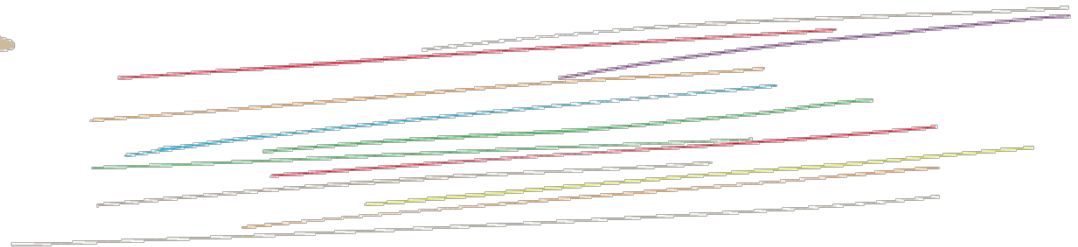




26^e congrès RANACLES
Évaluation(s) et autonomisation dans les centres de langues : pratiques et enjeux – 22-24 novembre 2018

Les outils et logiciels statistiques au service de l'évaluation certificative en anglais

Ecole doctorale
Sciences de l'Homme et de la Société
Université Lille Nord de France



Jean-François Brouttier
vendredi 23 novembre 2018

**Des outils
statistiques
complexes**

**Paradigme,
méthodologies et
objet de recherche**

**Évaluation et
autonomisation**

Constat

un lexique pointu en statistique

▀ **variable**

▀ **within-groups variable**

▀ **between-groups**

▀ **null hypothesis**

▀ **p value**

▀ **alpha level**

Constat

Peu de collègues à l'aise. Séminaires internes dans les laboratoires de linguistiques. (STL - M. Capliez, 2015)

➤ **variable** = basically, it means a factor

➤ **within-groups variable** = the "change" factor within an individual group; it can have several "levels"

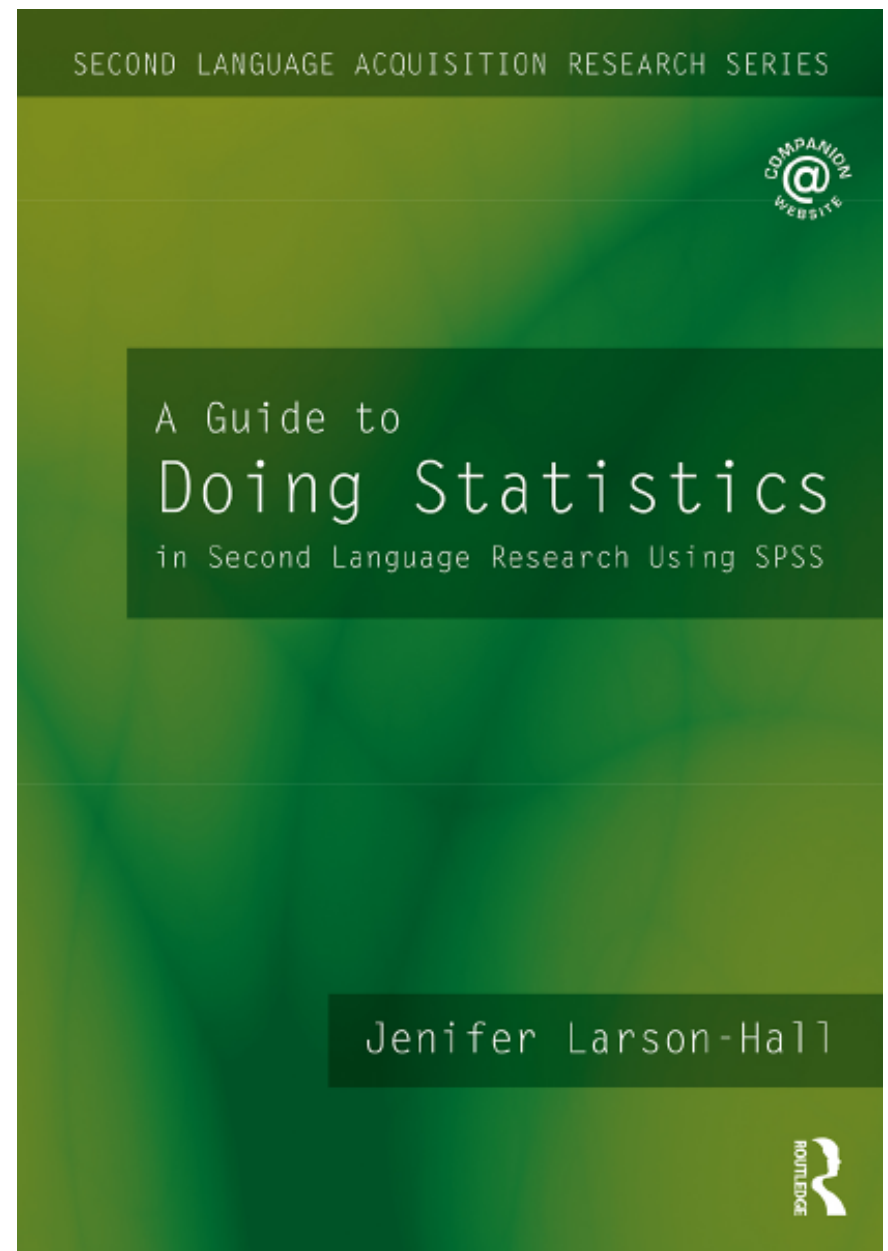
➤ **between-groups** = the "change" factor from one group to another

➤ **null hypothesis** = there is no significance

➤ **p value** = the final number that you obtain with a statistical test and which tells you whether the null hypothesis is validated or rejected

alpha level = the level of the p value from which we can consider that the null hypothesis is rejected or not. If the p value is superior to 0.05, then we must accept the null hypothesis (= > no significance). If the p value is inferior to 0.05, then we reject the hypothesis, and therefore we can conclude that there is a significant effect somewhere

Une référence qui fait autorité en la matière



De nombreuses pistes à explorer

5 Choosing a Statistical Test

- 5.1 Statistical Tests that Are Covered in This Book 130
- 5.2 A Brief Overview of Correlation 131
- 5.3 A Brief Overview of Partial Correlation 132
- 5.4 A Brief Overview of Multiple Regression 133
- 5.5 A Brief Overview of the Chi-Square Test of Independence 135
- 5.6 A Brief Overview of T-Tests 136
 - 5.6.1 A Brief Overview of the Independent-Samples T-Test 137
 - 5.6.2 A Brief Overview of the Paired-Samples T-Test 138
- 5.7 A Brief Overview of One-Way ANOVA 139
- 5.8 A Brief Overview of Factorial ANOVA 141
- 5.9 A Brief Overview of ANCOVA 143
- 5.10 A Brief Overview of Repeated-Measures ANOVA 143

Des outils et logiciels complexes à maîtriser

The screenshot displays two windows from the IBM SPSS Statistics software. The left window, titled "Employee data.sav [DataSet16] - IBM SPSS Statistics Data Editor", shows a data grid with columns for id, gender, bdate, and educ. The right window, titled "*Output15 [Document15] - IBM SPSS Statistics Viewer", displays the results of a Univariate Analysis of Variance. It includes a pie chart titled "Current Salary by Employment Category" and a table titled "Between-Subjects Factors".

Current Salary by Employment Category

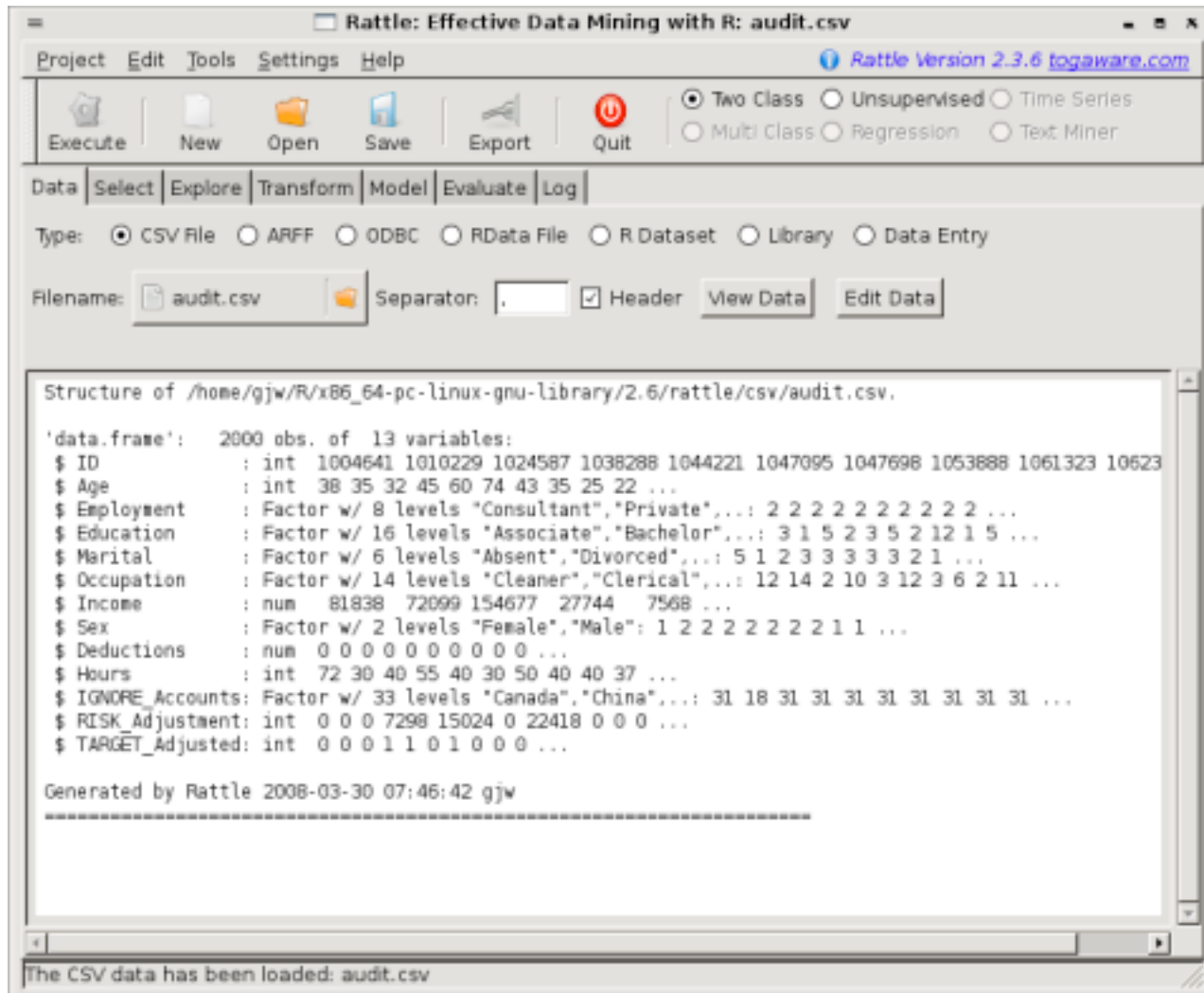
Employment Category	Percentage
Clerical	61.94%
Custodial	5.12%
Manager	32.94%

Univariate Analysis of Variance

Between-Subjects Factors

	Value Label	N	
Gender	f	Female	216
	m	Male	258
Employment Category	1	Clerical	363
	2	Custodial	27
	3	Manager	84
Minority Classification	0	No	370
	1	Yes	104

Des outils et logiciels complexes à maîtriser

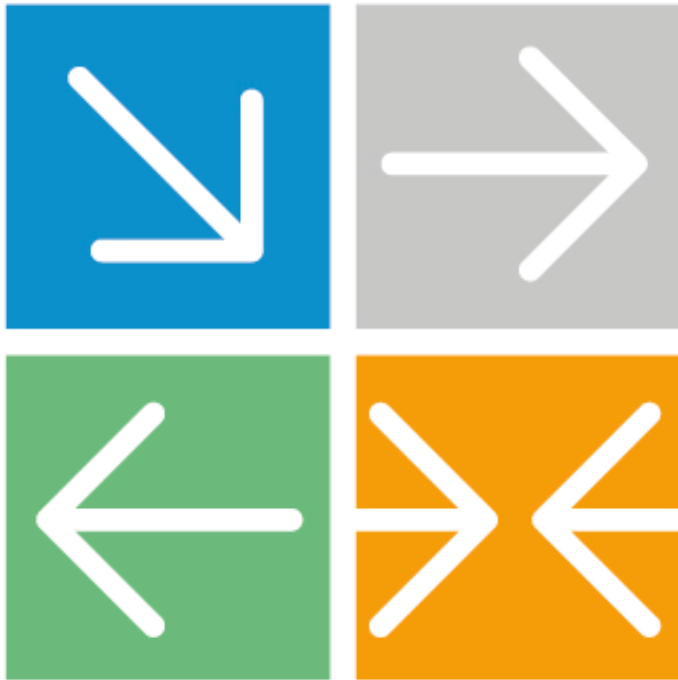


Des formations chronophages



2 journées pour une simple introduction
5 à 6 journées pour un début de maîtrise de l'un
ou l'autre outil

Quel paradigme ?



Certificat de compétences en langues de l'enseignement supérieur

Constat

L'hégémonie du mot test en anglais

▀ Language Tests

▀ Language Testers

ALTE : Association of Language Testers in Europe

▀ Language Testing

Postulat de départ

une polysémie en anglais, deux concepts en français

➤ **Test**

➤ **Certification**

Test, subst. masc.

stella.atilf.fr/

Le **T**résor
de la **L**angue
Française
informatisé

➤ *PSYCHOL., MÉD.* Épreuve élaborée en vue d'évaluer chez un individu son niveau mental, intellectuel, le développement de certaines facultés (physiques, morales, affectives), ses aptitudes (naturelles ou acquises) ou de tracer le portrait de sa personnalité.

Test, subst. masc.

stella.atilf.fr/

Le **T**résor
de la **L**angue
Française
informatisé

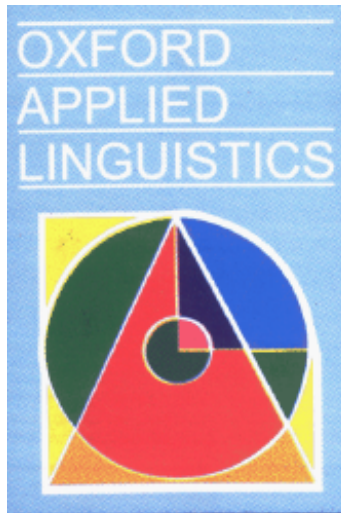
➤ *STAT.* Méthode permettant d'évaluer une population dont les caractéristiques ne peuvent être connues que par échantillonnage.

- 1893 « épreuve permettant de mesurer des phénomènes ou des aptitudes dans des conditions expérimentales précises » (ici p. réf. aux travaux du psychologue A. Binet)

- Empr. à l'angl. *test*, issu de l'a. fr. *test* « pot » (v. *têt* et *test*¹) et désignant, en m. angl., une coupelle de métallurgiste servant à isoler les métaux précieux, d'où son empl. pour désigner ce qui permet de déterminer la qualité ou la pureté de quelque chose (1594 ds *NED*)

- puis ce qui permet de mettre la foi de quelqu'un à l'épreuve et en partic. le serment introduit en 1673.
- dans différents domaines sc. et techn. au XIX^e s. au sens de « expérience, épreuve » et « élément nécessaire, ce qui permet une expérience »

Test
vs.
certification



Fundamental Considerations in Language Testing

Lyle F. Bachman

Certify (v.)

etymonline.com



ONLINE ETYMOLOGY DICTIONARY

mid-14c., "to declare the truth of," also "to vouch for or confirm" (an official record, etc.), from Old French certefiier "make certain, witness the truth of" (12c.). Also used in Middle English in broader senses of "inform, give notice to; instruct, to direct; to designate."

➤ *ENSEIGN. SUPÉRIEUR.* Acte délivré par l'enseignement supérieur, attestant de connaissances acquises, d'aptitude à quelque chose, mais ne correspondant pas à un degré et ne conférant pas de titre. (ex. donnée dans *Encyclop. pratique de l'éduc. en France*, 1960, p. 311).

Certificat

stella.atilf.fr/

Le **T**résor
de la **L**angue
Française
informatisé

- Assurance, garantie donnée à quelqu'un.
- 1380 *certificat* « document écrit attestant un fait »

Certification

stella.atilf.fr/

Le **T**résor
de la **L**angue
Française
informatisé

- Action de certifier par écrit.
- 1310 *certificacion*. Empr. au lat. médiév. *certificatio* « assurance, confirmation »

- certification
- Domaine : normalisation
- Attestation réalisée par une tierce partie, relative à des produits, des processus, des systèmes ou des personnes.

➤ En français, une certification est une formule signée, généralement accompagnée d'un sceau, qui établit l'identité d'un acte, d'un document ou d'une signature, ou encore, la conformité d'un produit à une norme.

➤ En France, le terme *certification* est aussi employé pour désigner l'opération qui authentifie les compétences et le savoir-faire d'un individu par rapport à une norme formalisée et qui valide des acquis expérimentiels.

Test vs. certification



Test vs. certification

évaluation certificative

◆ **Domaine**: Éducation-Formation. ◆ **Définition**: Évaluation sommative sanctionnée par la délivrance d'une attestation. ◆ **Équivalent étranger**: –

Source: *Journal officiel* du 16 juin 2007.

Histoire des méthodologies de l'enseignement des langues

Quelle histoire de l'évaluation ?

Evaluation

Spolsky (2000) > histoire des courants et méthodologies du *testing*

- les tests, par opposition aux certifications, permettent de situer la performance du candidat sur une échelle grâce à un score.
- 3 premières approches : pré-scientifique, structuraliste psychométrique et sociolinguistique-intégrative

Histoire des méthodologies de l'enseignement des langues

Quelle histoire de l'évaluation ?

Evaluation

Spolsky (2000) > histoire des courants et méthodologies du *testing*

- les tests, par opposition aux certifications, permettent de situer la performance du candidat sur une échelle grâce à un score.
- 3 premières approches : pré-scientifique, structuraliste psychométrique et sociolinguistique-intégrative

Histoire des méthodologies de l'enseignement des langues

Quelle histoire de l'évaluation ?

Tardieu (2013)

L'approche pré-scientifique s'inscrit dans une logique d'examens et de juges mais manque d'objectivité et de fiabilité. [...] Elle manque aussi de validité. [...] L'approche pré-scientifique est celle pratiquée au Baccalauréat, par exemple, qui ne s'appuie ni sur des théories du langage ou de l'apprentissage, ni sur une méthodologie scientifique.

Histoire des méthodologies de l'enseignement des langues

Quelle histoire de l'évaluation ?

Tardieu (2013)

L'approche structuraliste-psychométrique intervient pour répondre à la demande d'objectivité et de fiabilité, dans un souci d'évaluer la maîtrise du système linguistique (Mc Namara 2000), avec en contre partie un effet d'atomisation et de décontextualisation.

Histoire des méthodologies de l'enseignement des langues

Quelle histoire de l'évaluation ?

Tardieu (2013)

l'approche sociolinguistique intégrative : à partir des années 1960, la demande est en faveur d'une approche intégrée, testant à la fois la compétence linguistique fine et les savoir-faire (CO, CE, PO, PE) malgré les difficultés en termes de coût et d'attribution de scores en productions orales ou écrites. (Tardieu 2013 : 68)

Histoire des méthodologies de l'enseignement des langues

Quelle histoire de l'évaluation ?

Tardieu propose l'ajout d'un quatrième modèle :

pour compléter la classification de Spolsky, une quatrième approche, non psychométrique, devrait être ajoutée : 'l'approche sociolinguistique intégrative de type actionnel'. Cette approche s'appuie sur une conception socio-pragmatique de la communication appliquée à l'enseignement [...] (Tardieu 2013 : 71)

Histoire des méthodologies de l'enseignement des langues

Quelle histoire de l'évaluation ?

- certification CLES dans le prisme d'un *scoring plan* ou *scoring key* dont Purpura, nous apporte des éléments de méthodologie (à propos du First Certificate) :

I scored the *FCE Anchor Test* “right” or “wrong” according to a detailed scoring key and a set of guidelines for scoring. I scored each test twice, for accuracy, and entered the marks into the computer. I then matched and merged the data sets, dropping 278 cases due to missing data, leaving a total sample of 1,382 cases. (Purpura 1997 : 299–300)

Formation des enseignants

Quel héritage pour l'évaluation ?

Maquettes de formation : exemple

MEEF 2nd degré dans 2 ESPE
Master Didactique des langues
Aux États-Unis

Master 2 MEEF - Mention Second degré, parcours Professeur d'ANGLAIS stagiaire
 CMAN12 version 131

Année universitaire 2017-18 (mise à jour : juin 2017)

SEMESTRE 3 (CMAN032R)

Élément pédagogique	CM	TD	TP	ECTS	Coeff.	Session 1			
						Examen terminal			
						Étudiant en régime continu	Étudiant en régime ponctuel	Nature examen	Durée
UE 1 – CMANJ01U : Synergie langue/culture				5					
CMLV3A5X : Langue vivante (un choix)									
<i>LMEAG31E : Communication en Anglais</i>		18							
<i>CMLV3B3E : LV Espagnol</i>		18							
CMANJ21E : Faits de civilisation contemporaine et analyse du discours des media	20			5		Oui	Non	-	Dossier
UE 2 –CMANK01U : Didactique				12					
CMANK11E : Aide à la préparation de cours	10	20		6		Non	Oui	-	Oral (APP)
CMANK21E : Développement des 5 activités langagières dans le parcours d'apprentissage	10	20		6		Non	Oui	-	Ecrit 2h
UE 3 – CMANL01U : Contexte d'exercice du métier				3					

Formation des enseignants

		DEFINITION DU NOMBRE DE CHOIX PARMI				REPARTITION DES HEURES					
	Libellé :	Nature	Dispositif	Nombre de choix à réaliser	Nombre de choix possibles	Nombre d'heures CM par matière	Nombre d'heures TD/TP par matière	Nb d'heures total pour l'étudiant	CM Mutualisé Ou/Non	Nombre de parcours mutualisés	Crédits ECT
SEMESTRE 3						19	104	123			30
UE1	Disciplinaire					5	10	15			3
EC1	Approfondissement disciplinaire	Obligatoire	CM+TD			5	10	15			3
UE2	Didactique					9	21	30			14
EC1	Conception d'une séquence de cours (interventions ponctuelles possibles des LA/IPR)	Obligatoire	CM+TD			3	7	10			5
EC2	Mise en œuvre et gestion de classe (interventions ponctuelles possibles des LA/IPR)	Obligatoire	CM+TD			3	7	10			5
EC3	L'évaluation (interventions ponctuelles possibles des LA/IPR)	Obligatoire	CM+TD			3	7	10			4

Formation des enseignants

***ULille Master Sciences du langage – Parcours Didactique des langues
– Anglais Langue Vivante Étrangère***

***M1 un EC « Le CECR et l'évaluation » d'une UE « L'intervention
didactique » – 18h***

***ULCO Master Didactique des Langues / FLE et Anglais et ingénierie
de formation :***

M1 une UE « Didactique de l'évaluation » à 18h

M2 une UE « Evaluation et audit de formation » à 16h

Formation des enseignants

Un exemple aux États-Unis

Master of Education in Applied Linguistics – L2 Assessment Track

à Columbia Teachers College,

Taught and Research MA

[Programme Sheet](#)

Quel paradigme ?

Où se situe la rupture ? Quelle mutation ?

Les efforts se sont concentrés, du côté anglo-saxon, sur le rôle et l'efficacité de différents modèles d'évaluation formative pour réguler tant l'enseignement que l'apprentissage. Du côté francophone, les efforts se sont cristallisés autour de l'évaluation d'objectifs d'apprentissage formulés en termes de compétences. Ceux-ci font appel à la mobilisation des connaissances et des ressources internes et externes de l'élève dans l'exécution de tâches complexes s'inscrivant dans des familles de situations authentiques et hautement contextualisées. (Laveault 2013 : 31-32)

Contexte

Comment se conçoivent les tests et certifications en langues ?

entre courant méthodologique, secteur de formation et internationalisation des parcours...

▀ Secteur LANSAD

Langues pour spécialistes d'autres disciplines

▀ L'évaluation

enjeu important de tout parcours universitaire
peu de formation pour les enseignants en langues

▀ Tests

TOEIC, TOEFL, IELTS, BULATS
score chiffré sur une échelle
attestation valable deux ans

Contexte

Comment se conçoivent les tests et certifications en langues ?

entre courant méthodologique, secteur de formation et internationalisation des parcours...

■ Certifications

Cambridge First Certificate (B2), Advanced (C1), Proficiency (C2)
diplôme
valable à vie

■ le CLES

Alternative de l'enseignement supérieur français
MESRI, CPU, Coordination Nationale CLES
58 universités centres d'examens
inscription dans la perspective actionnelle (CECRL)

Enjeux

➤ **Évaluation certificative**

vers une évaluation authentique
évaluation interne vs externe
pratiques des Language Testers

➤ **Perspective actionnelle**

« considère avant tout l'utilisateur et l'apprenant d'une langue comme des acteurs sociaux ayant à accomplir des tâches (qui ne sont pas seulement langagières) dans des circonstances et un environnement donnés, à l'intérieur d'un domaine d'action particulier. » (Conseil de l'Europe 2001:15)

Enjeux

➤ Compréhension de l'oral

première sous-épreuve du CLES

questionnaire ressemblant parfois à du testing

tâche ?

➤ Définition d'une évaluation certificative dans une perspective actionnelle

➤ Enjeu de la mise en place de la certification

➤ Méthodologie de conception des épreuves

Questionnement

prenant appui sur un cadre théorique initial

➤ **Test vs Certification**

(Tardieu 2013 : 67)

➤ **La spécificité de l'activité langagière de compréhension de l'oral**

(Cornaire 1998 : 50–51)

(Guichon 2004 : 76–77)

➤ **La spécificité du CLES : scénario et lien entre les activités langagières**

(Nunan 2004 : 4)

Questionnement

prenant appui sur un cadre théorique initial

▀ La pertinence des cadres théoriques habituels pour traiter la question de l'évaluation en langues

(Bachman 1990)

(Spolsky 2000)

(Doucet 1996 :137–271)

(Tardieu 2013 : 68)

▀ La perspective actionnelle dans l'évaluation certificative

(Purpura 1997 : 299-300)

(Laveault 2013 : 31–32)

Méthodologie

Entre recherche-action et recherche développement

➤ **Analyse critique de l'existant**

sujets d'examens CLES diffusés depuis 2004

➤ **Analyse de résultats aux épreuves CLES**

traces d'éléments de l'activité de compréhension de l'oral dans la tâche de production écrite

➤ **Indicateurs de réussite à la compréhension de l'oral**

Comment mesurer la réussite à la CO ?

Performance (*scoring*) ou tâche ?

➤ **Expérimentation et groupe témoin**

pour une expérimentation d'un CLES bis en appliquant les méthodes du *scoring* lors de la conception du sujet (*scoring plan, blueprint*)

Orientation épistémologique et méthodologique ?

Quels cadres ? Quelles finalités ?

► **Language Testing?**

> bibliométrie comparative en cours entre publications du monde anglophone et les autres publications (francophones ? européennes ?). Postulat : différence d'1 pour 1000.

► **Si Language Testing > analyses statistiques**

équipes ? ETS à Princeton 30-40 chercheurs. Idem à Cambridge.
En France ? Dans le monde francophone ?

► **Rapport investissement / pertinence de la recherche ?**

Dissémination de la recherche dans la formation des enseignants ?
Accompagnement des concepteurs de sujets CLES ?
Préconisations pour des certifications ?

Inscription de la recherche dans son contexte

Quels cadres ? Quelles finalités ?

- **Holisme ?**
- **Approche systémique ?**
- **Théorie de la pensée complexe ?**
- **Perspective actionnelle et travaux du Conseil de l'Europe**
- **Redéfinition de l'activité de médiation dans le volume complémentaire du CECRL :**

Outre la médiation inter linguistique, cette approche prend aussi en compte la médiation liée à la communication et à l'apprentissage ainsi que la médiation sociale et culturelle. Cette approche plus ample a été choisie pour sa pertinence dans des classes de plus en plus diverses, en relation avec la diffusion d'EMILE (Enseignement de Matières par l'Intégration d'une Langue Etrangère) et parce que l'on considère de plus en plus que la médiation fait partie de tout apprentissage et tout particulièrement de l'apprentissage de la langue.

Inscription de la recherche dans son contexte

Quels cadres ? Quelles finalités ?

➤ **Redéfinition de l'activité de médiation dans le volume complémentaire du CECRL :**

Outre la médiation inter linguistique, cette approche prend aussi en compte la médiation liée à la communication et à l'apprentissage ainsi que la médiation sociale et culturelle. Cette approche plus ample a été choisie pour sa pertinence dans des classes de plus en plus diverses, en relation avec la diffusion d'EMILE (Enseignement de Matières par l'Intégration d'une Langue Etrangère) et parce que l'on considère de plus en plus que la médiation fait partie de tout apprentissage et tout particulièrement de l'apprentissage de la langue.

Bibliographie

- Bachman, Lyle F. 1990. *Fundamental Considerations in Language Testing*. (Oxford Applied Linguistics). Oxford: Oxford University Press.
- Conseil de l'Europe. 2001. *Cadre européen commun de référence pour les langues*. Paris : Didier. (22 March, 2017).
- Cornaire, Claudette. 1998. *La compréhension orale*. (Didactique Des Langues Étrangères). Paris : Clé International.
- Doucet, Patrick. 1996. *L'évaluation en question*. Compiègne : Université de Technologie de Compiègne.
- Guichon, Nicolas. 2004. *Compréhension de l'anglais oral et TICE : les conditions d'un apprentissage signifiant*. Nantes : Université de Nantes Linguistique.
- Laveault, Dany. 2013. Plus de convergences que de divergences : rencontre des conceptions francophone et anglo-saxonne de l'évaluation des apprentissage. *Le français dans le monde : Recherches et applications* 2013(53). 31-41.

Bibliographie

- Ministère de l'Éducation Nationale. 2000. *Arrêté du 22 mai 2000 portant création du certificat de compétences en langues de l'enseignement supérieur*. [https:// www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000582492](https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000582492) (13 October, 2017).
- Nunan, David. 2004. *Task-based Language Teaching*. (Cambridge Language Teaching Library). Cambridge (U.K.): Cambridge University Press.
- Purpura, James E. 1997. An Analysis of the Relationships Between Test Takers' Cognitive and Metacognitive Strategy Use and Second Language Test Performance. *Language Learning* 47(2). 289–325.
- Spolsky, Bernard. 2000. Language Testing in the Modern Language Journal. *The Modern Language Journal* 84:4 Special Issue : A Century of Language Teaching and Research: Looking Back and Looking Ahead, Part). 536–552.
- Tardieu, Claire. 2013. Les tests de langues : une nécessité ? *Le français dans le monde : Recherches et applications* 2013(53). 67–79.

Merci de votre
attention