

**Expérimentation d'une grille d'auto évaluation en ligne des stages pratiques en formation à l'enseignement de l'éducation physique et sportive en Algérie**

**Experimentation of an online self-assessment for practical training in teaching physical and sports education in Algeria**

Jaouad Alem, Ph.D<sup>1</sup>. et Hakim Hariti, Ph.D<sup>2</sup>.

<sup>1</sup>Université Laurentienne, Canada, [jalem@laurentian.ca](mailto:jalem@laurentian.ca)

<sup>2</sup>Université d'Alger 3, Algérie, [hariti.hakim@univ-alger3.dz](mailto:hariti.hakim@univ-alger3.dz)

**Résumé.**

En Algérie, Les étudiants-stagiaires en formation à l'enseignement de l'éducation physique et sportive sont exclusivement évalué et supervisés lors de leurs stages pratiques par des responsables universitaires et les superviseurs qui les accueillent dans les collèges et les lycées. Selon la recension des écrits, les faibles habiletés évaluatives des superviseurs et la simple prise en compte de la qualité du rapport de stage sont problématiques. Est-ce que les stagiaires algériens en EPS sont capables d'auto évaluer en ligne de façon valide, discriminante et précise leurs compétences professionnelles ? 150 stagiaires des instituts algériens d'EPS inscrits en master 2 (56 étudiantes, 94 étudiants) ont auto évalué leurs compétences avec un questionnaire composé de 26 items initiaux. La comparaison des moyennes montre que les auto évaluations sont plus élevées que les évaluations des rapports de stage ou que les moyennes académiques générales ; la comparaison des coefficients de variation montre qu'elles sont plus dispersées également. Après avoir éliminé 8 items déviants la consistance interne devient très élevée (Alpha de Cronbach = 0,94). La validité telle qu'estimée avec la corrélation entre l'impression générale du stagiaire et les trois moments des interventions (avant, pendant et après le stage) est démontrée dans la mesure où la corrélation la plus basse est de 0,95\*\*\*. Cette recherche a servi à proposer aux responsables universitaire une grille d'auto évaluation composée de 18 items, cette grille pourrait être utilisée dans un premier temps comme outil pour permettre d'effectuer un retour réflexif sur son expérience de stage. Il reste cependant à vérifier d'autres qualités métriques telle que la validité de construit par exemple.

**Mots Clés :** auto évaluation, qualités métriques, stage en enseignement de l'éducation physique et sportive.

**Abstract.**

In Algeria, student interns in training in physical education and sports education are exclusively evaluated and supervised during their practical internships by university officials and supervisors who host them in colleges and high schools. According to the literature review, the poor evaluation skills of supervisors and the simple consideration of the quality of the internship report are problematic. Are Algerian interns in physical education and sport capable of valid, discriminating and precise self-assessment online of their professional skills?

150 interns from Algerian institutes of sport enrolled in master 2 (56 females, 94 males) self-assessed their skills with a questionnaire of 26 initial items. The comparison of averages shows that the self-evaluations are higher than the evaluations of the internship reports or the general academic averages; comparing the coefficients of variation shows that they are also more dispersed. After eliminating 8 deviant items the internal consistency becomes very high (Cronbach's Alpha = 0.94). The validity as estimated with the correlation between the overall rating of the trainee and the three moments of the interventions (before, during and after the internship) is demonstrated to the extent that the lowest correlation is 0.95 \*\*\*. This research was used to propose to university officials a self-assessment grid made up of 18 items, this grid could be used as a first step as a tool to allow a reflective feedback on the internship experience. However, it remains to verify other metric qualities such as construct validity for example.

**Keywords:** self-assessment, metric qualities, interns in physical education and sports

## 1. Introduction

L'auto-évaluation des étudiants est un domaine important de recherche en éducation contemporaine et en psychologie de l'éducation (Falchikov, 1989; Falchikov et al. 1989). En effet, les recherches actuelles en pédagogie sur l'auto-évaluation, examinent les preuves empiriques accumulées et proposent de nouveau programme, méthodes, outils pour démontrer la fiabilité et la validité de l'auto-évaluation (Alem et al., 2009 et 2021; Brown et al., 2013; Dassé, 1989; Martel et al. 2011 ; Panadero et al., 2016).

Selon Evertson et al. (1985), la plus grande partie des acquis lors du stage pratique ne seront pas mobilisés en pratique enseignante réelle. « Il est alors sans doute plus que jamais nécessaire non seulement de discuter et de réviser le contenu et les objectifs des cours universitaires, mais aussi encadrer d'avantage l'apprentissage des étudiants et de développer les méthodes d'évaluation » (Pryor, 2010, page 266).

En effet, l'évaluation des compétences professionnelle des stagiaires ne devrait pas seulement concerner les simples savoirs comme les connaissances déclaratives, mais également et surtout l'acquisition des connaissances procédurales d'enseignement en référence au programme officiel telles que les méthodes et les processus développés par les stagiaires.

Certains superviseurs des stages qui sont responsables de la formation et de l'évaluation du stage pratique font peu d'observation de l'enseignement réalisée par le stagiaire et ne délivrent que des rétroactions souvent vagues et incomplètes (Tannehill et al., 1988 ; Ocansey, 1988). En effet, ces superviseurs s'attachent davantage aux démarches, aux méthodes, aux types de raisonnement mobilisés par les stagiaires qu'à des résultats ou qu'à des performances listées dans le référentiel (Paddeu et al., 2017).

Un référentiel de compétences doit s'assurer de refléter les compétences que mobilisent effectivement les enseignants, en intégrant les préoccupations de formation propres à l'institution universitaire, en explicitant les compétences à un niveau de généralité adapté, organisé et exploitable dans une perspective

d'évaluation de même qu'en anticipant les réalités professionnelles à venir (Chauvigné et al., 2010).

Selon St-Pierre (2004) « *Les objets à auto-évaluer peuvent être de diverses natures : la qualité d'un produit, d'une activité, d'un projet ; l'efficacité d'un processus ou d'une démarche ; des caractéristiques personnelles, ses points forts, ses points faibles, ses comportements, ses réactions, ses capacités, sa contribution, son intégration ; sa compétence dans un domaine spécifique, le degré, le niveau de développement de cette compétence, sa progression, son autonomie, l'état de ses ressources personnelles (connaissances, habiletés, attitudes), sa capacité à identifier et à exploiter des ressources externes pertinente* » (p. 35).

C'est en stage pratique que le futur enseignant se forme ; dans ce sens Abdallah-Pretceille (2003) affirme qu'à partir de sa propre expérience le stagiaire construit son identité professionnelle en s'appropriant les techniques d'objectivation de l'évaluation. Il s'agit donc pour lui de s'auto évaluer pour se former, pour forger son autonomie dans l'action qui lui permettra de mieux se connaître (Barbot et al., 2007).

En Algérie, il n'existe pas d'entente entre les universités et des lycées pour que des professeurs d'éducation physique et sportive accueillent et supervisent les stagiaires en formation à l'enseignement de l'EPS. Il serait donc hautement souhaitable de proposer une grille d'auto évaluation valide aux stagiaires algériens en EPS. Le but de cette recherche est d'analyser les qualités métriques de la grille d'auto-évaluation du stage pratique telle qu'utilisée à l'Université Laurentienne-Canada et administrée à une population d'étudiant algérien en EPS.

## **2. Cadre conceptuel**

Le milieu universitaire assiste à un changement de paradigme en matière de formation des professionnels. Boyer (2013) et De Ketele (2000) décrivent ce changement comme une évolution du rapport au savoir marquée par le délaissement d'une centration de l'enseignement vers une centration des apprentissages. En contexte de stage en enseignement, Chaubet et al. (2013) associent la compétence réflexive à l'engagement du stagiaire dans sa formation, à la prise de recul pour analyser les situations et à l'établissement de liens avec les connaissances issues de la recherche en éducation. Afin de vérifier si le stagiaire a atteint et intégré les compétences complexes attendues à la fin de sa formation, il devra pouvoir les réaliser dans des situations concrètes (Raïche, 1995).

Villeneuve (1994) définit le stage pratique comme 'une période de formation pratique qui s'insère dans le cours d'un programme d'études et qui s'effectue dans un milieu de travail préalablement choisi. C'est une expérience d'apprentissage supervisé visant l'acquisition des connaissances, habiletés et attitudes nécessaires à l'exercice d'une profession' (p. 19). Le stage pratique est une composante majeure de la formation pratique dans un programme d'études qui permet à l'étudiant d'acquérir, entre autres, de nouvelles connaissances, de développer ses compétences et de réfléchir sur son comportement en situation réelle (Dion, 2017).

Pour s'assurer que les étudiants auront développé et acquis les compétences exigées par le milieu professionnel, ils seront donc évalués, de façon formative ou/et formatrice, pendant le stage et de façon sommative/certificative au terme du stage. Cette évaluation à caractère formatif, même lorsqu'elle est certificative (Mottier-Lopez, 2010), a vraisemblablement un effet positif sur l'évaluation du stagiaire.

Boyer (2013) précise que l'utilisation d'échelles uniformes, tant pour l'évaluation formative que certificative, nécessite une interprétation arbitraire qui varie d'un formateur à l'autre, ce qui pourrait remettre en cause la validité et l'équité de l'évaluation. Surtout que l'enseignant associé à la supervision du stage est la personne qui exerce la plus grande influence sur le stagiaire (Dodds, 1985). Cependant, la sélection des enseignants associés est l'un des aspects les plus négligés dans les programmes de formation à l'enseignement (Griffin et al., 1981).

A ce sujet, Gouin et Hamel (2015) précisent que malgré le référentiel des compétences professionnelles pour l'accompagnement des stagiaires québécois, l'accompagnement en formation pratique présente encore des zones grises. De plus, les objectifs spécifiques de la formation du stagiaire et les objets d'évaluation peuvent varier selon le milieu d'appartenance du formateur évaluateur, le milieu scolaire ou le milieu universitaire (Portelance et Caron, 2015).

Desbiens (2009) a présenté quelques éléments majeurs de la problématique des compétences des encadreurs professionnels qui supervisent les stagiaires en enseignement au Québec :

- Les encadreurs professionnels sont rarement choisis sur la base de leurs habiletés en supervision parce que ces dernières sont imprécises et difficiles à mesurer.
- L'insuffisance, voire l'absence de contrôle de leurs qualifications, de leurs compétences et de leurs croyances sur les façons de devenir enseignant représente indéniablement un facteur susceptible de limiter la portée éducative d'un stage.
- Dans certains milieux, le fait d'avoir reçu une formation pèserait d'ailleurs assez peu dans le fait d'être recruté pour exercer cette fonction. Une étude publiée au Québec par Lacroix-Roy, Lessard et Garant en 2003 révèle en outre que 81,5% des encadreurs professionnels ont reçu moins de 30 heures de formation en cette matière, la moyenne s'établissant à 14,8 heures de formation.
- Koster, Korthagen, et Wubbels (1998) rapportent que les encadreurs ne sont pas toujours outillés pour réinvestir les concepts, les théories et les principes du champ de l'éducation dans leur travail auprès des stagiaires ce qui ne permet pas d'assurer la continuité entre ce qui est enseigné dans le programme de formation et dans les milieux de stage (Mitchell et Schwager, 1993).
- Les encadreurs professionnels sont peu enclins à observer de manière soutenue et systématique les stagiaires lorsqu'ils enseignent et manifestent ou rapportent des difficultés à évaluer de manière critique et discriminante la qualité des pratiques d'enseignement (Spallanzani et al., 1995).

Sur le plan pédagogique, l'auto évaluation est une des opérations cognitives de l'approche systémique empruntée à Richterich (1985). Barbot et Rosen-Reinhardt (2007) précisent que : « tout au long de sa formation, le futur-enseignant est donc amené à effectuer, de manière individuelle, un retour réflexif sur les cours suivis, sur les stratégies développées et sur les objectifs qu'ils se fixent » (p. 111).

L'auto-évaluation est définie plus précisément comme un processus par lequel les stagiaires 1) surveillent et évaluent la qualité de leur réflexion et de leur comportement lors des apprentissages et 2) identifient des stratégies qui améliorent leur compréhension et leurs compétences (McMillan et Hearn, 2008). Autrement dit, l'auto-évaluation se produit lorsque les stagiaires jugent leur propre enseignement pour améliorer les performances et lorsqu'ils identifient les écarts entre les performances actuelles et souhaitées.

Que nous enseignent la recension des écrits sur la capacité des stagiaires en formation à l'enseignement à auto évaluer leur propre expérience de stage ?

Godbout (1988) soutient que, de façon générale, les élèves sont tout à fait capables d'auto évaluer leurs apprentissages si les critères d'évaluation sont clairement expliqués.

Se basant sur une étude transversale conçue pour évaluer l'auto-perception des compétences professionnelles des enseignants italiens d'éducation physique, Tul et al.(2019) ont construit un questionnaire auto-administré conçu pour examiner un large éventail de compétences générales et spécifiques auprès de 484 enseignants du nord-est de l'Italie. Il apparaît que les approches didactiques représentent le premier facteur de l'efficacité et expliquerait 31,3% de la variance totale. Par ailleurs, les enseignants se sentent insuffisamment compétents dans certains domaines, tels que l'utilisation des technologies de l'information et de la communication, la communication en langue étrangère, les travaux de recherche scientifiques, la prise d'initiative et l'esprit du groupe.

Falchikov et Goldfinch (2000) rapportent des coefficients de corrélation moyens des évaluations par les pairs dans la pratique professionnelle ( $r = 0,54$ ), ce qui peut être plus problématique car dans les autres domaines d'évaluation académique, les corrélations sont bien plus élevées ( $r = 0,75$  pour les évaluations des productions écrites ;  $r = 0,83$  pour les évaluations des productions orales).

Tousignant et DesMarchais (2002) ont analysé la capacité d'auto-évaluation des étudiants stagiaires par rapport à leur propre performance académique dans un programme médical d'apprentissage par problèmes. 70 étudiants inscrits à leur troisième année d'un programme de quatre ans ont participé volontairement à l'étude en complétant un questionnaire qui mesurait leur capacité d'auto-évaluation à deux occasions différentes : 1) avant l'examen oral standardisé pour prédire leur performance, et 2) après un examen d'estimation de leur performance académique. L'exactitude de l'auto-évaluation a été étudiée par une analyse corrélationnelle entre l'auto-évaluation et la performance. L'étude a révélé que le pré-examen d'auto-évaluation n'était pas précis par rapport à la performance de l'examen oral ( $r$  variait de 0,042 à 0,243) ; cependant, la précision s'est légèrement améliorée lorsque l'étudiant s'auto évaluait a posteriori, mais restait peu significative ( $r$  variait alors de 0,257 à 0,334). En

conclusion, les étudiants ont démontré une faible capacité à auto évaluer leur propre performance.

Madrazo et al. (2018) ont analysé si l'auto-évaluation différait selon le genre des étudiants en médecine de l'Université d'Ottawa par rapport à l'évaluation de leurs pairs dans le cadre d'un examen clinique à faible enjeu. Pour examiner les différences entre les deux genres, les scores moyens de la note globale d'auto-évaluation ont été comparés aux scores de la note globale de l'évaluation par les pairs en utilisant pour la comparaison le test « *t* » d'échantillons indépendants. Dans l'ensemble, les scores d'auto-évaluation des étudiantes étaient significativement plus faibles que ceux de l'évaluation par les pairs (3,88 vs 4,67;  $p < 0,001$ ,  $d = 1,18$ ), alors qu'aucune différence significative n'a été trouvée entre les scores d'auto-évaluation et ceux de l'évaluation par les pairs pour les étudiants (4,64 vs 4,80;  $p = 0,228$ ,  $d = 0,32$ ). Cette étude fournit une preuve que la sous-estimation de l'auto-évaluation chez les étudiantes est observable même dans un contexte de formation à faibles enjeux.

Adib-Hajbaghery et al. (2012) ont réalisé une analyse corrélacionnelle entre l'auto-évaluation des compétences cliniques des stagiaires en sciences infirmières et l'évaluation de leurs enseignants auprès de 75 infirmières internes de l'Université des sciences médicales de Kashan. Ces chercheurs utilisèrent un questionnaire composé de 190 items liés à 15 catégories de compétences infirmières nécessaires pour prendre en charge différents soins médicaux et chirurgicaux des patients gravement malades. Une relation significative, positive et plutôt élevée a été observée entre le score moyen global d'auto-évaluation et les scores que les étudiants ont reçus de leurs enseignants dans les cours cliniques ( $r = 0,78$ ,  $p = 0,001$ ) et théoriques ( $r = 0,51$ ,  $p = 0,001$ ).

Compte tenu de ces quelques résultats de recherche, il devient légitime d'adresser la question de recherche suivante : est-ce que les stagiaires algériens en EPS sont capables d'auto évaluer de façon valide, discriminante et précise leurs compétences professionnelles ?

### **3. Méthodologie**

#### **3.1. L'échantillon des stagiaires**

Un échantillon de 150 stagiaires des instituts algériens d'EPS inscrits en master 2 (56 étudiantes, 94 étudiants) et réalisant un stage pratique dans divers établissements scolaires au niveau national (57 stages réalisés en enseignement moyen et 93 en enseignement secondaire). Le tableau 1 présente les caractéristiques de l'échantillon selon ces variables. Ils ont été invités durant le mois de février 2020 à compléter en ligne un questionnaire composé de 26 items d'autoévaluation du stage pratique. Ils ont complété en même temps une lettre de consentement éclairé où ils ont été informés de la nature de la recherche.

#### **3.2. Les variables à l'étude**

La grille d'évaluation (annexe 1) comporte 4 variables : une variable 'critère' (l'impression générale de la prestation du stagiaire (IGP, évaluée avec 10 items) et trois autres variables mesurées avec 16 items : le degré de préparation du stagiaire avant la séance d'intervention (avant, 5 items), la qualité de l'interaction avec son environnement pendant la séance (pendant, 8 items) et la

qualité de son intervention juste après la fin de la séance (après, 3 items). Les items sont évalués sur une échelle de Lickert à 3 modes (1 : non satisfaisant ; 2 : satisfaisant et 3 : très satisfaisant).

Tableau 1 : la taille de l'échantillon selon le genre et selon le palier d'enseignement

		<u>Palier</u>		<u>Total</u>
		<u>Collège</u>	<u>Lycée</u>	
<u>Genre</u>	<u>F</u>	<u>24</u>	<u>32</u>	<u>56</u>
	<u>H</u>	<u>33</u>	<u>61</u>	<u>94</u>
<u>Total</u>		<u>57</u>	<u>93</u>	<u>150</u>

### 3.3. Les quatre qualités métriques analysées

#### 3.3.1- Le degré de difficulté des items et leur degré de discrimination

Ces deux qualités métriques ont été évaluées respectivement par le biais de la comparaison des moyennes et du calcul du coefficient de variation CV (= 100\* écart type/ moyenne). Ce coefficient permet en effet de comparer des degrés de dispersion de variables qui n'ont pas été mesurées avec les mêmes échelles. C'est le cas des évaluations auto rapportées des rapports de stage (M stage) et de la moyenne générale académique (M acad), toutes les deux autos rapportées sur 20 points ; alors que la moyenne aux items est sur 3 points.

#### 3.3.2- La consistance interne des items :

Cette qualité métrique se définit comme le degré de constance qu'offrent les réponses des individus aux items variés d'un instrument de mesure. L'intérêt de la consistance interne porte sur l'analyse de la relation entre les items et le construit. La consistance interne est estimée par le coefficient alpha de Cronbach (1951). Bien qu'il existe plusieurs autres façons de calculer un indice de fidélité, cet indice est le plus populaire (Cortina, 1993 ; Croker et Algina, 1986).

#### 3.3.3- La validité concomitante :

Ce type de validité est l'une des deux formes de validité de critère (Organ et Ryan, 1995). La validité concomitante permet d'établir une relation entre un critère et une mesure à l'intérieur d'une même période de temps. Nous avons utilisé comme mesure concomitante de l'évaluation des compétences des stagiaires en intervention : l'impression générale de la prestation du stagiaire (en anglais : *overall rating*). Ce critère est, selon Shechtman (1998), un score qui possède des qualités métriques meilleures que les autres indicateurs. Bujold (2002) utilise le terme d'évaluation holistique pour définir ce score général et démontre également que ce score est plus valide que les autres types d'évaluation.

Le score général est mesuré par les 10 premiers items de la grille d'évaluation des stagiaires. La validité concomitante est donc le degré d'assurance que les

trois mesures (avant, pendant et après l'intervention des stagiaires) sont bien corrélées avec cet autre indicateur reconnu valide du succès en intervention : l'impression générale de la prestation du stagiaire durant son stage.

#### 4. Analyses et interprétations des résultats.

##### 4.1- L'analyse descriptive des variables à l'étude, le degré de difficulté des items et le degré de discrimination des items

Tableau 2 : les statistiques descriptives des variables étudiées en termes de moyenne (M), d'effectif (N), d'écart type (S) et de coefficient de variation (CV)

	M acad	M stage	overall rating	avant	pendant	après
M	11,76	13,32	2,24	2,18	2,17	2,16
N	150	150	150	150	150	150
S	1,30	0,94	0,36	0,35	0,46	0,83
CV	11,05	7,06	16,07	16,06	21,20	38,43

Le tableau 2 révèle que les moyennes académiques et au stage même si elles sont auto rapportées sur 20 points, sont bien plus faibles que l'impression générale, le degré de préparation, la qualité de l'interaction avec son environnement et la qualité de l'intervention juste après la fin de la séance tels que perçus par le stagiaire. En effet, rappelons les 26 items ayant été auto évalués sur une échelle de Lickert à 3 modes (1 : non satisfaisant ; 2 : satisfaisant et 3 : très satisfaisant) ; les auto évaluations sont donc jugées entre satisfaisantes et très satisfaisantes alors que les évaluations des responsables universitaires sont à peine au-dessus de la moyenne (entre 11,76 et 13,32 sur 20).

Le coefficient de variation montre bien combien les évaluations des responsables universitaires sont bien plus homogènes.

Ces résultats ne sont pas étonnants, Alem et al. (2021, 2012, 2009 et 2008) avaient aussi démontré que les rapports de stage étaient bien plus 'sévèrement' évalués par les responsables universitaires que les auto évaluations par les stagiaires.

##### 4.2- La consistance interne des items :

Robinson, Shaver, et Wrightsman (1991) et Clark et Watson (1995) proposent un seuil minimal acceptable du coefficient alpha de Cronbach de 0,80. La consistance interne des 26 items est donc acceptable ( $\alpha = 0,81$ ) ; les 26 items du questionnaire sont donc reliés entre eux.

Par contre comme le montre le tableau 3, il est possible de supprimer certains items déviants pour augmenter ce coefficient ; en effet certains ne sont pas ou presque pas du tout discriminants (la corrélation (item/item total corrigé) est

presque nulle). Pour rappel, Flanagan propose un seuil de  $r=0,20$  pour considérer qu'un item est discriminant.

Tableau 3 : la consistance interne des 26 items originaux

	Corrélation (Correction Item-Total )	Alpha de Cronbach si l'item est supprimé
it1	,081	,814
it2	,085	,814
it3	-,164	,821
it4	,021	,816
it5	,621	,795
it6	,592	,797
it7	-,003	,819
it8	,071	,880
it9	,621	,795
it10	,592	,797
it11	,621	,795
it12	-,003	,819
it13	,094	,814
it14	,592	,797
it15	,621	,795
it16	,592	,797
it17	,621	,795
it18	,592	,797
it19	,621	,795
it20	,592	,797
it21	,621	,795
it22	,592	,797
it23	,621	,795
it24	,592	,797
it25	,621	,795
it26	,592	,797

En éliminant 8 items déviants, il devient possible d'augmenter la consistance interne des 18 items restants (alpha = 0,94). Le tableau 4 indique les degrés de discrimination de chacun des items restants.

Tableau 4 : la consistance interne et les degrés de discrimination des 18 items restants

	Corrélation (Correction Item-Total )	Alpha de Cronbach si l'item est supprimé
it5	,679	,939
it6	,659	,939
it9	,679	,939
it10	,659	,939
it11	,679	,939
it14	,659	,939
it15	,679	,939
it16	,659	,939
it17	,679	,939
it18	,659	,939
it19	,679	,939
it20	,659	,939
it21	,679	,939
it22	,659	,939
it23	,679	,939
it24	,659	,939
it25	,679	,939
it26	,659	,939

### 4.3- La validité concomitante

Le tableau 5 présente la validité concomitante définie comme les corrélations entre l'évaluation de la prestation du stagiaire avant, pendant et après son intervention auprès du groupe classe avec la mesure critère : l'impression générale de la prestation du stagiaire.

Tableau 5 : la validité concomitante

	évaluation du stagiaire avant l'intervention	évaluation du stagiaire pendant l'intervention	évaluation du stagiaire après l'intervention
impression générale de la prestation du stagiaire (ou overall rating)	0,95***	1***	0,95***

Puisque les corrélations sont toutes significatives et très élevées, la validité concomitante est démontrée. Par contre, si nous avons pris comme critère la moyenne académique ou encore l'évaluation du rapport de stage par le

responsable universitaire, les corrélations ne seraient pas significatives comme le montre le tableau 6.

**Tableau 6** : les corrélations entre l'impression générale du stagiaire et les évaluations du responsable universitaire

	évaluation du rapport de stage	La moyenne académique
impression générale de la prestation du stagiaire	0,09 ; p=0,27 non significatif	-0,01 ; p=0,91 non significatif

Le tableau 6 montre bien combien les responsables universitaires des stages ne peuvent se baser que sur les rapports de stage pour évaluer la qualité de l'expérience pratique réalisée, la qualité des interventions professionnelles n'est pas du tout évaluée. Concernant l'évaluation du rapport de stage, nos résultats sont consistants aussi bien avec nos propres analyses d'autres cohortes de stagiaires marocains et sénégalais (Alem et al., 2008) qu'avec ceux de la méta analyse de Fraser, Walberg, Welch et Hattie (1987) qui avaient rapporté que la corrélation entre le succès académique et le succès professionnel ne dépassaient pas 0,16.

### Conclusion

L'analyse de la consistance interne des 26 items qui auto évaluent la prestation des stagiaires révèle que ceux-ci sont reliés entre eux ( $\alpha = 0,91$ ), mais qu'il est possible d'augmenter la fidélité de cette mesure en supprimant 8 items déviants ( $\alpha = 0,94$ ).

La validité concomitante est démontrée dans la mesure où la corrélation la plus basse est significative ( $R=0,95^{***}$ ) entre la mesure critère (i.e l'impression générale du stagiaire) et les 18 items restants de la grille d'évaluation.

L'analyse des trois évaluations selon le coefficient de variation indique que les auto évaluations des stagiaires sont trop élevées et très dispersées par rapport aux évaluations des responsables universitaires.

En conclusion, alors que cette étude avait pour point de départ un questionnaire composé de 26 items, les analyses statistiques suggèrent d'en retenir 18 pour proposer aux responsables des stages pratiques en EPS des items qui pourraient servir aux stagiaires comme outil d'auto évaluation et de retour réflexif sur leur expérience de stage. Une pratique courante dans les universités canadiennes est d'accorder entre 10 % et 20 % des points à l'auto évaluation à condition que le stagiaire justifie des auto évaluations.

Il conviendrait de performer d'autres analyses statistiques sur les évaluations des cohortes des stagiaires des années académiques à venir pour vérifier si la nouvelle grille d'évaluation épurée des items inutiles possède encore des qualités métriques similaires voire supérieures.

Il conviendrait également d'analyser d'autres types de validité telle que la validité de construit par exemple.

### Références bibliographiques

- Abdallah-Preteceille, M. (2003). *Former et éduquer en contexte hétérogène : pour un humanisme du divers*. Paris : Anthropos.
- Adib-Hajbaghery, M., Karbasi-Valashani, K., et Heidari-Haratmeh, A. (2012). Correlation of clinical skills self-assessment of nursing internship trainees with their teachers' evaluation. *Nurs Midwifery Stud*, 1 (2), 94-99.
- Alem, J., et Hariti, H. (2021). Une grille d'évaluation des stages pratiques en formation à l'enseignement de l'éducation physique et sportive. *Propulsion*, (34) 1, 48-51.
- Alem, J., et Boudreau-Larivière, C. (2012). Evaluation of an Internship Assessment Grid for Francophone Physical and Health Education Student Interns. *The Canadian Journal for the Scholarship of Teaching and Learning*, 3 (1). <http://dx.doi.org/10.5206/cjsotl-rcacea.2012.1.5>
- Alem, J.; Boudreau-Larivière, C., et Kpazai, G. (2009). La validation empirique des items d'une grille d'évaluation des stages pratiques en intervention en activité physique et santé. 77ième Congrès de l'association francophone pour le savoir (ACFAS): La science en français... une affaire capitale ! 11-15 mai 2009, Ottawa, [http://www.acfas.ca/congres/a\\_propos.html](http://www.acfas.ca/congres/a_propos.html)
- Alem, J.; Kpazai, G.; Sané, O.; Dadouchi, F., et Couture, R.; (2008). La prédiction de l'engagement professionnel des futurs enseignants d'éducation physique et sportive. *Annales Scientifiques de l'Université "Alexandru Ioan Cuza" Iasi, Roumanie, Section Sciences de l'éducation*. 12, 125-146.
- Barbot, M.J., et Rosen-Reinhardt, E. (2007). Élaboration d'un processus d'auto-évaluation et outillage en formation d'enseignants au niveau master. *Recherches en Didactiques : Les Cahiers Théodile*, (8), 100-126.
- Boud, D., et Falchikov, N. (1989). Quantitative studies of student self-assessment in higher-education: A critical analysis of findings. *Higher Education*, 18(5), 529-549.
- Boyer, L. (2013). *Co construction d'un modèle cognitif de l'apprentissage d'une compétence en vue d'assurer la validité et l'équité de son évaluation : Le cas de la compétence « Exercer un jugement clinique infirmier »* (Thèse de doctorat). Université de Sherbrooke, Québec.
- Brown, G. T. L., et Harris, L. R. (2013). Student self-assessment. In J. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367-393). Thousand Oaks, Californie: Sage
- Bujold, N. (2002). La supervision pédagogique. Vue d'ensemble. In M. Boutet et N. Rousseau (dir), *Les enjeux de la supervision pédagogique des stages* (pp. 9-23). Ste-Foy : Les Presses de l'Université Laval.
- Chaubet, P., Correa Molina, E., et Gervais, C. (2013). Considérations méthodologiques pour aborder la compétence à «réfléchir » ou à «faire réfléchir» sur sa pratique en enseignement. *Phronesis*, 2(1), 28-40.
- Chauvigné, C., Coulet, J.C., et Gosselin, P. (2006). *Journées : Compétences, emploi et enseignement supérieur*. Rennes : Université de Bretagne.

- Clark, L.A. et Watson, D. (1995). Constructing Validity: Basic Issues in Objective Scale Development. *Psychological Assessment*, 7, ( 3), 309-319
- Cortina, J.M. (1993). What is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, 78, ( 1), 98-104
- Croker, L., et Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Belmont, CA: Wadsworth.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 31, 93-96.
- Dassé, B. (1989). *L'implication des étudiants dans le processus d'évaluation formative des apprentissages en éducation physique*. Thèse de doctorat, Faculté des Sciences de l'Éducation, Département d'éducation physique, Université Laval.
- De Ketele, J.M. (2000). En guise de synthèse : Convergences autour des compétences. In C. Bosman, F. M. Gerard, et X. Roegiers (dir.), *Quel avenir pour les compétences* (pp.187-191). Bruxelles : De Boeck Université.
- Desbiens, J.F. (2009). Bien enseigner ne s'apprend pas tout seul ! *Séminaire pensée et culture*. École des sciences de l'éducation. Université Laurentienne. Sudbury (Ontario).11 février 2009.
- Dion, I. (2017). De la théorie à la pratique : l'exemple des stages de la Division de la gestion de documents et des archives de l'Université de Montréal. *Archives*, 46 (2), 91-112.
- Dodds, P. (1985). Delusion of "Worth-it-ness": Field Experiences in Elementary Physical Education Teacher Education Programs. In H. Hoffman and J. Rink (Eds). *Physical Education Professional Preparation: Insights and Issues* (pp. 90-109). Washington D.C.: American Alliance for Health, Physical Education, Recreation and Dance.
- Evertson, C.M., Hawley, W. D. et Zlotnik, M. (1985). Making a difference in educational quality though teacher education. *Journal of Teacher Education*, 40, 56-58.
- Falchikov, N., et Boud, D. (1989). Student Self-Assessment in Higher Education: A Meta-Analysis. *Review of Educational Research*, 59 (4), 395-430.
- Falchikov, N., et Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research*, 70 (3), 287-322.
- Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Synthesis of educational productivity research. *Journal of Educational Research*, 11(2), 145-252.
- Godbout, P. (1988). Les prélèvements d'information en activité physique : la mesure. In Gaëtan Morin (Ed.), *La supervision de l'intervention en activité physique*. (pp. 233-299). Montréal : Gaëtan Morin Éditeur.
- Guyton, E., et Farokhi, E. (1987). Relationships among academic performance, basic skills, subject matter knowledge and teaching skills of teacher education graduates. *Journal of Teacher Education*, 38, 37-42.
- Gouin, J.A., et Hamel, C. (2015). La perception de formateurs de stagiaires quant au développement et à l'évaluation formative des quatre compétences liées à

- l'acte d'enseigner. *Canadian Journal of Education / Revue canadienne de l'éducation*, 38(3), 1-27.
- Griffin, G.A., Hughes, R., Defino, M.E., et Barnes, S. (1981). *Student Teaching: A Review. Report no. 9007*. Austin, TX: Research and Development Centre for Teacher Education.
- Madrazo, L., Lee, C. B., McConnell, M., et Khamisa, K. (2018). Self-assessment differences between genders in a low-stakes objective structured clinical examination (OSCE). *BMC research notes*, 11(1), 393. <https://doi.org/10.1186/s13104-018-3494-3>.
- Martel, D.; Gagnon, J. et Godbout, P. (2011). « Auto-évaluation d'élèves et prédiction de l'évaluation de l'éducateur physique à leur égard : incidences sur leurs comportements en éducation physique », *eJRIEPS*, 22. <https://doi.org/10.4000/ejrieps.4721>
- McMillan, J.H., et Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Horizons éducatifs*, 87 (1), 40-49.
- Mottier Lopez, L. (2010). Perspectives conclusives. Repenser la complémentarité des évaluations formative et certificative des compétences professionnelles. In L. Béclair, C. Lebel, N. Sorin et A. Roy (éds.), *Régulation et évaluation des compétences en enseignement* (p. 237-249). Ottawa, Ontario: Presses Universitaires d'Ottawa.
- Ocansey, R. (1988). The Effects of a Behavioral Model of Supervision on the Supervisory Behaviors of Cooperating Teacher. *Journal of Teaching in Physical Education*, 8, 46-63.
- Organ, D. W. et Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel Psychology*, 48, 775-802.
- Paddeu, J., et Veneau, P. (2017). L'approche par « compétences » dans l'enseignement professionnel français : impasse et non-sens pédagogique au regard des pratiques enseignantes. *Cahiers de la recherche sur l'éducation et les savoirs*, 16, 95-115.
- Panadero, E., Brown, G.T.L. et Strijbos, J.W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, 28(4), 803-830.
- Portelance, L., et Caron, J. (2015). Objet de tension en contexte d'alternance : l'évaluation des apprentissages du stagiaire. Communication au symposium « *Apprendre et enseigner en contexte d'alternance : vers la définition d'un noyau conceptuel propice à l'exploration de problématiques et d'objets diversifiés* », Réseau éducation et formation (REF), Montréal, Québec.
- Pryor, J., et Crossouard, B. (2010). Challenging formative assessment: disciplinary spaces and identities. *Assessment & Evaluation in Higher Education*, 35(3), 265-276.
- Raïche, G. (1995). *L'approche par compétences, qu'est-ce que ça change ? La mesure de l'intégration des apprentissages dans une perspective cognitiviste*. Actes du 7e colloque de l'A.R.C. (pp. 135-139). Québec : Cégep de Jonquière.

- Richterich, R. (1985). *Besoins langagiers et objectifs d'apprentissage*. Paris : Hachette.
- Robinson JP, Shaver PR, & Wrightsman LS (1991). *Measures of personality and social psychological attitudes*. San Diego: Academic Pr.
- Shechtman, Z. (1998). Agreement between Lay Participants and Professional Assessors: Support of a Group Assessment Procedure for Selection Purposes. *Journal of Personnel Evaluation in Education*, 12(1), 5-17.
- St-Pierre, L. (2004). L'habileté d'autoévaluation : pourquoi et comment la développer ? *Pédagogie collégiale*, 18 (1), 33-38.
- Tannehill, D., et Zakrajsek, D. (1988), What's Happening in Supervision of Student Teachers in Secondary Physical Education. *Journal of Teaching in Physical Education*, 8, 1-12.
- Tousignant, M., et DesMarchais, J.E. (2002). Accuracy of student self-assessment ability compared to their own performance in a problem-based learning medical program: a correlation study. *Advances in Health Sciences Education*, 7 (1), 19-27.
- Tul, M., Leskošek, B., et Kovač, M. (2019). The Professional Competences of Physical Education Teachers from North-Eastern Italy. *CEPS Journal*, 9 (1), 103-120.
- Villeneuve, L. (1994). *L'encadrement du stage supervisé*. Montréal, Québec : Éditions Saint-Martin.

**Annexe 1** : Les 26 items qui composaient la grille d'auto évaluation initiale

<b>A. Impression générale :</b>	
La voix, le volume et la tonalité étaient-ils clairs et inspirants ?	Item 1
Le stagiaire avait-il créé et avait-il maintenu des sessions agréables et productives ?	Item 2
Le stagiaire pouvait-il modifier son modèle d'entraînement /d'enseignement en fonction du niveau de compétence du groupe ?	Item 3
La contribution globale du stagiaire était-elle positive ?	Item 4
Le stagiaire avait-il une attitude professionnelle ?	Item 5
Le stagiaire était-il ponctuel?	Item 6
Quelle était la qualité de sa tenue vestimentaire?	Item 7
Quel était son degré de respect vis-à-vis les apprenants?	Item 8
Quel était son degré de charisme?	Item 9
Le stagiaire avait-il une capacité d'écouter, de comprendre les besoins exprimés et de reformuler les questions posées?	Item 10
<b>B. AVANT le début de la séance d'enseignement:</b>	
À quel point le stagiaire était-il préparé pour chaque session d'entraînement/ d'enseignement?	Item 11
Les compétences visées étaient-elles formulées de façon claire? Correspondaient-elles aux niveaux des apprenants ?	Item 12
Les exercices prévus étaient-ils progressifs (du plus faciles au plus difficiles) ? Présentaient-ils un bon niveau de défi?	Item 13
Quelle était la qualité du choix et l'utilisation optimale du matériel ?	Item 14

Les plans des séances étaient-ils clairs? Sont-ils illustrés avec des schémas?	Item 15
<b>C. PENDANT la séance :</b>	
Les pratiques/séances d'enseignement étaient-elles bien organisées, claires et progressives ?	Item 16
La prise en main du groupe-classe (appel, explication de la séance, gestion du matériel)	Item 17
Le déroulement de la séance proprement dite : démonstrations claires, au bon moment, selon les besoins, démonstrations illustrées avec des exemples	Item 18
Le stagiaire fournissait aux athlètes/élèves des feedback appropriés et appliquait les corrections nécessaires	Item 19
Le stagiaire était en mesure d'obtenir la coopération des apprenants	Item 20
Le stagiaire était en mesure d'obtenir un temps d'apprentissage (active learning task) optimal.	Item 21
Le stagiaire effectuait des balayages visuels réguliers et complets (se place de façon à observer tous les apprenants)	Item 22
A la fin de la séance: synthèse de la leçon avec les apprenants.	Item 23
<b>D. APRES la séance :</b>	
Le stagiaire était-il capable de réfléchir sur son action, d'être pro actif ?	Item 24
Quelle était la qualité de son rapport avec son superviseur?	Item 25
Quelle était la qualité des questions posées au superviseur ?	Item 26