AIPTLF – BOLOGNA 2004
13ème Congrès de psychologie du travail et des organisations

ETUDE PROSPECTIVE DE LA VALIDITÉ DES ITEMS DE DÉSIRABILITÉ SOCIALE D'UN TEST DE PERSONNALITÉ INSPIRÉ DU MODÈLE Á CINQ FACTEURS

Anne CONGARD [1], Pascal ANTOINE[2], Pierre COLLOMB[3], Stephan IVANCHAK[4]

Communication : Orale C066

Thème : Recrutement et sélection

[1] Université de Provence - Aix- Marseille I - Centre de Recherche PsyCLÉ (EA 3273) – 29, avenue Robert Schuman – 13621 Aix en Provence Cédex 1.
Courriel : anne.congard@laposte.net.
[2] Université Lille 3 – UPRES Temps, Emotions et Cognitions – UFR de Psychologie – BP 149 – F59653 – Villeneuve d'Ascq Cedex. Courriel : pascal.antoine@univ-lille3.fr.
[3] Marine Mobilité Brest – 8, rue Colbert, BP17, 29240 BREST.
Courriel :clb421@club.internet.fr.
[4] Université de Nice-Sophia antipolis – Laboratoire de Psychologie expérimentale et quantitative – Pôle universitaire St Jean d'Angely – 24, avenue des diables bleus – 06 357 NICE Cedex 4. Courriel : tivanchak@free.fr.

ETUDE PROSPECTIVE DE LA VALIDITÉ DES ITEMS DE DÉSIRABILITÉ SOCIALE D'UN TEST DE PERSONNALITÉ INSPIRÉ DU MODÈLE Á CINQ FACTEURS

Anne CONGARD, Pascal ANTOINE, Pierre COLLOMB, Stephan IVANCHAK

Résumé

Dans le cadre de la construction d'un test de personnalité inspiré du modèle à cinq facteurs, à visée de recrutement, il nous a semblé important de construire une échelle de désirabilité sociale (DS) permettant de quantifier ces biais de réponses.
Nous allons décrire dans cette étude les différentes étapes de construction de cette échelle et les méthodes employées pour tester la validité de ces items. Les résultats nous permettent, d'une part, d'élaborer une échelle de désirabilité sociale adaptée et, d'autre part, d'identifier les items de l'inventaire qui sont le plus sujet à la DS.

Mots-clés : désirabilité sociale, personnalité, autoduperie, hétéro-duperie, big five.

I Problématique

Bien que les inventaires de personnalité soient largement utilisés en recrutement (Bruchon-Schweitzer & Ferrieux, 1991), cette pratique reste freinée par la crainte d'une déformation des réponses (Vom Hofe & Levy-Leboyer, 1993). La DS est associée de longue date à l'évaluation de la personnalité (Tournois, Mesnil et Kop, 2000) et de nombreux tests comportent des échelles évaluant ce phénomène. Elle constitue la critique la plus fréquente adressée aux tests de personnalité utilisés pour la sélection du personnel.
Dans le cadre de la construction d'un inventaire de personnalité de type Big Five, nous avons donc souhaité, d'une part, analyser la validité d'une échelle additionnelle permettant au recruteur d'évaluer la désirabilité sociale des réponses et, d'autre part, étudier la désirabilité sociale propre à chacun des autres items de l'inventaire. Cette démarche participe à l'étude de la validité de construit de notre outil.

II Cadre théorique

II.1. Deux approches du concept de désirabilité sociale

Dans la littérature, deux courants principaux semblent s'intéresser plus particulièrement au concept de désirabilité sociale : la psychologie différentielle et la

psychologie sociale. **Les psychologues différentialistes** définissent la DS comme la *:« tendance à déformer les descriptions de soi dans un sens favorable »* (McCrae & Costa, 1983, p. 884). Pour Tournois et al. (2000), elle renvoie à « *la tendance à produire une image de soi favorable* » (p. 220). Paulhus (1984) identifie deux composantes : l'illusion de soi et la gestion des impressions. L'illusion de soi (ou autoduperie) fait référence à une évaluation optimiste de soi, non consciente, automatique (« *tromper soi-même en toute bonne foi* »). La gestion des impressions (ou hétéro-duperie) correspond, quant à elle, à une stratégie délibérée de réponse pour fournir à autrui une représentation de soi estimée favorable (« *tromper autrui consciemment* »). Pour identifier ces deux dimensions, ces chercheurs utilisent les méthodes corrélationnelles et notamment l'analyse factorielle.

En psychologie sociale, la désirabilité sociale correspond à la tendance à compléter les items d'un questionnaire en se référant aux normes sociales et à la pression normative plutôt que de se présenter tels que l'on est vraiment. Le concept de désirabilité sociale est associé dans ce cadre à celui d'autoprésentation et renvoie à « un ensemble de stratégies plus ou moins intentionnelles, plus ou moins contrôlées qu'un individu met en œuvre dans le but d'obtenir une approbation sociale » (Dubois & Beauvois, 2001, p. 392). Pour mettre en relief ce type de phénomène, les chercheurs ont élaboré un certain nombre de protocoles d'expérimentations dont le plus connu est le *paradigme d'autoprésentation* (Jellison & Green, 1981). Il s'agit de mettre les individus face à trois consignes différentes : une consigne dite « standard » (répondre aux questions posées le plus sincèrement possible), une consigne d'autoprésentation normative (répondre au questionnaire de manière à se faire bien voir, c'est-à-dire donner une image positive de lui-même) et enfin une consigne d'autoprésentation contre-normative (répondre pour donner une mauvaise image de soi-même). Les psychologues sociaux utilisent essentiellement des méthodes basées sur la comparaison de moyennes pour observer ce type de phénomène.

Les psychologues différentialistes et sociaux définissent et étudient la DS de façons distinctes. Cependant, que l'on considère la DS comme une caractéristique intrinsèque et stable de l'individu ou comme le résultat d'une pression sociale, un consensus émerge concernant la déformation des réponses dans des situations d'évaluation comportant un enjeu.

II.2 Mesures de la désirabilité sociale

La DS peut altérer la validité de construit d'un instrument de mesure et les conclusions qui peuvent être faites à partir d'une évaluation, puisqu'elle constitue une source de variance non directement liée au trait qu'on souhaite mesurer (Holden & Fekken, 1989). Quelle que soit l'option théorique dans laquelle on se positionne, il semble donc important de pouvoir quantifier la DS dans le cadre de la construction d'un inventaire de personnalité à visée de recrutement.

La technique la plus utilisée pour évaluer ce phénomène est la construction d'une échelle de désirabilité sociale. Ces échelles sont, pour la plupart, intégrées à l'évaluation. Comme le précisent Tournois et al. (2000), on retrouve ce type d'échelle sous diverses appellations : échelle de mensonge, échelle de bonne impression, besoin d'approbation sociale. Plusieurs outils actuels comportent des

échelles permettant de mesurer ce phénomène : l'EPI (Eysenck & Eysenck, 1975), l'Alter Ego (Caprara, Barbaranelli & Borgogni, 1997) et le MMPI2 (Hathaway & McKinley, 1996). Les interprétations de ces échelles sont multiples et variées. Les auteurs de l'Alter Ego font référence à trois notions différentes pour interpréter le score de DS : *« Gestion de l'impression sur soi », « Conformisme social », et « Effet de l'estime de soi »*. Cette échelle reflète bien le concept de désirabilité sociale dans sa complexité et les difficultés inhérentes à son évaluation mais il est en pratique difficile de faire la part entre ces trois phénomènes à partir d'un score unique. Il en est de même pour le MMPI2, où l'interprétation de l'échelle de DS fait référence à différents éléments : profils de personnalité pathologique, conformisme social et estime de soi.

En définitive, le concept de désirabilité sociale est complexe à comprendre et à définir. Les échelles qui l'évaluent sont donc multiples ce qui rend l'interprétation de ces scores floue et ambiguë. Dans le cadre de notre recherche nous souhaitons être en mesure d'apporter notre contribution à la compréhension de ce phénomène.

III. Méthodologie

III. 1. Description de l'échantillon

71 cadres (dont 70 hommes, âgés de 33 ans en moyenne) ont complété à deux reprises l'inventaire de personnalité[5] comportant une échelle de désirabilité sociale en 12 items. Ils sont vu au cours d'une formation qualifiante à deux reprises.

III. 2. Procédure

Les participants ont complété à deux reprises le test de personnalité, à une semaine d'intervalle, avec deux consignes inspirées du paradigme d'autoprésentation (Jellison et Green, 1981) :

(a) une consigne d'auto-présentation dite « standard » : *Imaginez que vous souhaitez en savoir plus sur votre personnalité et que vous choisissez de répondre à un questionnaire. Ce questionnaire ne comporte aucun enjeu pour vous. Il vous permet simplement d'identifier un certain nombre de traits saillants de votre personnalité. Ce questionnaire a été construit pour permettre à chacun de se situer par rapport à un certain nombre de traits de personnalité. Essayez d'être le plus sincère possible. Si vous hésitez sur une phrase, pensez à ce que diraient vos amis ou vos proches.*

(b) une consigne d'autovalorisation dite normative : *Vous êtes conviés à un entretien d'embauche pour un poste qui correspond tout à fait à vos compétences et aspirations actuelles. Cet entretien représente un enjeu majeur dans votre carrière sous tous points de vue. Les autres postes qui vous ont déjà été proposés n'étaient pas si intéressants que celui-ci. Vous souhaitez vous présenter au mieux au recruteur. N'hésitez pas à insister sur les traits de votre personnalité que vous souhaitez mettre en valeur. Imaginez que vous êtes face à*

[5] L'inventaire de personnalité comprend 252 items dont la structure a été validée sur 736 sujets.

un recruteur pour cet entretien d'embauche et qu'il vous propose de remplir ce questionnaire.

Dans chacune des conditions, les participants devaient présenter leur personnalité sur une échelle graduée en 5 points : de A (tout à fait vrai) à E (tout à fait fausse). Ils étaient séparés en deux groupes : le premier groupe (N= 28) passait la consigne « standard » puis la consigne « autovalorisation» ; l'ordre inverse était proposé à l'autre groupe (N= 43).

III.3. Hypothèses de travail

La première partie de nos hypothèses concerne notre échelle de désirabilité sociale. Nous faisions l'hypothèse que tous les items de celle-ci varieraient d'une consigne à l'autre. Cependant, le caractère inconscient des items d'autoduperie ferait ressortir une variation moindre de ces items par rapport aux autres items de l'échelle.

La seconde partie vise à étudier les autres items de l'inventaire. Nous pensons qu'il était possible d'identifier dans le modèle à 5 facteurs les caractéristiques qui sont les plus affectées par la DS. Conformément à la littérature, nous pouvions nous attendre à des variations sur les items de conscience, névrosisme et d'extraversion (Ones, Viswesvaran & Reiss, 1996).

IV. Résultats

Les résultats généraux indiquent que l'ordre des consignes n'influence pas les réponses des sujets (effet de groupe non-significatif).

IV. 1. Etude de validité des items de l'échelle de désirabilité sociale

Les comparaisons de moyennes (t) réalisées sur les items de l'échelle de désirabilité sociale intégrés dans l'inventaire, nous indiquent que 10 items sur 12 présentent des différences de moyennes significatives ($t(71) < 3,5$; $p < .001$). Les réponses aux items *« Je dis toujours des choses favorables aux autres »* et *« Je suis toujours courtois, même avec les gens désagréables »* ne semblent pas affectés par les conditions expérimentales. De même, les items *« Quand on me critique, je l'accepte toujours »* et *« J'ai parfois l'impression qu'il y a une barrière entre moi et les autres »* présentent des différences de moyennes significatives avec une taille de l'effet modérée ($\eta^2 > 0,5$). Seul un item présente une taille de l'effet forte *(« Je suis parfois inquiet »*). Dans la littérature, ce type d'item est souvent présent dans les échelles de DS. Nos résultats montrent que les items d'autoduperie de l'échelle sont affectés par ce biais au même titre que les items d'hétéroduperie. Ils nous conduisent ainsi à douter du caractère inconscient de cette première dimension.

IV. 2. Etudes des variations des items de personnalité de l'inventaire : aspects généraux

185 items sur 252 présentent des différences de moyennes significatives dans les deux conditions expérimentales soit 73 % d'entre eux. Parmi ces items 102 (40%) présentent une taille de l'effet que l'on peut qualifier de moyenne (supérieure à 0,50) et 21 items avec une taille de l'effet forte (supérieure à 0,80).

Facteurs	Nb d'items (%) avec $50<\eta^2<80$	Nb d'items (%) avec $\eta^2 >80$
Agréabilité	8 (11%)	0 (0%)
Conscience	23 (31%)	5 (25%)
Extraversion.	16 (21,9%)	6 (30%)
Névrosisme	19 (26%)	5 (25%)
Ouverture	7 (9,6%)	4 (20%)
Total	73 (100%)	20 (100%)

Tableau 1: Nombre d'items de chaque facteur ayant un η^2 supérieur à 50 et 80.

Chaque facteurs de personnalité est évalué à partir de 48 items ainsi toutes les échelles du test disposent du même nombre d'items. Le tableau 1 présente le nombre d'affirmations de chaque facteur affectées par la condition expérimentale. Il s'agit, essentiellement, d'items de C*onscience, Névrosisme et Extraversion*. Les items d'O*uverture* et d'A*gréabilité* sont sous-représentés par rapport aux autres dimensions. Seuls, quatre facteurs de personnalité sont affectés fortement ($\eta^2 >80$) par la désirabilité sur cinq (*extraversion, névrosisme, ouverture, conscience*).

IV. 3. Appartenance théorique de ces items

Si l'on regarde avec attention le contenu de chacun de ces items, on remarque que dans chaque facteur, seulement certains éléments semblent désirables. Le **facteur conscience** est affecté par la DS sur des items portant sur l'organisation dans le travail, la fiabilité dans les activités professionnelles confiées, l'impulsivité versus la prudence dans les paroles exprimées et/ou les activités réalisées et enfin l'ambition. Les items les plus sujets à des variations renvoient à l'organisation et à la planification dans le travail ($\eta^2<80$). Dans le **facteur extraversion**, les items présentant les tailles de l'effet les plus fortes font références au caractère dynamisque et au management. Nous pouvons également observer des variations sur des items relatifs à l'aisance en groupe dans le domaine personnel et/ou professionnel et les capacités d'intégration dans un groupe. Dans le **facteur névrosisme**, émergent des items relatifs à la dépression, à la colère ainsi qu'aux comportements impulsifs. Les items d'anxiété et de colère sont les plus sujets aux variations. Dans le **facteur agréabilité**, il s'agit des items se référant à la chaleur et l'empathie (facilité des contacts avec autrui, soutien et attention aux autres). Enfin, dans **le facteur ouverture**, les items évaluant l'attrait pour les rêveries et l'imagination sont les plus touchés par la DS. Nous retrouvons également des items qui concernent l'intérêt pour le changement (voyage, sortir de la routine) et l'attrait pour la culture.

IV. 4. Le statut des items négatifs ou péjoratifs

Il semble que les items négatifs soient plus sujets à la désirabilité sociale que les items positifs. En effet, l'inventaire comprenait, au départ, 45 % d'items négatifs, on en compte 52 % parmi ceux qui présentent des différences significatives avec une taille de l'effet supérieure à .50.

IV. 6. Analyses des items qui ne semblent pas affectés par la désirabilité sociale

Les affirmations qui ne sont pas affectées par les changements d'une consigne à l'autre sont surtout celles qui expriment une idée sur la vie en général et/ou sur des opinions. Il s'agit également d'items se référant à la dimension *agréabilité*.

Discussion et conclusion

La notion de désirabilité sociale joue un rôle central dans le débat, tant au point de vue méthodologique que théorique, autour des tests de personnalité. Les résultats de ce travail nous permettent de mieux comprendre ce concept et son influence sur les différents facteurs de personnalité. Nous retrouvons les résultats de la littérature (Ones, Viswesvaran & Reiss, 1996). En effet, il semble que les facteurs *névrosisme, conscience* et *extraversion* sont les plus touchés par le DS. Cependant, cet effet ne concerne que certains éléments de ces grandes dimensions : la planification et l'organisation pour le facteur *conscience*, l'anxiété et la colère pour le *névrosisme*, le dynamisme et les compétences managériales pour l'*extraversion*, la chaleur et l'empathie pour l'*agréabilité* et enfin, l'ouverture à l'imagination et à la rêverie pour le facteur *ouverture*.

Nos résultats montrent également que la mise en œuvre de ce type de biais n'est pas inconsciente comme le présentait Paulhus (1984). En effet, les items d'autoduperie, insérés dans l'échelle de DS, varient d'une condition à l'autre. Pour construire notre échelle de DS, nous allons conserver les 10 items de désirabilité sociale identifiés comme présentant des variations d'une consigne à l'autre. Pour la compléter, nous ferons un choix parmi les vingt items de personnalité qui présentent une taille de l'effet supérieur à .80. Ce choix se fera en fonction des deux facettes que nous souhaitons mesurer : hétéro-duperie et gestion des impressions sur soi.

Cette étude mérite d'être approfondie. Il nous faut recueillir des protocoles plus diversifiés en terme de sexe, d'age et de catégories socio-professionnelles pour que notre population représente au mieux celle concernée par ce test. Certains travaux montrent, en effet, que le niveau scolaire et culturel des participants peut avoir une incidence sur la mise en œuvre de ce biais. Les personnes les moins cultivés auraient plus de mal à le mettre en oeuvre par méconnaissance des normes et conventions sociales. Nos résultats seront également mis à l'épreuve grâce à l'utilisation conjointe d'une échelle de DS validée dans la littérature : le DS36 (Tournois et al., 2000).

Nous souhaitons également pouvoir réduire le biais de désirabilité sociale en utilisant d'autres stratégies. Une des solutions, considérée comme la plus valide, est de signifier dans la consigne l'existence d'une échelle de mensonge dans le test et d'indiquer qu'elle permet de détecter les tentatives délibérées de déformer des réponses (Hough & Paullin, 1994). Cette technique sera testée dans nos prochains travaux. Ainsi, nous comparerons les réponses des sujets répondant avec une consigne précisant l'existence d'une échelle de mensonge et sans.

Progressivement, nous introduirons une méthode qui permettra de corriger les résultats du test de personnalité en se basant sur cette échelle de désirabilité sociale. Une des solutions commune est, en effet, de s'attacher à étudier les covariances entre traits de personnalité et l'échelle de désirabilité sociale. Traditionnellement on

pense que ces covariances sont des indicateurs de biais sur les échelles (Ellingson & Sackett, 1999).

Bibliographie :

Bruchon-Schweitzer, M., & Ferrieux, D. (1991). Une enquête sur le recrutement en France. *Revue Européenne de Psychologie Appliquée*, 41, 1, 9-17.

Caprara, G.V., Barbaranelli, C., & Borgogni, L. (1997). *Alter Ego : les cinq facteurs fondamentaux de la personnalité*. Paris : Les Editions et Applications Psychologiques.

Dubois, N., & Beauvois, J-L (2001). Désirabilité et utilité : deux composantes de la valeur des personnes dans l'évaluation sociale. *L'orientation scolaire et professionnelle*, 30, 3, 391-405.

Hathaway, S.R., & Mc Kinley, J.C. (1996). *Manuel de l'inventaire Multiphasique de personnalité du Minnesota – 2.* Paris : Les éditions du centre de psychologie appliquée.

Holden. R.R., & Fekken, G.C. (1989). Three common social desirability scales: Friends, acquaintances, or strangers. *Journal of Research in Personality*, 23, 180-191.

Hough, L., & Paullin, C. (1994). Construct-oriented scale construction. In G.S. Stokes, M. D. Mumford & W.A. Owens (Eds.), Biodata handbook (pp. 127-131). Palo Alto, CA : CPP *Books*.

Ellingson, J.E., & Sackett, P. R. (1999). Social desirability corrections in personality measurement: Issue of applicant comparison and construct validity. *Journal of Applied Psychology,* 84, 2, 155-166.

Eysenck, H.J., & Eysenck, S.B.G. (1975).*Manual of the Eysenck Personnality Questionnaire.* Sevenoaks: Hodder and Stoughton.

Jellison, J.M., & Green, J. (1981). A self-presentation approach to the fundamental attribution error : the norm of internality. *Journal of Personality and Social Psychology,* 40, 643-649.

McCrae, R.R., & Costa, P.T.Jr. (1983). Social desirability scales : More substances than style. *Journal of consulting and Clinical Psychology*, 51, 6, 882-888.

Ones, D., Viswesvaran, C., Reiss, A.D. (1996). Role of social desirability in personality testing for personnel selection: The Red Herring. *Journal of Applied Psychology,* 81, 6, 660-679.

Paulhus, D. (1984). Two-component models of social desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.

Tournois, J., Mesnil, F., & Kop, J-L. (2000). Autoduperie et hétéroduperie : un instrument de mesure de la désirabilité sociale. *Revue européenne de psychologie appliquée, 50* (1), 219-232.

Vom Hofe, A., & Levy-Leboyer, C. (1993). Evaluation of the use of personality tests in personnel selection in France. *Revue européenne de Psychologie Appliquée*, 43, 3, 221-227.