Gérard, 2004, 2005 ; Gerard, 2008). Alors qu'en se basant sur la littérature (Roegiers, 2000, 2001; De Ketele, 2008) et les indications épistémologiques des **concepteurs des programmes de SVT** (MEEPEN, 2008; MEN, 2015), en ce qui concerne les définitions choisies de la notion de compétence, on perçoit aisément et clairement qu'il faut des **situations-problèmes** pour l'apprentissage et l'évaluation des SVT, selon l'APC. Gérard (2005, 2006, 2008) parle de **situation complexe** à distinguer d'une situation compliquée. La complexité appelle dès lors l'adaptabilité, car on ne sait jamais à quoi s'attend exactement face à une situation complexe » (Gerard, 2008, p. 169). Le problème ne pourrait s'expliquer que dans le fait que les professeurs ne soient pas bien outillés, en termes de formation en APC, pour concevoir de telles situations.

Concernant la **conformité**, nous avons constaté que l'on continue toujours à évaluer, plus, des **connaissances** mais surtout **des habiletés** que de **compétences** (Diouf, 2020, 2021 ; Diouf & Kane, 2021). Donc, ces résultats viennent confirmer ceux que nous avons obtenus en analysant des épreuves d'évaluation, de 2011 à 2014, administrées dans les Inspections d'Académie de Saint-Louis et de Dakar (Diouf, 2020).

Ainsi, ces résultats ramènent et posent, avec acuité, le problème de la **validité des épreuves** de SVT, administrées dans l'enseignement général du Sénégal. En effet, on ne peut pas déclarer avoir évalué une ou des compétence(s) et proposer une ou des situation(s) qui ne permet(tent) qu'à évaluer les habiletés, (De Ketele, Gérard, & Roegiers, 1997) ; car, ces épreuves sont composées de situations dominées par des questions de **type II** mais « carantes » en question de **types III et IV** (Diouf, 2020, 2021 ; Diouf & Kane, 2021). En considérant ces résultats, l'on est en droit de dire, qu'en SVT, on est tombé dans le piège de l'excès d'évaluation des habiletés, en voulant évaluer, certes, des compétences. Alors que les habiletés et les savoir-faire ne sauraient substituer les compétences, puisque d'après la clarification qu'on en avait faite, les habiletés sont des éléments constitutifs (ressources) de la compétence. (Jonnaert, 2002; Diouf, 2020).

## Conclusion

Cette recherche a montré qu'au niveau de la **pertinence** et de la **validité** des situations d'évaluation des compétences au sein des épreuves standardisées, administrées dans l'enseignement moyen général sénégalais, il y a encore beaucoup d'efforts à fournir. En effet, il y a que presque **49%** de nos épreuves qui sont **pertinentes** et **valides**, et parmi elles, il n'y a que **10%** qui sont **exemplaires**, autrement dit **bien pertinentes** et **bien valides**. Les **41%** sont des épreuves à retravailler (jusqu'à **30%** des cas) ou à refaire carrément (pour les **10** à **11%**). Cependant, les critères de **pertinence** et de **validité** à retravailler n'ont pas tous les mêmes niveaux d'alignement. Certains sont moins pertinent et valides que d'autres. De façon remarquable, la **contextualité** ou **signifiante** fut un critère pour lequel, seulement **6,25%** des épreuves étaient **bien pertinent et valides** et **12,50%** **pertinent et valides** ; et la **qualité des documents supports**, la **conformité** des situations et leur **complexité** sont aussi à retravailler. Cependant, **l'efficacité de l'instrumentation** et la **présence de documents supports**, dans les situations d'évaluations des compétences, sont à des niveaux d'alignement acceptables, dépassant les **50%** de **bien pertinent et valides** et **pertinent et valides**.

Après une dizaine d'années de mise en œuvre du programme conçu et écrit selon l'APC (MEEPEN, 2008), les épreuves d'évaluation des apprentissages de SVT sont toujours dominées par les situations **d'évaluation d'habiletés** (correspondant aux questions de **type II**) ; moins de **situations de connaissances** (correspondant aux questions de **type I**) ; et peu de **situations de stratégies** et de **compétences** (correspondant respectivement aux questions de types **III et IV**) (Diouf, 2020, 2021 ; Diouf & Kane, 2021). En voulant évaluer des compétences, on tombe dans l'évaluation excessive d'habiletés.

Mais comment expliquer cet état de fait, que les épreuves sont **moins pertinentes et valides** du point de vue **fond (ou contenu)** des situations d'évaluation des compétences ?

Nous pourrions, dans d'autres recherches, trouver une partie des explications dans le programme officiel, notamment dans ses éléments de modèle théorique de référence sous-tendant l'APC, et dans le référentiel de la définition des épreuves ; l'analyse des caractéristiques socioprofessionnelles (**ancienneté dans l'enseignement, diplômes, date d'obtention du diplôme professionnel, formation à l'évaluation et références des professeurs pour la conception des épreuves**) ; et des aspects de pilotage et institutionnels, concernant le **non-respect** de la **démarche évaluative** et des **valeurs de l'évaluation**.

## Références bibliographiques

- De Ketele, J.-M. (2008). L'approche par compétences: au-delà du débat d'idées, un besoin et une nécessité d'agir. Dans P. Jonnaert, M. Ettayebi, & R. Opertti, *Logique de compétences et développement curriculaire: Débats, perspectives et alternative pour les systèmes éducatifs* (pp. 61-78). Paris: L'Harmattan.
- De Ketele, J.-M., & Gérard, F.-M. (2004, Juillet 19). La validation des épreuves d'évaluation selon l'approche par les compétences. *Mesure et Évaluation en Éducation*, 28(3), pp. 1-26.
- De Ketele, J.-M., & Roegiers, X. (1993). *Methodologie du recueil d'informations. Fondements des méthodes d'observations, de questionnaires, d'interviews et d'études de documents*. Bruxelles: De Boeck.
- De Ketele, J.-M., Gérard, F.-M., & Roegiers, X. (1997). L'évaluation et l'observation scolaires : deux démarches complémentaires. *Éducatives - Revue de diffusion des savoirs en Éducatives*(12), pp. 33-37.
- encyclopédie pour aujourd'hui*. (pp. 17-31). Paris: ESF Editeur.
- Diouf, P. B. (2021). *Alignement didactique entre outils d'évaluation sommative et directives du programme de SVT : cas d'épreuves administrées, dans l'enseignement moyen général du Sénégal, par les professeurs de l'Inspection d'Académie (IA) de Dakar*. Mémoire de Master en sciences de l'éducation non publiée, Université Cheikh Anta Diop de Dakar, Dakar.
- Diouf, P. B. (2020, Juillet). Analyse didactique d'outils d'évaluation sommative en SVT a la lumière de l'APC : cas d'épreuves administrées dans les inspections d'académie de Dakar et Saint-Louis. *Liens Nouvelle Série, I*(29), pp. 154-175.
- Diouf, P. B., & Dieng, B. D. (2020, Décembre). Evolution et enjeux actuels de l'évaluation des apprentissages et théories de l'apprentissage : un éclairage par les apports théoriques de l'évolution du statut de la connaissance de De Ketele (2008). *Liens Nouvelle Série*(30), pp. 125-144.
- Diouf, P. B., & Kane, I. (2021, Décembre). Objectifs d'apprentissage évalués dans les épreuves standardisées en SVT, de l'Inspection d'Académie (IA) de Dakar, de 2016 à 2020. *ASSEMPE\_Revue Universitaire des Sciences de l'Education*, 64-77.
- Durand, M.-J., & Chouinard, R. (2012). *L'évaluation des apprentissages. De la planification de la démarche à la communication des résultats*. Montréal, Québec: Editions Marcel Didier inc.
- Durand, M.-J., Sakho, I., & Ducharme, M. (2014). La situation de compétences: analyse des productions proposées en mathématique. Dans M.-J. Durand, & N. Loye, *L'instrumentation pour l'évaluation: La boîte à outils de l'enseignant évaluateur* (pp. 169-187). Québec: Marcel Didier inc.
- Gérard, F.-M. (2005, octobre 24-26). L'évaluation des compétences par des situations complexes. *Actes du Colloque de l'Admee-Europe, IUFM Champagne-Ardenne, Reims*.
- Gérard, F.-M. (2008). La complexité d'une évaluation des compétences à travers des situations complexes : nécessités théoriques et exigences du terrain. Dans M. Ettayebi, R. Opertti, & P. Jonnaert, *Logique de compétences et développement curriculaire : débats, perspectives et alternative pour les systèmes éducatifs* (pp. 167-183). Paris : pp.: L'Harmattan.
- Jonnaert, P. (2002). *Compétences et socioconstructivisme: Un cadre théorique*. Bruxelles: De Boeck.
- Machado, C. A. (2008). La pertinence dans un développement curriculaire. Dans M. Ettayebi, R. Opertti, & P. Jonnaert, *Logique de compétences et développement curriculaire: Débats, perspectives et alternative pour les systèmes éducatifs* (pp. 13-16). Paris: L'Harmattan.
- Mboup, M. (2003, Mars 19). Mémoire de DEA en Sciences de l'Education. *EVALUATION DES ACQUIS SCOLAIRES. Le paramétrage et l'habillage des épreuves peuvent-ils influencer les méthodes et stratégies d'enseignement/apprentissage des sciences de la Vie et de la Terre ?* Dakar, Sénégal: CUSE-ENS/UCAD.

- MEEPEM. (2008, Mai). PROGRAMME DES SCIENCES DE LA VIE ET DE LA TERRE de L'ENSEIGNEMENT MOYEN. Dakar, Sénégal: MEEPEM.
- MEN. (2015). *Programmes des sciences de la vie et de la terre de l'enseignement secondaire General. Classes de secondes (L et S), Premières (L2, S1 et S2), Terminales (L2, S1 et S2)*. Dakar: Ministère de l'éducation nationale du Sénégal.
- Roegiers, X. (2000, 2001). *Une pédagogie de l'intégration. Compétences et intégration dans l'enseignement* (éd. 1e et 2e). Bruxelles: De Boeck.
- Sakho, I. (2017, Novembre). *Élaboration et validation d'un modèle d'épreuve certificative en lecture et écriture, axé sur l'approche par compétences (APC) en 6e année du primaire au Sénégal*. Montréal: UNIVERSITE DE MONTREAL.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Québec, Canada: Editions du Renouveau pédagogique Inc.
- Tardif, J. (2006). *L'évaluation des compétences. Documenter le parcours de développement*. Montréal, Québec: Chenelière Education.
- Tarillon, L. (2006). Quelques apports des 5 premiers chapitres de Gérard Scallon. Grenoble, Isère, France. Récupéré sur [www.ac-grenoble.fr/disciplines/ses/Content/stages/FC\\_pedago\\_2007/Fiches\\_de\\_lecture/Scallon\\_evaluation\\_competence.htm](http://www.ac-grenoble.fr/disciplines/ses/Content/stages/FC_pedago_2007/Fiches_de_lecture/Scallon_evaluation_competence.htm)