



HAL
open science

Validité de construit et validité prédictive d'une épreuve de jeu de rôle dans le cadre d'un centre d'évaluation chez des officiers de Marine.

Borteyrou Xavier, Rasclé Nicole Bruchon-Schweitzer, Anne Congard

► To cite this version:

Borteyrou Xavier, Rasclé Nicole Bruchon-Schweitzer, Anne Congard. Validité de construit et validité prédictive d'une épreuve de jeu de rôle dans le cadre d'un centre d'évaluation chez des officiers de Marine.. Perspectives différentielles en psychologie, 2008. hal-03437142

HAL Id: hal-03437142

<https://hal.science/hal-03437142>

Submitted on 12 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Validité de construit et validité prédictive d'une épreuve de jeu de rôle dans le cadre d'un centre d'évaluation chez des officiers de Marine.

Borteyrou Xavier¹

Rasclé Nicole

Bruchon-Schweitzer Marilou

Collomb Pierre

Congard Anne

Problématique

Les centres d'évaluation (CE) comprennent en général plusieurs types de mises en situation individuelles et collectives (in basket, entretiens structurés, épreuve de groupe, jeu de rôle) au cours desquelles les candidats sont évalués par des psychologues et par des managers. Ces CE sont utilisés soit pour recruter des candidats soit pour établir un bilan de compétences.

Aujourd'hui, la littérature concernant la validité des CE est pléthorique. Si notre travail se rapporte à la validité de construit d'une épreuve de jeu de rôle, il nous semble important d'aborder rapidement la validité de contenu et la validité prédictive des CE car il est difficile d'envisager ces trois validités de manière séparée (Binning & Barrett, 1989).

La validité de contenu consiste à se demander si les comportements évalués dans le cadre d'un CE sont représentatifs des conduites futures. Étant donné que les CE sont construits sur le principe de la cohérence comportementale (Wernimont & Campbell, 1968), la validité de contenu est le point fort des CE. La démarche d'un CE consiste effectivement à faire correspondre point par point, les qualités évaluées, et les qualités à prédire (Callinan & Robertson, 2000), ceci grâce aux analyses de poste.

La validité prédictive consiste à se demander si les évaluations faites durant un CE permettent de prédire la performance ultérieure des évalués. Si l'on considère l'évaluation globale d'un CE, le coefficient de validité prédictive moyen retrouvé dans la littérature est de 0.37 (Gaugler, Rosenthal, Thornton, & Bentson, 1987). Ce coefficient est satisfaisant mais moins élevé que celui obtenu grâce à d'autres méthodes de recrutement comme par exemple les échantillons de travail ($r=.54$), les entretiens structurés ($r=.51$) et les tests de connaissances professionnelles ($r=.48$)

¹ Laboratoire de psychologie EA 3662. Université Bordeaux 2. 3 ter place de la victoire. 33 076 Bordeaux cedex. E-mail : xavier.borteyrou@u-bordeaux2.fr

(Schmidt & Hunter, 1998). Ces résultats posent un problème épineux : un CE (évaluation totale) posséderait une validité prédictive inférieure à celles des épreuves qu'il contient (Hoeft & Schuler, 2001). Autrement dit, l'information apportée par une évaluation globale serait moins prédictive vis-à-vis de la performance professionnelle que la connaissance des résultats à chacune des épreuves. Ce paradoxe a été récemment soulevé dans le cadre d'une méta-analyse menée par Arthur et al. (2003). Il est donc important d'isoler ce qui est réellement évalué dans les épreuves d'un CE en s'intéressant à leur validité de construit.

La validité de construit consiste à se demander quels sont les construits psychologiques évalués dans le cadre d'un CE. Or, depuis 1979, les différentes études menées ne parviennent pas à isoler précisément ces construits. (Lievens, 2002; Lievens & Klimoski, 2001). Ce constat fait dire à Russel & Domm (1995) : « les CE sont prédictifs de la réussite professionnelle mais on ne sait pas ce qu'on mesure ! ». Dans le cadre de la présente recherche, nous avons pris en compte des solutions méthodologiques proposées par la littérature pour améliorer la validité de construit des CE (Lievens & Klimoski, 2001). Ces solutions incluent le recours à des grilles comportementales qui vont permettre de rendre la procédure d'évaluation plus descriptive qu'évaluative (Thornton & Zorich, 1980) et la formation des évaluateurs à l'utilisation de ces grilles.

Méthode

Notre population comprend 125 officiers de marine évalués entre 1998 et 2001. Chacun d'eux a été évalué à l'aide de grilles comportementales par des officiers de marine plus anciens et des psychiatres formés à l'utilisation de ces grilles. Toutes les précautions méthodologiques citées dans la littérature ont été prises (formation des évaluateurs, analyses de postes, grilles comportementales très précises, standardisation des épreuves).

Les épreuves de Jeu de Rôle, ainsi que les grilles utilisées ont été les mêmes entre 1999 et 2001. Chaque épreuve est évaluée par 2 personnes qui remplissent d'abord séparément une grille d'observation puis se réunissent pour coter de manière consensuelle. Nous disposons donc de 291 observations sur l'ensemble de ces années.

Dans un premier temps, nous avons réalisé une analyse descriptive des items afin de ne retenir que ceux qui se distribuent sur les trois modalités. Nous avons éliminé les items dont le skewness approchait de 3 ou -3 et le kurtosis était supérieur à 4. Cette première étape nous a conduit à supprimer 8 items non discriminatifs sur 50. Nous avons soumis les 42 items restants à une analyse factorielle en maximum de vraisemblance suivie de rotations obliques de type promax. Cette méthode d'extraction nous a permis d'identifier 6 facteurs dont la valeur propre est supérieure à 1 (critère de Kaiser). Ces résultats sont présentés au tableau 1.

Résultats

Le premier facteur correspond clairement à une attitude de confrontation agressive et à une attitude autoritaire (chercher à gagner en s'opposant à l'autre, agresser, se poser comme détenteur de la vérité) qui évoque le type de leadership « autoritaire »

décrit par Lewin (1975). L'alpha de Cronbach de ce facteur est de 0.81 ce qui est satisfaisant.

Les items qui saturent le deuxième facteur décrivent dans quelle mesure le stagiaire prend en compte les sentiments et les besoins de l'autre, caractéristique déjà retrouvée par Arthur et al (2003) sous le nom de considération/conscience des autres. En outre, ce facteur est très homogène avec un alpha de Cronbach de 0.87.

Le troisième facteur ne comporte que des items saturés négativement. Il renvoie à une attitude correspondant à la volonté d'éviter le problème, de " noyer le poisson ", attitude qui avait également été mise en évidence par les analyses de poste. L'intitulé du pôle positif de cette dimension devrait donc correspondre à une attitude franche et directe dans le cadre de la négociation. L'homogénéité de ce troisième facteur est moins satisfaisante avec un alpha de Cronbach de 0.64 mais il s'agit là d'une dimension originale (peu retrouvée dans la littérature sur les CE) et il nous semble intéressant de la conserver.

Le quatrième facteur comporte également des items saturés négativement dont le contenu renvoie à un manque d'assurance, à une remise en question voire à une certaine anxiété. Avec un alpha de Cronbach de 0.49, on ne peut pas considérer qu'il s'agit d'un facteur homogène.

Le cinquième facteur correspond clairement à une expression orale en termes d'aisance, de clarté. Ce facteur est fréquemment retrouvé dans les CE sous la dénomination de « communication orale » (Arthur, Day, McNelly, & Edens, 2003). Seul un item (être attentif et concentré) ne correspond pas précisément à la thématique de ce facteur. L'homogénéité de ce facteur est satisfaisante avec un alpha de Cronbach de 0.87.

Le sixième facteur fait état de difficultés dans le discours, de confusion et comprend seulement deux items qui apparaissent sémantiquement cohérents. Une vérification de la consistance interne de ce facteur indique cependant qu'il n'est pas homogène avec un alpha de Cronbach de 0.36. Nous avons donc décidé d'exclure ce facteur du reste de nos analyses.

Discussion.

Nous sommes parvenus à isoler quatre dimensions interprétables et qui présentent des qualités psychométriques satisfaisantes. Ces dimensions (attitude autoritaire, considération, franchise, communication orale) ont été identifiées de manière exploratoire ce qui est nouveau. Notre travail trouve cependant des limites dans le fait que deux dimensions sur six (attitude affirmée et confusion) ne présentent pas les qualités psychométriques requises en termes d'homogénéité. De plus, nous avons mené des analyses factorielles sur l'ensemble des observations alors qu'il aurait fallu dans l'idéal mener des analyses séparées pour les évaluateurs psychologues et les évaluateurs de la Marine Nationale. On peut effectivement penser, avec Sagie & Magnezy (1997) que les évaluations se structurent différemment selon le type d'observateurs (les managers identifieraient moins de dimensions que des psychologues car plus soumis à l'effet de halo). Malheureusement, notre effectif ne nous permettait pas de mener des analyses

factorielles séparément. Ces résultats sont malgré tout encourageants puisque la validité de construit est considérée comme le « talon d'Achille des CE » depuis le début des années 80. Nous avons mis en évidence que sous certaines conditions, la validité de construit des CE peut être améliorée sans toutefois être comparable à celle des tests de personnalité ou d'intelligence.

Arthur, W., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion related validity of assessment center dimension. *Personnel Psychology*, 56, 125-154.

Binning, J. F., & Barrett, G. V. (1989). Validity of personnel decision: a conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology*, 74(3), 478-494.

Callinan, M., & Robertson, I. T. (2000). Work sample testing. *International Journal of Selection and Assessment*, 8(4), 248-260.

Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology*, 72(3), 493-511.

Hoelt, S., & Schuler, H. (2001). The conceptual basis of assessment centre ratings. *International Journal of Selection and Assessment*, 9(1), 114-123.

Lewin, K. (1975). *Psychologie dynamique*. Paris: PUF.

Lievens, F. (2002). Trying to understand the different pieces of the construct validity puzzle of assessment centers: An examination of assessor and assessee effects. *Journal of Applied Psychology*, 87(4), 675-686.

Lievens, F., & Klimoski, R. J. (2001). Understanding the assessment centre process: where are we now. *International Reviews of Industrial and Organizational Psychology*, 16, 245-286.

Russel, C. J., & Domm, D. R. (1995). Two field tests of an explanation of assessment center validity. *Journal of Occupational and Organizational Psychology*, 68, 25-47.

Sagie, A., & Magnezy, R. (1997). Assessor type, number of distinguishable dimension categories, and assessment center construct validity. *Journal of Occupational and Organizational Psychology*, 70, 103-108.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implication of 85 years of research findings. *Psychological Bulletin*, 124(2), 262-274.

Thornton, G. C. I., & Zorich, S. (1980). Training to improve observer accuracy. *Journal of Applied Psychology*, 65, 351-354.

Wernimont, P. F., & Campbell, J. P. (1968). Signs, samples, and criteria. *Journal of Applied Psychology*, 52, 372-376.

Tableau 1 : Saturation des items de la grille du jeu de rôle sur les 6 facteurs extraits par analyse factorielle par maximum de vraisemblance, suivie de rotations promax.

	F 1	F 2	F 3	F 4	F 5	F 6
Valeur propre	6.75	3.70	1.89	1.69	1.17	1.02
Pourcentage de variance expliquée	15.70	8.60	4.40	3.92	2.71	2.37
Chercher à gagner en s'opposant à l'autre	0,72					
Agresser	0,65					
Se poser comme détenteur de la vérité	0,62					
Montrer que l'autre a tort	0,61					
Réfuter d'emblée	0,52					
Chercher à aboutir en dominant	0,51					
Contester le propos de l'autre	0,48					
Montrer que l'on a raison	0,43					
Chercher à aboutir en affaiblissant	0,41					
Faire preuve de malveillance pour aboutir	0,30					
Orienter son esprit vers la compréhension de l'autre		0,74				
Tenir compte des réactions des interlocuteurs		0,68				
Accorder de l'estime à ce qu'il dit		0,63				
Avoir une attitude d'ouverture		0,61				
Fait preuve d'une conception interactive		0,60				
Accepter et intégrer d'autres idées		0,60				
Chercher à comprendre		0,60				
Entrer dans la logique de l'autre		0,58				
Engager une discussion constructive		0,55				
Avoir le réflexe de questionner		0,52				
Etre chaleureux		0,41				
Acquiescer d'emblée		0,34				
Faire preuve de ruse pour aboutir			-0,64			
Chercher à aboutir en manipulant			-0,56			
Se perdre dans les détails			-0,54			
Esquiver la logique de l'interlocuteur			-0,47			
Encaisser ce que dit l'autre, subir				-0,57		
Etre pris au dépourvu par les questions				-0,49		
S'engager à douter				-0,48		
Présenter des signes d'anxiété				-0,39		
S'exprimer avec aisance					0,63	
Etre attentif et concentré					0,63	
Présenter un discours structuré					0,56	
Parler de manière concise et claire					0,56	
S'exprimer en développant son discours					0,55	
Chercher à expliquer plutôt qu'à influencer					0,50	
Se baser sur des faits pour argumenter					0,45	
Réaliser un dosage équitable du temps de parole					0,30	
Parler de manière confuse						0,76
Ratés dans le discours						0,57
Alpha de Cronbach	0.81	0.87	0.64	0.49	0.78	0.36