

**RESUME DES RESULTATS DES RECHERCHES EFFECTUEES A L'UNIVERSITE DE
LAUSANNE SOUS LA DIRECTION DE ROLAND CAPEL, AVEC LA COLLABORATION DE
SES ASSISTANTES ET ASSISTANTS ET LA PARTICIPATION DE SES ETUDIANT(E)S
Concernant
LE PROBLEME DE LA FALSIFICATION DES QUESTIONNAIRES A EVALUATION
SUBJECTIVE UTILISES EN SITUATION AVEC ENJEU**

**Francis GENDRE
Rapport technique (2017)**

Synthèse de la littérature

La Falsification a d'abord été conçue comme en relation avec la Désirabilité Sociale une caractéristique objective et unidimensionnelle des items (comme la popularité par exemple) pour devenir une caractéristique subjective et multidimensionnelle des réponses d'un sujets aux items proposés, ce qui rend le problème de sa détection et de sa correction éventuelle beaucoup plus complexe.

Les questionnaires de personnalité et d'attitudes ont longtemps été négligés dans les procédures de sélection car considérés comme n'ayant qu'une validité négligeable voire nulle et être sujets à de nombreux biais de réponse les rendant inutilisables.

Le développement des techniques de méta analyses et la simplification du domaine de la personnalité apporté par le modèle généralement accepté des 5 Grands Facteurs (et de leurs facettes) a permis depuis les années 90 d'avoir une meilleure estimation de la valeur potentielle des dimensions de personnalité, dont la validité directe et incrémentale est maintenant considérée comme non négligeable et fait de l'évaluation de la personnalité une procédure incontournable dans les procédures de sélection à coté de l'évaluation des aptitudes (Intelligence générale) et de l'interview d'embauche (structuré), surtout quant elle est associée à une analyse des exigences du poste mis au concours en terme des traits de personnalité requis (approche confirmatoire).

La validité des dimensions de personnalité est aussi augmentée en tenant compte des effets modérateurs et médiateurs.

Dans ce contexte la falsification, bien que toujours possible, est considérée comme n'ayant pas un effet suffisant pour neutraliser la validité des dimensions de la personnalité, même si elle l'amointri et rend la décision d'embauche moins précise et moins efficace.

Les solutions proposées pour prévenir ou contrer la Falsification se résument à l'avertissement sur les conséquences d'une détection dont l'efficacité n'est pas nulle mais ne résout pas véritablement le problème et l'utilisation de formes ipsatives censées protéger des biais de tendance centrale, acquiescence et désirabilité. Ces formes ont été critiquées d'un point de vue mathématiques comme ne pouvant pas permettre les comparaisons interpersonnelles que seuls des scores normatifs absolus autorisent.

En fait les formes normatives habituelles (réponses en échelles de Likert) ne sont pas plus absolues que celles ipsatives et il y a peu de différences dans les comportements de ces deux approches. Le vrai problème est que les formes ipsatives en éliminant des items la désirabilité, élimine des scores la désirabilité Vraie des sujets qui est l'information la plus importante par rapport à la décision d'embauche !

Certains voient dans l'utilisation de la théorie de réponse à l'item une voie prometteuse mais qui n'a pas encore fait ses preuves.

La solution au problème de la falsification nécessite de mieux comprendre le processus de falsification au niveau des réponses à chaque item et en tenant compte de la perception qu'a le sujet de ses capacités à falsifier et des caractéristiques idéales du poste à pourvoir de manière à modéliser ce processus avec une précision suffisante pour pouvoir envisager des solutions opérationnelles.

Recherches effectuées à LAUSANNE

En ce qui me concerne voilà comment je considère les choses: Dans la réalité il y a 2 types de situations les situations sans enjeu dont le type est la situation de conseil ou c'est le sujet qui est demandeur et qui a donc tout intérêt à être sincère pour pouvoir obtenir l'aide qu'il sollicite, pour qu'elle soit la plus adéquate... Dans cette situation on ne se pose pas le problème de la falsification ou non.....et il y a les situations avec enjeu dont le type est la situation de sélection ou c'est un commanditaire (le recruteur) qui est le demandeur et le sujet qui répond a tout intérêt à paraître proche de ce que le commanditaire recherche s'il veut avoir les meilleures chances d'être embauché. C'est là que se pose le problème de la falsification et de la désirabilité locale, c'est à dire à la fois ce que le commanditaire désire (la Désirabilité locale proprement dite) et la Désirabilité locale perçue ce que le sujet qui répond pense être la Désirabilité locale. Le problème est que s'il est possible de connaître la Désirabilité locale en demandant au recruteur de remplir le test en ce sens, il n'est pas possible de connaître la Désirabilité locale Perçue, qui peut différer d'un sujet à l'autre même s'il y a sans doute des similitudes et que l'on peut plus ou moins déterminer d'après le profil moyen des candidats, la Désirabilité Locale Perçue Moyenne, sinon il faut demander aux candidats de définir leur Désirabilité Perçue ce qui semble difficile (mais cependant pas impossible).

Ceci étant dit nous n'avons pas accès à un terrain adéquat, à la réalité d'une situation de sélection pour pouvoir étudier, modéliser et valider un processus de détection et de correction éventuelle des données individuelles.

C'est pourquoi nous devons nous contenter de notre terrain à savoir des étudiant(e)s en psychologie qui ne sont guère représentatif des candidats à un poste disons de cadre dans l'industrie et qui ne sont pas demandeurs mais en quelque sorte obligé de participer à une expérience de "simulation", en répondant à un questionnaire d'abord en situation dite Sincère et ensuite en situation dite Biaisée. Il y a donc manifestement un problème à la fois de motivation et de connaissance du monde du travail qui fragilise les résultats que l'on peut trouver et qui peut mettre en doute la Sincérité des protocoles en situation dite sincère, car il risque toujours d'avoir des individus farfelus qui répondent un peu n'importe comment. Il en est de même pour la situation dite Biaisée ou ce même phénomène peut aussi se produire.

Cependant nous devons faire avec ce que nous avons et c'est ce que nous avons fait, ainsi que d'ailleurs d'autres chercheurs travaillant en milieu universitaire.

Quoiqu'il en soit nous avons mené plusieurs campagnes de testage selon la distinction "situation sincère" vs "situation biaisée" avec d'abord une version "courte" de Socrate, puis la version normale et avec des consignes et/ou des conditions de passations quelques peu différentes (comme de répondre en même temps à la situation sincère et à la situation biaisée), ce qui introduit une difficulté supplémentaire dans la synthèse des résultats.

Cependant dans ces analyses nous avons été guidée par le modèle de la falsification

spécifique à la méthode fonctionnelle que j'ai élaboré et qui est $O = V + c (DS - V)$. En nous basant sur la littérature existante nous avons d'abord retenu la notion d'une Désirabilité Sociale Générale comme source quasi Universelle de Falsification, et nous avons défini cette Désirabilité comme la première composante principale des items de Socrate, par analogie avec la première composante de tests d'aptitudes qui définit le facteur Général. C'est aussi ce que propose plusieurs auteurs plutôt que d'avoir recours à des juges pour définir la Désirabilité Sociale Générale.

Et nous avons effectué sur les autres composantes retenues des rotations Varimax pour obtenir les autres dimensions fondamentales de Socrate, et qui par constructions sont indépendantes de la Désirabilité. L'idée étant que ces dimensions seraient peu ou pas susceptibles d'être falsifiées et que la quasi totalité de la falsification se concentrerait sur la dimension DS. Si cela était le cas le modèle de la falsification permettrait de pouvoir estimer la falsification et de la corriger et ceci en relation avec une des solutions proposées à l'époque par de nombreux chercheurs et censée contrecarrer la falsification : les formes ipsatives. Elles consistent à présenter les items sous forme à choix forcé de groupes de 3 ou 4 items et de choisir celui qui vous convient le mieux et le moins bien, ces groupes d'items étant appareillés du point de vue de la Désirabilité Sociale. Le raisonnement étant le suivant : Les items présentés étant de même Désirabilité le sujet ne peut plus falsifier en fonction de la Désirabilité, et donnera alors une réponse "sincère".

Ce qui semble logique mais oublie une chose essentielle c'est que de ce fait la Désirabilité disparaît de l'instrument, aussi bien la Désirabilité falsifiée que celle Vraie : les qualités désirables que possède réellement le sujet et c'est précisément cette Désirabilité que recherche le Recruteur. Les formes ipsatives sont donc des formes amputées de l'essentiel, de ce que recherche le recruteur!

Cependant l'utilisation conjointe d'une forme habituelle et d'une forme ipsative dans le cadre de la méthode fonctionnelle qui isole la dimension de Désirabilité pouvait permettre théoriquement de résoudre le problème et même d'estimer la Désirabilité Vraie du sujet.

Nous avons donc tout d'abord créé 2 formes parallèles courtes de Socrate, une présentée de façon habituelle à réponses en échelle de Lickert et une ipsative, et ces deux formes ont été passées en situation dite sincère et en situation dite Biaisée (les formes étant inversées celle en échelle de Lickert devenant ipsative et celle ipsative devenant en échelle de Lickert).

. 103 étudiant(e)s ont passé deux fois Socrate Sélection, la forme A ou la forme B, la première fois avec des instructions de « sincérité » et la seconde (en inversant les formes) avec des instructions de « sélection » pour un poste impliquant affirmation de soi, dynamisme, efficacité (Désirabilité Générale vs. réservé, tolérant, aimable).

Pour notre expérimentation, nous avons calculé les éléments suivants : la moyenne et le sigma des réponses en échelle de Likert Vrai et Falsifié, les stratégies de réponses Vraie et Falsifiées, pour les réponses en échelle de Likert et pour celles à choix forcé, et les cohérences (corrélations multiples) qui leur sont associées. De même que les convergences (produit des stratégies normées) entre les différentes stratégies.

Résultats

Session « sincère »

2.1.1. Corrélations

Nous avons remarqué deux choses, en relation avec la théorie de la falsification : la première est : les corrélations sont élevées, (compte tenu de la brièveté des formes,) entre les éléments correspondants de la stratégie en échelle de Likert et ceux de la stratégie à choix forcé, .70 et .71 pour les dimensions Modération et Indépendance, et .48 et .47 pour celles Amabilité et Conventionnel. Ceci confirme l'hypothèse que la méthode à choix forcé, restitue assez bien la stratégie Vraie, indépendante de la Désirabilité, **tout au moins en situation « sincère »**.

La seconde est la faible corrélation (.32) entre la Cohérence Tronquée, en échelle de Likert et celle de la Cohérence à choix forcé. En théorie, elles devraient être identiques.

Moyennes

Les moyennes des éléments correspondants des stratégies Indépendantes de la Désirabilité en échelle de Likert et à choix forcé : Nous avons remarqué la différence importante (.48 et .38) entre les cohérences, au bénéfice des échelles de Likert, alors qu'en théorie elles sont censées être identiques !

Ceci s'explique certainement par le fait que la forme à choix forcé est moins fidèle que celle en échelle de Likert, ce qui n'est guère mentionné dans la littérature sur le sujet. En effet le mode à choix forcé est constitué de 18 trios, dont 2 réponses peuvent être considérées comme indépendantes, soit 36 réponses indépendantes, alors que le mode en échelle de Likert les 54 réponses sont indépendants. La diminution attendue de la Cohérence pour une telle diminution du nombre d'items indépendants est exactement de .10, celle constatée.

Session « biaisée »

Corrélations

Les corrélations entre les éléments correspondants des stratégies en échelle de Likert et à choix forcé, sont relativement élevées, .55 et .66 pour Modération et Amabilité et .44 et .24 pour Conventionnel, mais cependant plus faibles que pour la session « sincère », ce qui va dans le sens attendu. Nous avons constaté également la faible corrélation entre la Cohérence Tronquée en échelle de Likert (.35) et celle de la Cohérence à choix forcé pour les mêmes raisons que précédemment.

. Moyennes

Les moyennes des éléments correspondants des stratégies Indépendantes de la Désirabilité en échelle de Likert et à choix forcé ont mis en évidence que la différence entre les Cohérences Tronquées, est plus faible (.07) que pour la session « sincère » ($r = .34$, contre .44 précédemment), alors qu'en théorie elle devrait être plus forte ! Ceci nous a démontré que nos sujets ne se sont pas comportés comme le prévoit la théorie

Inter sessions

Corrélations

Les corrélations intersessions entre les éléments des stratégies en échelle Likert et à choix forcé, sont rarement significatives et alors de valeurs modestes. En particulier les corrélations entre stratégies à choix forcé intersessions sont pratiquement inexistantes, contrairement à la théorie qui fait l'hypothèse que la stratégie à choix forcé observé en situation de « sélection » reproduit celle observée en situation « sincère ».

Moyennes

Les différences de moyennes les plus fortes concernent la Désirabilité (-.82), la Moyenne (-.70) et la Cohérence (-.51), en accord avec ce qui était prévisible.. La

désirabilité permet donc d'identifier clairement la situation de sélection ! Les autres différences sont de peu importance, même si elles montrent qu'en situation de sélection les sujets se décrivent comme plus aimables, plus conventionnels et plus dirigeant qu'en situation sincère. En ce qui concerne les réponses à choix forcé, il y a peu de différences et elles sont modestes : les sujets se décrivent plus agressifs et plus créatifs en situation de sélection qu'en situation « sincère ». **Ce peu de différence va dans le sens escompté, mais ceci est trompeur comme l'a montré les analyses de corrélations.** En fait les moyennes ont peu changées mais leur écart type c'est réduit de moitié en passant de la situation sincère à celle biaisée, indiquant que ceux qui avaient des scores élevés les ont baissés et ceux ayant des scores bas les ont élevés conformément au modèle de la falsification propre à la méthode fonctionnelle.

Conclusion

L'analyse des moyennes et des inter-corrélations des éléments des stratégies associées aux réponses en échelle de Likert et à choix forcé obtenues sous deux instructions : sincère ou frelaté, a confirmé un certain nombre d'hypothèses associées à la théorie de la falsification.

Les réponses en échelle de Likert et à choix forcé mesurent bien des dimensions analogues, indépendantes de la désirabilité, mais les choix forcés le font avec une précision (fidélité) moindre due à la dépendance d'une des trois réponses proposées. La situation « frelaté » induit une forte augmentation de la dimension Désirabilité, en accord avec la théorie, ainsi que de la moyenne des réponses et de leur Cohérence, en ce qui concerne des échelles à choix forcé elles ne présentent pratiquement aucune relations entre les sessions, ce qui est contraire à la théorie de la falsification dite classique et amène à faire l'hypothèse que la désirabilité locale perçue par les sujets est différente de celle proposée.

Nouvelles analyses

Suite à cette première expérimentation nous avons effectué de nouvelles prises de données avec des formes complètes et non plus raccourcies, aussi bien pour la forme en échelle de Likert que pour la forme à choix forcés, et essayé une forme dite multi ipsative.

L'analyse des moyennes et écart types en situation sincère et biaisé a montré pour la forme Likert une forte augmentation de la dimension de Désirabilité, et beaucoup plus faiblement des dimensions Conventionnelle, Agressivité et Action, et par une diminution de l'écart type de toutes les dimensions, indice d'une perte d'information.

La corrélation avec la convergence est positive pour la Cohérence en situation sincère et négative en situation biaisée, indiquant pour cette dernière que la falsification à tendance à faire augmenter la Cohérence c'est à dire à produire une image plus stéréotypée.

Les corrélations avec les éléments de la stratégie, montrent une forte corrélation positive avec la dimension Désirabilité en situation sincère et une forte corrélation négative en situation biaisée. Ceci indique d'une part que ceux qui en situation sincère se décrivent comme possédant des qualités désirables ne falsifient pas ou peu et que par

contre ceux qui en situation biaisé se présentent comme possédant des qualités désirables sont ceux qui ont le plus falsifié ! **La falsification est donc directement liée à l'incitation à la falsification c'est à dire l'écart perçu entre ce que l'on est et ce qui est désirable de paraître.**

Cependant la Désirabilité n'explique pas tout et il existe une Désirabilité résiduelle faite de Conservatisme, d'Agressivité et d'Action qui brouille quelque peu le problème.

Les relations entre les éléments de la stratégie normée sincère et biaisé montre que la falsification invalide la dimension de Désirabilité, amoindri fortement celles de Conservatisme, Amabilité et Action et un peu moins celle d'Indépendance.

La discrimination sincères /falsificateurs (y compris imposteurs) de même que les corrélations entre éléments de la stratégie et Convergence en situation biaisée met en évidence les mêmes dimensions : la Désirabilité en tout premier lieu, et de façon moindre la Désirabilité résiduelle.

Ces résultats confirment ceux que nous avons trouvé dans les expérimentations antérieures, qui utilisaient les réponses Likert et Ipsatives, et qui avait mis en évidence que la Désirabilité n'était pas le seul facteur de falsification et qu'il existait une Désirabilité résiduelle qui invalidait la forme ipsative contrairement aux théories en vigueur en plus d'éliminer la dimension de Désirabilité.

Nous avons conçu la méthode Multi-ipsative pour pallier à ce second problème en montrant qu'en situation sincère réponse Likert et Multi ipsatives pouvaient être considérées comme identique (sans élimination de la dimension de Désirabilité). Nous avons l'espoir que la complexité de la forme multi-ipsative rendrait la falsification plus difficile voire au mieux impossible ! Nous sommes obligés de constater qu'il n'en est rien et que les sujets sont capables de falsifier la forme multi-ipsative d'une manière analogue à la falsification de la forme Likert en situation biaisée.

Effet de la falsification sur la validité de Socrate

Résumé

La Falsification peut être définies comme la décision de modifier ce qui serait une réponse vraie, sincère en une réponse différente dite altérée et ceci en fonction de la perception qu'a le sujet de la désirabilité propre à la situation, dans le but d'en obtenir un avantage reel ou symbolique (être recruté ou être apprécié).

Il existe autant de désirabilité que de situation mais en pratique seule la Désirabilité Sociale et Professionnelle Générale a été considérée

De manière classique il est habituel d'essayer d'éliminer ce biais soit en éliminant la Désirabilité par régression, soit en présentant au sujet un choix entre items appareillés du point de vue de la désirabilité sociale et professionnelle générale. Ceci élimine certes la désirabilité mais aussi des qualités réelles mais désirables que pourraient posséder le sujet.

La méthode fonctionnelle propose une technique originale et sophistiquée qui lui est propre, qui permet d'isoler la désirabilité sociale et professionnelle et de créer un espace de mesure indépendant de celle-ci, sans avoir à modifier ni la présentation naturelle des items ni la façon d'y répondre. Le principe en est le suivant : dans le domaine des aptitudes la première composante principale correspond au **facteur G** (ou intelligence générale), une dimension générale, positive, qui transcende les facteurs spécifiques. Dans le domaine des attitudes il en est de même et dans le cas des compétence il s'agit d'une dimension qui opérationnalise ce qui peut être appelé **la Désirabilité Sociale et Professionnelle Générale.**

Socrate est destiné à être utilisé en situation de sélection comme une alternative aux solutions classiques consistant à éliminer la Désirabilité soit par régression soit au niveau des items en les appareillant. L'idée est que si un candidat falsifie ses réponses en fonction de la Désirabilité, ses résultats aux autres dimensions indépendantes de la désirabilité seront altérés et diminués et qu'en ne tenant compte que des résultats indépendants de la Désirabilité, les falsificateurs seront défavorisés par rapport à ceux qui répondent sincèrement, et la justice ainsi rétablie.

Les résultats de Socrate comporte des indices de contrôle, des échelles fondamentales, des échelles fonctionnelles et des échelles externes.

Une expérimentation a donc été menée en proposant à 115 étudiants en psychologie de répondre de manière sincère à SOCRATE, puis en s'imaginant en situation de sélection. Ces mêmes étudiant ont d'autre part répondu de manière sincère au NEO-PI.

L'analyse de ces données a montré qu'en situation simulée de sélection, nos sujets ont modifiés leurs réponses de manière à obtenir un score élevé à la dimension de Désirabilité proposée.

En ce qui concerne les indices de contrôle, la falsification a provoqué une forte augmentation du Niveau de réponse, une diminution importante de l'écart type de tous les indices et scores aux échelles ainsi qu'une augmentation de la Cohérence Totale et de la Fiabilité.

Ceci indique que la falsification a diminué l'information véhiculée, qui est devenue, à la fois plus stéréotypée et plus positive.

En ce qui concerne les échelles fondamentales, comme prévu, le score de la dimension Désirabilité a fortement augmenté, mais les scores des autres dimensions indépendantes de la Désirabilité ont aussi augmentés pour 3 des 4 dimensions impliquées, assez fortement pour Amabilité et plus modestement pour Modération et Travailleur, Indépendant restant inchangé.

Ceci indique que la Désirabilité « agissante », ou Désirabilité Perçue, bien que fortement corrélée à la Désirabilité proposée, a une partie indépendante de cette Désirabilité (ou Désirabilité Perçue résiduelle) qui s'exprime dans certaines des échelles fondamentales indépendantes de la Désirabilité.

L'effet conjoint de la Désirabilité Proposée et de la Désirabilité Résiduelle Perçue explique les augmentations et les diminutions des scores des échelles fonctionnelles et externes qui ne sont que des illustrations des différentes combinaisons possibles des échelles fondamentales.

Ainsi pour les échelles fonctionnelles, la falsification augmente 6 de leurs scores, à savoir : combativité, esprit d'initiative, organisation, efficacité, pragmatisme et sens du management et diminue 5 de leurs scores, à savoir : esprit de synthèse, indépendance, autorité, persuasion et besoin d'intégration à une équipe.

Pour les échelles externes les scores de Stabilité (Névrosisme -), Extraversion et Conscience augmentent, et ceux d'Ouverture et d'Amabilité diminuent.

Ces résultats montrent que le dispositif de SOCRATE qui devait aboutir à ce que tous les scores des échelles fondamentales indépendantes de la Désirabilité diminuent sous l'effet de la falsification, ce qui devait rendre la falsification contre productive et qu'il devait en être de même pour les échelles fonctionnelles et externes calculées sans DS, ne fonctionne que partiellement du fait des différences existant entre la Désirabilité Proposée et la Désirabilité Perçue. L'intérêt de la Falsification en reste cependant amoindrie et celui de la sincérité augmenté,

Un autre problème subsiste : quel est l'effet de la falsification par rapport à la validité de SOCRATE ? L'analyse des relations entre SOCRATE Sincère et SOCRATE Biaisé et le NEO-PI Sincère nous ont apporté des informations à ce sujet, qui ont montrées que la falsification diminuait et altérait la validité conceptuelle de SOCRATE.

Ainsi pour les indices de contrôle, la Cohérence purifiée (sans DS) est associée en situation Sincère, à la Stabilité et à l'Amabilité alors qu'en situation Biaisée, si elle reste associée à l'Amabilité, par contre, elle est associée négativement (Névrosisme) à la Stabilité.

Le Niveau, aussi bien en situation Sincère que Biaisée, est associé à toutes les dimensions du NEO-PI mais plus faiblement en situation Biaisée.

En ce qui concerne les échelles fondamentales, la falsification a détruit toutes les relations avec la dimension Conscience, diminué celles avec les dimensions d'Extraversion, d'amabilité et de Névrosisme et laissé intactes celles avec la dimension Ouverture.

En ce qui concerne les échelles fonctionnelles, que ce soit en situation Sincère ou Biaisée, avec ou sans DS, les relations avec la dimension Ouverture, cohérentes et élevées, restent intactes tandis que celles avec la dimension Conscience, nombreuses, cohérentes et élevées en situation Sincère (avec ou sans DS) disparaissent totalement en situation Biaisée. Les relations avec les dimensions Extraversion et Amabilité, nombreuses, cohérentes et élevées en situation Sincère (avec ou sans DS), diminuent en situation Biaisée et même changent de nature pour Extraversion. Les relations avec la dimension Névrosisme sont plus nombreuses et plus élevées en situation Biaisée qu'en situation Sincère.

En ce qui concerne les échelles externes, avec ou sans DS, en situation Sincères les relations observées sont nombreuses, cohérentes et élevées, tandis qu'en situation Biaisée elles sont plus rares, plus basses et plus problématiques.

En bref, la falsification même si elle n'invalide pas complètement les résultats, diminue fortement l'information valide véhiculée et la rend moins précise, à l'exception de celle relative à la dimension Ouverture qui semble ici être indépendante à la fois de la Désirabilité proposée et de la Désirabilité perçue.

Finalement l'analyse des différences entre les scores Biaisés et ceux Sincère (ou falsification) a montré que l'ampleur de la falsification était fonction de l'Incitation à la falsification c'est à dire de l'écart existant entre l'image de soi sincère et le perception de celle désirable. Il y a donc autant de type de falsification qu'il y a de possibilité d'avoir une ou plusieurs déficiences par rapport à la Désirabilité perçue.

En conclusion, la falsification est un problème majeur dans l'utilisation des épreuves d'évaluation subjectives car elle rend la procédure de sélection à la fois injuste, en favorisant la duplicité au dépend de la sincérité, et de surcroit inefficace en détruisant une grande partie de la validité réelle des instruments utilisés.

CONCLUSIONS

Les recherches que nous avons menées ont montrées que le problème de la falsification tel qu'on peut l'étudier à partir de simulations effectuées avec des étudiants en psychologie était un problème complexe car il déborde largement de ce qu'on appelle habituellement la Désirabilité Sociale. Les procédures préconisées pour remédier à ce problème à savoir les test à choix multiples appareillent les items en fonction de la Désirabilité Générale sont totalement inefficace d'une par parce qu'ils éliminent la Désirabilité Vrai aussi bien que celle Falsifiée et aussi parce qu'ils sont malgré tout falsifiables en fonction de ce que nous avons appelé la Désirabilité Locale Résiduelle et aussi beaucoup moins fidèles que les tests en réponse de Likert.

Nos études ont montré l'importance de l'Incitation à la falsification c'est à dire l'écart existant entre le profil sincère et la Désirabilité perçue. Plus cet écart est grand et plus la tendance à falsifier est élevée.....et aussi le fait que la Désirabilité Perçue intervient en plus de la Désirabilité Proposée. En fait il existe autant de Désirabilités Perçues que de sujets et de situations même si l'influence de la Désirabilité Générale est la plus importante.

La situation simulée avec enjeu induit donc des réponses qui augmentent fortement la moyenne de la dimension Générale (DS) mais aussi de façon moindre pratiquement toutes les autres dimensions et surtout elle diminue de moitié l'écart type de toutes les dimensions même celles qui paraissent n'être que peu affectées par la falsification si on ne s'en teneit qu'aux moyennes. Cela indique une perte importante d'informations pour toutes les variables mêmes celles qui semblaient peu touchées. Cela indique aussi que le modèle de la falsification que nous avons élaboré s'applique aussi aux dimensions de Désirabilité moyenne, la falsification ayant alors tendance à rendre ces scores moyens pour tout le monde.

Ces altérations que la falsification apportent aux réponses ne sont évidemment pas anodines par rapport à la validité des résultats obtenus en situation avec enjeu (même simulé). Que ce soit par rapport à la validité des dimensions mesurées en situation biaisée et en situation sincère ou par rapport à un test externe comme le NEO PI passé en situation sincères, la validité des résultats biaisés est considérablement affectée pour être quasiment nulle pour la dimension DS de Socrate ou celle Conscience du NEO PI , mais toutes les autres dimensions sont également affectées de façon plus ou moins marquée. Seule la dimension Ouverture semble épargnée quelque peu, ce qui peut être que circonstanciel et fonction de la Désirabilité Perçue spécifique aux conditions de notre expérimentation.

Nous continuons nos recherches dans le but de déterminer une procédure fiable permettant d'identifier les protocoles sincères et ceux falsifiés et parmi ceux falsifiés le niveau de la falsification.

Nouvelles recherches effectuées à Lausanne

Francis GENDRE

Rapport technique (2017)

1. Nouvelles recherches effectuées à LAUSANNE

2. En nous basant sur la littérature existante nous avons d'abord retenu la notion d'une Désirabilité Sociale Générale comme source quasi Universelle de Falsification, et nous avons défini cette Désirabilité comme la première composante principale des items de Socrate, par analogie avec la première composante de tests d'aptitudes qui définit le facteur Général. C'est aussi ce que propose plusieurs auteurs plutôt que d'avoir recours à des juges pour définir la Désirabilité Sociale Générale.
3. Et nous avons effectué sur les autres composantes retenues des rotations Varimax pour obtenir les autres dimensions fondamentales de Socrate, et qui par constructions sont indépendantes de la Désirabilité. L'idée étant que ces dimensions seraient peu ou pas susceptibles d'être falsifiées et que la quasi totalité de la falsification se concentrerait sur la dimension DS. Si cela était le cas le modèle de la falsification permettrait de pouvoir estimer la falsification et de la corriger et ceci en relation avec une des solutions proposées à l'époque par de nombreux chercheurs et censée contrecarrer la falsification : les formes ipsatives.
4. Elles consistent à présenter les items sous forme à choix forcé de groupes de 3

ou 4 items et de choisir celui qui vous convient le mieux et le moins bien, ces groupes d'items étant appareillés du point de vue de la Désirabilité Sociale. Le raisonnement étant le suivant : Les items présentés étant de même Désirabilité le sujet ne peut plus falsifier en fonction de la Désirabilité, et donnera alors une réponse "sincère".

5. Ce qui semble logique mais oublie une chose essentielle c'est que de ce fait la Désirabilité disparaît de l'instrument, aussi bien la Désirabilité falsifiée que celle Vraie : les qualités désirables que possède réellement le sujet et c'est précisément cette Désirabilité que recherche le Recruteur. Les formes ipsatives sont donc des formes amputées de l'essentiel, de ce que recherche le recruteur!

RESULTATS DES DERNIERES ANALYSES EFFECTUEES

1. Introduction.

La Falsification, c'est l'action consistant à modifier une réponse Sincère qui pourrait être donnée dans une situation sans enjeu et où le répondant est aussi le demandeur, pour donner une réponse Altérée qui tend à se rapprocher de ce que le répondant considère comme une réponse Désirable, dans une situation avec enjeu et où le demandeur est aussi celui qui peut accorder le gain de l'enjeu.

L'analyse de la Falsification implique d'étudier :

La motivation à Falsifier : elle dépend essentiellement de l'importance de l'enjeu pour celui qui est candidat, si c'est vital pour lui ou simplement un plus par rapport à sa situation actuelle.

L'incitation à Falsifier : elle dépend essentiellement de l'écart qui existe entre la réponse Sincère et la Réponse Désirable. Plus cet écart est grand et plus l'incitation à Falsifier sera grande.

L'importance de la Falsification : elle dépend essentiellement de 3 facteurs : la motivation à falsifier, l'incitation à falsifier et le culot du répondant. Elle peut aller de 0 à 100% du comblement de l'écart entre réponse Sincère et réponse Désirable

Le culot du répondant : c'est une caractéristique de la personnalité du répondant qui dépend de ses valeurs morales, de sa propension à prendre des risques et de sa conviction d'être capable de falsifier sans être soupçonné de l'avoir fait, mais aussi de l'importance de l'enjeu. Il se mesure par l'importance de la falsification, le % de l'écart entre la réponse Sincère et la réponse Désirée, comblé dans la réponse Altérée.

2. Nouvelles Analyses.

Nous avons utilisé 3 Cohortes d'étudiant soit un total de 461 sujets ayant participé à une expérience de simulation en répondant d'abord sous des consignes de Sincérité et une seconde fois avec des consignes de réponse en situation avec enjeu.

2.1. Description de la Falsification

Les indices de Contrôle

La falsification a créé des réponses plus Cohérentes et plus Fidèles c'est à dire plus stéréotypée et un élargissement du domaine considéré comme attractif.

Les échelles Fondamentales

En ce qui concerne les échelles fondamentales, la falsification a produit une forte augmentation de la dimension de Désirabilité Générale, une augmentation moyenne de la dimension Amabilité, une augmentation faible de la dimension Conservation et une très faible diminution de la dimension Activité et nulle de celle Individualisme.

Les échelles Externes (Big Fives)

En ce qui concerne les échelles externes la falsification a produit une forte augmentation de l'échelle de Stabilité et de celle de Conscience, une augmentation moyenne de l'échelle d'Extraversion une diminution moyenne de celle d'Ouverture et pratiquement nulle de celle d'Amabilité .

Conclusions concernant les différences de moyennes

En situation « Sincère », notre échantillon se caractérise par l'image de personnes ouvertes à autrui, courtoises et aimables, plutôt individualistes et indépendantes, empathiques et altruistes , plutôt stables et plutôt extraverties.

En situations « Biaisée » l'image a changée et devient celle de personnes très ouvertes à autrui , très stables , entreprenantes et assurées , consciencieuses et travailleuses, extraverties et aimables et quelque peu conservatrice et assez indépendantes.

En bref on est passé de l'image d'étudiants sociables et normaux à l'image du cadre idéal, trop parfait pour être vrai !

On peut en conclure que cette dernière image correspond à ce que nous appelons la Désirabilité Locale Perçue et que bien que la désirabilité locale perçue soit par nature individuelle cette Désirabilité Locale Perçue Moyenne semble être partagée par une majorité de nos sujets

Les écarts types.

On pourrait croire que les dimensions et échelles dont la moyenne n'ont pas ou peu changées lorsque l'on passe de la situation Sincère à la situation Biaisée ne sont pas affectées par la falsification, mais en fait il n'en est rien, car la falsification agit aussi sur la variance des dimensions et échelles, c'est à dire sur la quantité d'information véhiculées par ces dimensions et échelles.

En passant d'une situation à l'autre les écarts types des dimensions et échelles sont réduits de moitié. L'information véhiculée (la variance) par toutes les dimensions et échelles a été réduite d'un quart ce qui se traduit nécessairement par une réduction notable de la validité des informations recueillies en situation biaisée par rapport à celle recueillies en situation sincères.

Les corrélations.

Nous avons calculé les corrélations entre les échelles fondamentales et les échelles externes (échelles OCEAN du modèle des BIG FIVE) en séparant les données dites Sincères de celles dites Biaisées

Les dimensions fondamentales des caractéristiques des items sont par construction orthogonales mais ceci n'implique pas nécessairement que les échelles fondamentales et encore moins que les échelles externes, le soient.

En ce qui concerne les échelles fondamentales, en situation Sincère, elles sont pratiquement orthogonales.

En situation Biaisée, l'on constate des corrélations assez fortes de l'ordre de .30

En ce qui concerne les échelles externes on constate également une augmentation des corrélations en situation Biaisée pouvant aller jusqu'à .50..

En ce qui concerne les relations entre les échelles fondamentales et les échelles externes on constate également une augmentation en nombre et en valeurs des corrélations quand on passe de la situation Sincère à la situation biaisée, pouvant aller jusqu'à .20 d'augmentation

Conclusions concernant les corrélations

La falsification a augmenté la valeur des inter corrélations entre les échelles fondamentales, de même qu'entre les échelles externes, ainsi qu'entre les échelles fondamentales et les échelles externes, ce qui est sans doute la conséquence de

l'adoption d'un modèle unique : la Désirabilité Locale Perçue, qui induit, par exemple, des augmentations conjointes des dimensions DSG et AMA ou Conscience et Stabilité.

2.2. Les effets de la Falsification sur la validité.

Nous avons analysé l'influence de la Falsification sur la validité des échelles fondamentales et externes de SOCRATE vis à vis d'un critère interne et vis à vis d'un critère externe. Le critère interne concerne les résultats obtenus en situation Sincère et le critère externe les résultats obtenus au NEO PI répondu en situation sans enjeu.

Validité interne

Les corrélations entre les indices de contrôles sont faibles de l'ordre de .20, cependant les indices concernant les caractéristiques du vecteur de réponses MOY et VAR sont plus fortes de l'ordre de .50..

Les corrélations entre les échelles fondamentales sincères et falsifiées sont faibles de l'ordre de .30

Les corrélations pour les échelles externes sont comparables, également de l'ordre de .30.

En estimant la fidélité de ces échelles aux alentours de .75 *la validité a été diminuée de plus de la moitié*, ce qui est à mettre en parallèle avec la diminution des écarts types quand on passe de la situation Sincère à la situation Biaisée.

Validité externe

En situation Sincère les validités des échelles externes corrélerent avec les échelles correspondantes du NEO PI de .50 à .70. En situation Biaisée elles vont .07 à .28 .

La validité est pratiquement annulée pour C et N (qui sont les composantes les plus importantes de DSG) ainsi que pour A et diminuées de moitié pour O et des 2/3 pour E.

Conclusions

La validité des résultats obtenus en situation Biaisée, est diminuée de moitié vis à vis de critères internes et annulée pour les dimensions les plus falsifiées ou diminuée de plus de moitié pour les autres, vis à vis de critères externes.

La perte d'information peut donc être considérable et il est donc essentiel de pouvoir quantifier la falsification et identifier les protocoles les plus falsifiés de manière à pouvoir utiliser Socrate en situation de sélection. C'est ce que nous avons fait

2.3. La quantification de la Falsification.

Ayant décrit la falsification nous l'avons quantifiée par l'inverse de la convergence entre le profil « sincère » et le profil « falsifié » ce qui n'est possible que dans le cadre de la méthode fonctionnelle.

Nous avons regroupé nos données en fonction du degré de falsification (ou son contraire l'honnêteté) et effectué des analyses de validations pour chacun des 6 groupes ainsi créés.

L'analyse de la validité « interne » et « externe » des dimensions de Socrate pour chacun de nos 6 groupes montre une dégradation rapide de ces validités en fonction du degré de falsification et seuls les deux groupes les plus « Honnêtes » conservent des validités acceptables.

Nous avons alors abordé le problème par le biais de la motivation qui est la composante essentielle de la Falsification, en prédisant « l'état d'esprit » ou en d'autres termes la situation dans laquelle a été recueilli les données : Sincère ou Biaisée. Pour éviter le reproche de ne reposer que sur des données « simulées » nous avons développé la prédiction de la falsification à partir de données réelles : un échantillon de sujet ayant répondu à Socrate en situation sans enjeu, de conseil et un échantillon de sujets ayant répondu en situation avec enjeu, de recrutement.

Ultime analyses : PROSE (Probabilité d'un biais de sélection)

L'effet de la probabilité d'un biais de sélection sur la validité de SOCRATE

Nous avons étudié l'effet du biais de sélection sous de nombreux angles : d'abord l'effet de la situation, avec et sans enjeu, en situation réelle et en situation simulée et ensuite l'effet de l'indice de biais de sélection (PROSE) en fonction de sa valeur plus ou moins forte, en situation simulée

L'effet de la situation

En situation simulée nous avons étudié la validité interne et externe de SOCRATE. Nous ne présentons ici, pour simplifier la discussion et la présentation, que les validités relatives aux échelles externes de SOCRATE mesurant les dimensions des BIG FIVES : O (ouverture), C (conscience), E (extraversion), A (amabilité) et N (névrosisme ou stabilité -).

277 sujets (étudiants en psychologie) ont répondu à SOCRATE avec des instructions de sincérité (sans enjeu) et avec des instructions de situation de sélection en vue d'obtenir un poste à responsabilité (avec enjeu). La validité des réponses données en situation avec enjeu vis-à-vis des réponses données en situation sans enjeu sont de .37 pour O, .29 pour C, .34 pour E, .47 pour A et .28 pour N. Par rapport à la fidélité de ces échelles qui sont de l'ordre de .75 et plus on peut admettre *que la situation a induit une diminution de la validité des échelles de l'ordre de la moitié.*

Ces mêmes sujets ont également répondu en situation sans enjeu au test NEOPI mesurant les mêmes dimensions OCEAN. La validation externe des échelles de SOCRATE en situation sans enjeu est de .49 pour O, .64 pour C, .68 pour E, .47 pour A et .59 pour N, tandis qu'en situation avec enjeu on obtient .26 pour O, .07 pour C, .26 pour E, .16 pour A et .16 pour N.

L'enjeu diminue considérablement la validité externe de O et E et annule complètement la validité des dimensions les plus sensibles C, N et A.

L'effet de la probabilité du biais de sélection

La probabilité du biais de sélection a été déterminée à partir de données réelles impliquant un échantillon de 400 sujets ayant répondu à SOCRATE en situation sans enjeu et 400 sujets ayant répondu à SOCRATE en situation avec enjeu (sélection ou promotion) et en calculant une équation de régression multiple logistique concernant le critère de la situation avec ou sans enjeu., permettant de calculer la probabilité d'appartenir à l'un ou l'autre groupe. Cette équation a permis de classer correctement 68% des sujets de la situation sans enjeu et 71% de ceux avec enjeu. Appliquée aux étudiants de notre expérience de simulation elle a permis de classer 83% des sujets ayant répondu avec des consignes de sincérité et 85% de ceux ayant répondu avec des consignes d'enjeu. Nous avons appelé cette probabilité PROSE.

Comme en situation réelle nous ignorons si le répondant a répondu de manière sincère ou biaisée nous avons utilisé toutes nos données étudiantes (sincères et biaisées) et nous les avons scindées en 5 groupes en fonction de PROSE : 1 de 0 à 27%, 2 de 28 à 42%, 3 de 43 à 57%, 4 de 58 à 67% et 5 de 66 à 89%. Et pour chacun de ces groupes nous avons calculé la relation entre les échelles OCEAN de SOCRATE et celles du NEOPI (répondu en situation sans enjeu).

Le groupe 1 a 114 sujets, le 2 108, le 3 112, le 4 123 et le 5 97. Nous avons constaté 5% de protocoles répondu en situation avec enjeu simulé dans le groupe1, 22% dans le groupe 2, 50% dans le groupe 3, 16% dans le groupe 4 et 2% dans le groupe 5.

Pour le groupe 1 les validités : .44 pour O, .68 pour C, .60 pour E, .44 pour A et .40 pour N

Pour le groupe 2 les validités : .50 pour O, .52 pour C, .59 pour E, .45 pour A et .33 pour N.

Pour le groupe 3 les validités : .40 pour O, .31 pour C, .55 pour E, .40 pour A et .16 pour N.

Pour le groupe 4 les validités : .15 pour O, .12 pour C, .33 pour E, .23 pour A et .09 pour N.

Pour le groupe 5 les validités : .21 pour O, -.05 pour C, .11 pour E, .11 pour A et .08 pour N.

Cependant comme la sélection des groupes s'est effectuée sur la base de PROSE, une variable extérieure aux échelles utilisées, un phénomène indirect de restriction of range peut être attendu. Nous avons donc corrigé ces valeurs à l'aide de la formule préconisée par THORNDIKE pour ce genre de situation.

Pour le groupe 1 les validités : .55 pour O, .71 pour C, .52 pour E, .56 pour A et .77 pour N.

Pour le groupe 2 les validités : .62 pour O, .48 pour C, .56 pour E, .25 pour A et .61 pour N.

Pour le groupe 3 les validités : .35 pour O, .28 pour C, .51 pour E, .39 pour A et .44 pour N.

Pour le groupe 4 les validités : .33 pour O, .13 pour C, .40 pour E, .23 pour A et .12 pour N.

Pour le groupe 5 les validités : .36 pour O, -.40 pour C, .03 pour E, .50 pour A et .14 pour N.

Ce sont ces valeurs qui ont été utilisées pour construire le graphique qui fait la synthèse des résultats de nos recherches sur le biais de sélection et que nous pouvons résumer ainsi : Le biais de sélection existe et il se manifeste d'une manière analogue en situation réelle et en situation biaisée ainsi que l'atteste l'efficacité de sa quantification (PROSE), développée sur des données réelles et appliquée aux données simulées (contre validation).

Le biais de sélection est le résultat d'un processus complexe qui découle essentiellement de l'importance de l'enjeu et de l'incitation à falsifier ou contraste entre le profil sincère d'un sujet et le profil idéal supposé permettre de remporter l'enjeu.

Le biais de sélection a pour effet principal d'augmenter le score des échelles « désirables » mais aussi de façon plus insidieuse de diminuer l'information véhiculée par toutes les échelles en réduisant leurs variances d'environ $\frac{1}{4}$ ainsi que de modifier la structure des échelles en diminuant la diversité des contenus par l'augmentation des inter corrélations.

Tout ceci a pour conséquence ultime **d'altérer sérieusement la validité** tant interne qu'externe de l'instrument ainsi que le démontre avec force et clarté le graphique.

Les conséquences de la connaissance de PROSE

Le graphique met en évidence que les deux dimensions des BIG FIVES qui sont les plus importantes pour la prédiction de la réussite dans de très nombreuses professions, (ainsi qu'il ressort des méta analyses effectuées sur ce sujet) à savoir CONSCIENCE et STABILITE (N-) présentent une détérioration quasi linéaire de leurs validités en fonction de l'augmentation de PROSE pour aboutir à une validité nulle (non significative) pour les groupes 4 et 5 et même inversée pour C dans le groupe 5 ! La troisième dimension à être désirable et valide, bien que de façon moindre, Extraversion présente aussi une décroissance quasi linéaire mais un peu moins marquée mais aboutissant aussi à des validités nulles pour les groupes 4 et 5.

Ouverture et Amabilité dont la désirabilité est contestable dans ce contexte ne sont pratiquement pas ou peu affectées par l'augmentation de PROSE et de façon non linéaire.

Ce résultat ne peut qu'alerter les psychologues qui utilisent des tests d'évaluations subjectives en situation avec enjeu (sélection ou promotion), car sans la connaissance de la probabilité d'un biais de sélection ils risquent fortement de se fonder sur des données non fiables et de recommander les sujets dont le profil est le plus désirable, ceux qui correspondent le mieux aux exigences du poste mais qui sont aussi ceux qui ont le plus biaisé leurs réponses, de façon à mieux correspondre au profil idéal. Pire en agissant ainsi ils risquent de recommander des sujets dont la dimension CONSCIENCE, au lieu d'être simplement impossible à connaître serait faible.

A ce stade de nos constats il est bon de rappeler que PROSE est une probabilité et non une certitude et que même si c'est improbable un PROSE élevé peut correspondre à un protocole valide et qu'un PROSE bas peut correspondre à un protocole biaisé. Il en découle que l'activité du psychologue en situation avec enjeu est une activité complexe qui exige de l'expertise, des instruments adéquats pouvant détecter et quantifier le biais de sélection, de l'expérience et de l'intuition et qui comme toute activité humaine est sujette à des aléas et à des erreurs ; La connaissance de PROSE permet de minimiser ces erreurs en utilisant les recommandations suivante : en présence d'un PROSE très élevé de l'ordre de 75 ou plus il est préférable, compte tenu des risques encouru, de déclarer qu'il est impossible de pouvoir donner un avis sur la personnalité ou les compétence du sujet à partir de ses réponses à SOCRATE. En présence d'un PROSE bas de l'ordre de .40 ou moins le protocole peut être déclaré valide, mais il est peu probable qu'il corresponde bien aux exigences du poste.

C'est donc sur la partie « centrale » de PROSE, entre .40 et .60, la zone d'indécision concernant la nature du protocole : Sincère ou Biaisé, qu'il faut se concentrer car il s'agit de protocoles qui présentent des caractéristiques désirables sans toutefois correspondre exactement au profil idéal et qui sont soit sincère soit qui ont quelque peu augmenté leurs qualités et minimisé leurs défauts.

Il faut donc essayer de démêler le vrai du faux en s'appuyant sur le curriculum vitae et sur les entretiens, sur ses impressions et sur ses intuitions, sur son sens clinique.

Si on veut prendre plus de risque, et que l'on a une grande expérience et une grande confiance en son sens clinique, on peut tenter de découvrir **la perle rare** dans les protocoles ayant un PROSE élevé, entre .60 et .75.

CONCLUSIONS

Au terme de ces 5 années de recherches nous pensons avoir exploré ce qu'il nous était possible de faire dans la situation qui est la nôtre : essayer d'étudier le mécanisme de la Falsification, non pas à partir de la réalité d'une sit

uation avec enjeu, mais d'une simulation de cette situation, effectuée par des étudiants en psychologie qui n'ont pas une grande expérience du monde du travail. Malgré les problèmes et les limites d'une telle situation les résultats auxquels nous sommes parvenus nous semblent susceptibles d'être généralisés aux situations réelles.

En effet nous avons pu montrer que des personnes incitées à falsifier un questionnaire comme SOCRATE était capable de le faire et ceci dans la direction d'une dimension que nous appelons la Désirabilité Locale Perçue, qui est suffisamment commune à l'ensemble des participants pour donner des résultats cohérents et qui sont les suivants : La falsification existe et elle dépend essentiellement d'une part de la motivation à falsifier qui est fonction de l'importance de l'enjeu (ou de l'incitation à le faire en situation de simulation) et d'autre part de L'Incitation à Falsifier c'est à dire de la différence qui existe entre l'image que le sujet a de lui même et de l'image qu'il a du profil idéal qui est recherché par le commanditaire, et qui est fonction de la description du poste mis au « concours » et des connaissances que le sujet peut en avoir pour avoir occupé des postes similaires.

La falsification, en relation avec des postes de cadres à responsabilité (le poste proposé pour la simulation) a pour effet de faire augmenter les scores de certaines échelles qui sont en relation avec la Désirabilité Sociale et avec la Désirabilité Professionnelle Générale, c'est à dire en ce qui concerne Socrate la dimension de Désirabilité Sociale et professionnelle Générale, de Amabilité et de façon moindre de Conventionnel et en terme des Big Five de Stabilité (DS Sociale), de Conscience (DS Professionnelle) et de façon moindre de Extraversion.

Il ressort de ceci que la Désirabilité Locale Perçue Moyenne pour un poste de cadre à responsabilité est l'image d'une personne Assurée, Dynamique et Entreprenante, Equilibrée et Stable, Consciencieuse et Travailleuse, Conciliante et Diplomate, et quelque peu Conventionnelle. Il nous semble que ce portrait peu s'appliquer à un grand nombre de situations de cadres à responsabilités, ce qui est une garantie de l'utilité de nos résultats et de leur transposition possible dans la réalité du monde professionnel.

La falsification a aussi pour effet de faire augmenter certains indices de contrôle propre à Socrate comme la Cohérence, la Moyenne et l'indice IF. En d'autre termes la falsification rend la description qui est faite plus stéréotypée, plus large et plus positive et donnant à la dimension DS une plus grande importance par rapport aux autres dimensions.

Enfin la falsification a pour effet de diminuer la variance de toutes les dimensions de moitié, ce qui indique une forte perte d'information véhiculée par chacune des dimensions même pour celles dont la moyenne n'a pas été augmentée. Cela indique que les différences qui pouvaient exister entre les sujets sont largement gommées pour faire apparaître un groupe compact autour de la Désirabilité Locale Perçue Moyenne.

Ceci confirme que la méthode ipsative qui est souvent préconisée pour lutter contre la Falsification est totalement inefficace d'une part parce qu'elle fait disparaître la dimension de Désirabilité qui est la dimension la plus importante, ensuite que les autres dimensions restent affectées par la falsification et qu'enfin les scores ipsatifs sont moins fiables que ceux en échelles de Likert du fait de la perte de degré de liberté, ainsi que nous avons pu le montrer lors de nos premières expérimentations.

Ayant décrit la falsification nous l'avons quantifiée par l'inverse de la convergence entre le profil « sincère » et le profil « falsifié » ce qui n'est possible que dans le cadre de la méthode fonctionnelle.

Nous avons regroupé nos données en fonction du degré de falsification (ou son contraire l'honnêteté) et effectué des analyses de validations pour chacun des 6 groupes ainsi créés.

Les résultats ont confirmés de manière éclatante nos constatations concernant la dégradation de la validité en situation falsifiée.

En effet l'analyse de la validité « interne » et « externe » des dimensions de Socrate pour chacun de nos 6 groupes montre une dégradation rapide de ces validités en fonction du degré de falsification et seuls les deux groupes les plus « Honnêtes » conservent des validités acceptables.

Nous avons développé une formule permettant pour chaque protocole de calculer sa probabilité d'être « Honnête » ou « Falsifié », et ceci à partir de données réelles, non simulées,

Que nous avons appelé PROSE. Nous avons pu montrer la validité de cette formule pour les données simulées : en ne retenant que les sujets ayant une probabilité de .60 d'être « Honnête » on pouvait garantir une validité satisfaisante alors que, en ne retenant les sujets ayant une probabilité de .60 d'être « Falsificateur », c'est à dire ceux qui paraissent les plus désirables, la validité devenait nulle.

Au terme de cette série de recherches et de résultats nous pensons qu'ils sont susceptibles d'être dès à présent utilisés en pratique.

Dernières analyses (2017) : PROSEC

Probabilité de biais de sélection (falsification) conditionnelle

Francis GENDRE

Rapport Technique

INTRODUCTION

Nous avons pu montrer que PROSE, développé à partir de données réelles à l'aide d'une équation de régression logistique prédisant « l'état d'esprit » (ou situation avec ou sans enjeu), était très efficace pour prédire la situation « sincère » ou « simulée » d'étudiant(e)s en psychologie.

Nous avons pu montrer que, en mélangeant les protocoles sincères et simulés et en déterminant 5 groupes de sujets ayant PROSE très faible, faible, moyen, fort et très fort, la validité des échelles OCEAN de SOCRATE vis à vis de celles du NEO PI, répondu en situation sincère, diminuait de façon quasi linéaire pour les 3 dimensions les plus désirables professionnellement à savoir N- (Stabilité) et C (conscience) ainsi que E (Extraversion), et devenant non significatives (nulles) pour les groupes 4 et 5.

Nous en avons déduit que PROSE plus qu'une simple prédiction de l'état d'esprit était également un indicateur de l'intensité de la falsification, un indicateur de la validité du protocole.

Il en résulte qu'en situation de sélection la « bonne » décision consistera à déclarer invalides les protocoles ayant PROSE élevé et très élevé. Cependant le problème est que PROSE est d'une part aussi un indicateur de la désirabilité du profil du répondant par rapport à un poste à responsabilité et d'autre part une probabilité et non une information absolue. Ainsi, même improbable, un PROSE élevé peut être « sincère » et ce type de protocole est celui qui est le plus recherché.

C'est pourquoi il nous a semblé essentiel de pouvoir détecter ces « perles rares », qui présentent toutes les qualités désirables de manière sincère et non simulées.

Il en découle l'idée de calculer pour chacun des groupes de sujets ayant PROSE très bas, bas, moyen, élevé et très élevé un PROSEC, une **probabilité de biais de sélection**

conditionnelle, en utilisant la même procédure que pour PROSE mais sur des données recueillies en situation « sincère » et « simulée » et pour chacun des niveaux de PROSE. Nous avons utilisé les données de 922 protocoles, ayant été répondu par 461 étudiants une fois en situation sincère et une fois en situation simulée.

Le groupe 1 va de PROSE .05 à .259, le 2 de .26 à .415, le 3 de .416 à .563, le 4 de .564 à .663 et le 5 de .664 à .889.

Ces limites ont été choisies pour avoir des groupes à peu près égaux ayant au moins 160 sujets pour satisfaire aux exigences de stabilité d'une équation de régression multiple comprenant 5 prédicteurs ($30 \times 5 + 10$)

GROUPE 1, PROSE TRES FAIBLE.

Il est constitué de 184 sujets dont 4 sont des falsificateurs.

Les corrélations simples avec les prédicteurs sont toutes non significatives à l'exception de MOY (.25) et il en est de même pour les poids bétas (MOY .25), aboutissant à une corrélation multiple de .30. La régression multiple logistique classe correctement 95% des cas, exactement le % de départ.

Le nombre de sujets « falsificateurs et la faible validité d'un seul prédicteur la Moyenne ne permet pas d'améliorer la situation de départ.

Conclusions : les sujets classés dans le groupe 1 en fonction de PROSE doivent être déclarés *Sincères*, avec un risque minime de 5%.

Groupe 2, PROSE FAIBLE.

Il est constitué de 185 sujets dont 18,4% de falsificateurs.

Les corrélations simples et les poids bétas (régression multiple utilisant tous les prédicteurs) mettent en évidence 5 prédicteurs COT (.14 et .20), MOY (.59 et .55), AMA (.25 et .20) auxquels nous avons rajouté VAR (-.02 et -.11). et IND (.11 et .11)

L'équation de régression multiple utilisant ces prédicteurs aboutit à une corrélation multiple de .52 avec les poids bétas suivants : COT .07, MOY .57, VAR -.11, AMA .12 et IND .11.

L'équation de régression logistique aux coefficients suivants : constante -24.765), 1.832 (COT), 6.427 (MOY), -3.808 (VAR), 3.299 (AMA) et 3.601 (IND) aboutit à la table de contingence suivante :

	1	2	TOTAL	%correct
1	146	5	151	96.69
2	16	18	34	52.94
TOTAL	162	23	185	88.65

Conclusions : Prosec, en attribuant les sujets dont la probabilité est supérieure à .50 au groupe 2 (falsifié), permet de classer correctement 164 des 185 sujets au lieu de 162. Le gain est minime. On peut donc comme pour le groupe 1 déclarer les sujets classés dans le groupe 2 comme *Sincères*, mais avec un risque de l'ordre de 18%, que l'on peut légèrement réduire en utilisant PROSEC.

Groupe 3, PROSE moyen.

Il est constitué de 186 sujets dont 51,1% de falsificateurs.

Les corrélations simples et les poids bétas (régressions utilisant tous les prédicteurs) mettent en évidence 5 prédicteurs COP (.24 et .17), MOY (.63 et .60), VAR (.09 et -.14), CONS (.20 et .09) et AMA (.35 et .12).

L'équation de régression multiple utilisant ces prédicteur aboutit à une corrélation multiple de .67, avec les poids bétas suivants COP .11, MOY .59, VAR .01 CONS .13 et AMA .11.

L'équation de régression logistique aux coefficients suivant : -21.911 (constante), 3.181 COP, 5.744 MOY, -3.316 VAR, 2.319 CONS et 2.978 AMA, aboutit à la table de contingence suivante :

	1	2	Total	%correct
--	---	---	-------	----------

1	76	15	91	83.52
2	17	78	95	82.11
Total	93	93	186	82.80

Conclusions : PROSEC en attribuant les sujets dont la probabilité est supérieurs à .50 au groupe 2 (falsificateurs) permet de classer correctement 83% des cas contre 50% en utilisant le base rate (tous sincères ou tous falsifiés). Le gain est donc considérable avec un risque de 17% tout a fait acceptable.

Cependant il faut aussi comparer au fait d'attribuer les sujets ayant PROSE supérieur à 50 au groupe 2 et ceux inférieur au groupe 1. La table de contingence est :

	1	2	Total	%correct
1	54	39	93	58.06
2	37	56	93	60.22
Total	91	95	186	59.14

Le gain de 83% par rapport à 59% reste considérable.

Groupe 4, PROSE élevé.

Il est constitué de 185 sujets dont 86,5 % sont des falsificateurs.

Les corrélations simples et les poids béta de l'équation utilisant tous les prédicteurs permet de mettre en évidence 5 prédicteurs : COP 5.36 ET -.29), MOY (.50 et .35), VAR (.03 et -.15) CONS (.43 et .43) et AMA (.39 et .45).

L'équation de régression multiple utilisant ces prédicteurs aboutit à une corrélation multiple de **.66** avec les poids béta suivants :COP -.16, MOY .38, VAR -.11, CONS .39 et AMA .36.

L'équation de régression logistique aux coefficients suivants :-16.297,(constante) -6.955 COP, 4.594 MOY, -2.014 VAR, 9.823 CONS et 11.510 AMA, aboutit à la table de contingence suivante :

	1	2	Total	%correct
1	14	11	25	56
2	4	156	160	97.52
Total	18	167	185	91.89

Conclusions : PROSEC en attribuant au groupe 2 (falsificateurs) les sujets ayant une probabilité supérieure à .50, permet d'obtenir une efficacité globale de 92% au lieu de 87% si on attribuait tous les sujets au groupe des falsificateurs. Le gain semble assez faible mais en fait il est considérable car au lieu d'éliminer 18 sujets qui sont les plus importants car ce sont des « perles » rares, cela permet d'en sauver 14 mais avec un risque élevé de 44%, car ils sont mélangés avec 11 falsificateurs mais cela en vaut la peine.

GROUPE 5 , PROSE très élevé.

Il est constitué de 182 sujets dont 89.6 % sont des falsificateurs.

Les corrélations simples et les poids béta d'une régression multiple utilisant tous les prédicteurs permet de mettre en évidence les 5 prédicteurs suivants : MOY 5.32 et .16), DSG (.17 et .23), CONS (.32 et .47), AMA (.13 et .42) et ACT (.08 et .20).

L'équation de régression multiple utilisant ces prédicteurs aboutit à une corrélation multiple de **.53** avec les poids béta suivants :.16 MOY, .28 DSG, .43 CONS, .32 AMA et .20 ACT.

L'équation de régression logistique aux coefficients suivants -14.654 (constante), 0.104 MOY, 22.649 DSG, 16.354 CONS, 16 .131 AMA et 7.911 ACT, aboutit à la table de contingente suivante

	1	2	Total	%correct
1	9	10	19	47.37

2	5	159	163	96.93
	14	169	182	91.76

Conclusions : PROSEC en attribuant au groupe 2 (falsificateurs) les sujets ayant une probabilité supérieure à .50 permet d'obtenir une efficacité globale de 92% au lieu de 90% en attribuant tous les sujets au groupe des falsificateurs. Le gain paraît très faible mais comme pour le groupe 4 il est considérable car au lieu d'éliminer 14 sujets qui sont les plus importants car ce sont des « super perles » rares, cela permet d'en sauver 9 mais avec un risque élevé de 53% car ils sont mélangés à 10 falsificateurs.

Nous avons ensuite étudié le regroupement des groupes 1 et 2 et des groupes 4 et 5.

Groupe 1 et 2, PROSE faible.

Il est constitué de 369 sujets dont 11,2% sont des falsificateurs.

L'équation de régression multiple utilisant les prédicteurs suivants .183 COT, -.158 COP, .470 MOY, .121 CONS et .181 AMA aboutit à une corrélation multiple de .53.

L'équation de régression logistique aux coefficients suivants -22.045(Cte) 4.156 COT, -3.532 COP, 5.601 MOY, 3.294 CONS et 4.445 AMA aboutit à la table de contingence suivante :

	1	2	Total	%correct
1	324	2	326	99.39
2	21	22	43	51.16
Total	345	24	369	93.77

Ce résultat est un peu meilleur qu'en séparant les groupes 1 et 2, et doit donc lui être préféré.

Groupe 4 et 5, PROSE élevé.

Il est constitué de 367 sujets dont 88% de falsificateurs.

L'équation de régression multiple utilisant les prédicteurs suivants COP, MOY, VAR, CONS, AMA aboutit à une corrélation multiple de .58.

L'équation de régression logistique aux coefficients suivants 0.389 (Cte), -0.397 COP, 0.360 MOY, -0.171VAR, 1.098 CONS et .709 AMA aboutit à la table de contingence suivante :

	1	2	Total	%correct
1	20	24	44	45.45
2	5	318	323	98.45
Total	25	342	367	92.10

L'addition des groupes 4 et 5 donne

	1	2	Total	% correct
1	23	21	44	52.27
2	9	314	323	97.21
Total	32	335	367	91.82

Ce résultat étant un peu meilleur qu'en séparant les groupes 4 et 5 il doit lui être préféré.

Conclusions. Ces résultats ayant été obtenus sur nos données étudiantes nous allons maintenant étudier notre groupe réel en faisant une tri partition.

GROUPE REEL.

Nous avons effectué une tri partition analogue à celle effectuée sur les données étudiantes.

Le groupe 1 comprend 260 sujets, le groupe 2, 151 sujets et le groupe 3, 253 sujets. Les corrélations simples sont très faibles la plus forte pour le groupe 1 étant de .12 pour DSG, pour le groupe 2 de .11 pour COP et pour le groupe 3 de .15 pour DSG. Les corrélations multiples utilisant toutes 10 les variables sont faibles pour tous les groupes : .17 pour le groupe 1, .26 pour le groupe 2 et .22 pour le groupe 3. Dans ces conditions on ne peut pas améliorer la prédiction découlant de PROSE. Nous devons donc nous contenter des résultats obtenus en situation de simulation sur des étudiant(e)s en psychologie de sexe féminin en majorité..... Ces résultats sont cependant probants et nous pensons qu'ils sont susceptibles de pouvoir être utilisés en pratique avec les précautions d'usage concernant le passage d'une simulation à la réalité.

L'influence de l'utilisation de PROSE et des PROSEC sur la validité des protocoles des étudiants sincères et falsifiés vis à vis du NEO PI répondu en situation sincère.

ECHELLES OCEAN DE SOCRATE ET DU NEO PI

Nous allons examiner l'évolution des validités conceptuelles des échelles OCEAN- de Socrate vis à vis des mêmes échelles du NEO PI pour différents groupes de sujets.

Tous les sujets (N = 554)

O	C	E	A	N
37	37	47	34	-33

Tous les « sincères » (N = 277)

O	C	E	A	N
49	64	68	47	-59

Tous les « falsificateurs » (N = 277)

O	C	E	A	N
26	07	26	16	-16

Tous ceux ayant PROSE inférieur à 50 (N = 268)

O	C	E	A	N
49	58	60	43	-46

Tous ceux ayant PROSE supérieur ou égal à 50 (N = 286)

O	C	E	A	N
19	02	32	24	-02

Nous constatons que PROSE détermine 2 groupes dont les validités conceptuelles sont très proches des validités conceptuelles des « sincères » et des « falsifiés ».

Elles sont quelques peu inférieures pour le premier groupe (inf à 50) car quelques falsificateurs s'y cachent et pour le second groupe (sup à 50) elles diffèrent quelques peu du fait que des sincères s'y cachent également.

Le groupe 1, PROSE inférieur à 40 (N = 214)

O	C	E	A	N
50	56	63	50	-50

Ceux ayant PROSEC inférieur à 50 (prédits « sincères ») (N = 203)

O	C	E	A	N
51	57	64	50	-54

Ceux ayant PROSEC supérieur à 50 (prédits falsificateurs) N = 11

O	C	E	A	N
27	60	58	66	-23

Les validités conceptuelles sont pratiquement identique même si elles sont très légèrement supérieures après élimination des 11 sujets détectés « falsificateurs » dont les validités conceptuelles sont de même niveau et très proche des « sincères » Vrais. Ce résultat confirme que pour ce groupe on peut déclarer tous les protocoles valides, PROSEC n'amenant pas d'amélioration.

Le groupe 2, PROSE compris entre 40 et 59 (N = 118)

	O	C	E	A	N
	38	41	48	38	-39

Ceux ayant PROSEC inférieur à 50, prédits « sincères » N = 61

	O	C	E	A	N
	46	64	66	48	-68

Ceux ayant PROSEC supérieur ou égal à 50 (prédits « falsificateurs ») N = 57

	O	C	E	A	N
	30	-06	23	27	06

Ce résultat confirme l'efficacité de la séparation des prédits sincères de ceux prédits falsificateurs pour ce groupe central.

Le groupe 3, Prose supérieur à 60 (N = 222)

	O	C	E	A	N
	23	19	27	14	-27

Ceux ayant PROSEC inférieur à 50 (prédits sincères) N = 14

	O	C	E	A	N
	70	45	45	30	-56

Ceux ayant PROSEC supérieur ou égal à 50 (prédits falsificateurs) N = 208

	O	C	E	A	N
	18	19	26	12	-28

Ce résultat mettant en évidence la validité conceptuelle des sujets prédits « sincères » par PROSEC confirme l'intérêt de PROSEC pour identifier dans ce groupe des « perles rares » qui s'y cachent.

ECHELLES FONDAMENTALES DE SOCRATE ET OCEAN DU NEO PI.

Pour étudier les relations entre les échelles fondamentales et les échelles OCEAN du NEO PI nous avons calculé les corrélations simples, les corrélations multiples ainsi que les poids betas qui y sont associés.

Tous les sujets « sincères ». N = 277

	Validités					Bétas					R2	R
	DSG	CONS	AMA	ACTI	IND	DSG	CONS	AMA	ACTI	IND		
O	.034	-.428	.101	-.251	-.014	-.042	-.481	.022	-.266	.146	.272	.477
C	.342	.347	-.130	.355	.170	.435	.455	.034	.339	.075	.427	.654
E	.385	-.465	.336	.227	-.346	.313	-.293	.333	.270	-.180	.492	.701
A	-.202	.010	.417	.099	-.022	-.162	.027	.436	.171	.024	.231	.481
N	-.397	-.155	-.284	.235	.071	-.509	-.318	-.339	.203	.019	.386	.621

Dans le groupe sincère les échelles du NEO PI les mieux prédites sont Extraversion, Conscience et Névrosisme, Ouverture et Amabilité l'étant moins, ce qui correspond par

ailleurs aux observations effectuées sur les relations entre les échelles OCEAN de SOCRATE.

Du fait de la quasi orthogonalité des échelles fondamentales les validités et les poids béta sont très proches.

Tous les sujets « falsifiés ». N = 277

	Validités					Bétas					R2	R
	DSG	CONS	AMA	ACTI	IND	DSG	CONS	AMA	ACTI	IND		
O	-.038	.113	.120	-.145	.022	-.001	.122	.116	-.145	.022	.048	.219
C	.129	-.068	.141	.062	.010	.135	-.045	.162	.071	.060	.048	.219
E	.064	.036	.024	-.065	.083	.076	.027	.034	-.080	.091	.019	NS
A	.028	.145	.000	.071	.043	.046	.153	.027	.067	-.006	.028	NS
N	.042	.005	-.019	-.040	-.002	.045	.011	-.019	-.046	-.001	.004	NS

Dans le groupe falsifié les échelles Extraversion, Amabilité et Névrosisme ne sont plus prédites tandis que Ouverture et Conscience ne le sont que très faiblement, mettant en évidence la forte détérioration de la validité de Socrate produite par la falsification.

PROSE inférieur à 50. N = 264

	Validités					Bétas					R2	R
	DSG	CONS	AMA	ACTI	IND	DSG	CONS	AMA	ACTI	IND		
O	.040	-.442	.111	-.271	.016	.003	-.489	.053	-.286	.204	.303	.550
C	.281	.363	-.178	.361	.189	.301	.386	-.057	.316	.027	.354	.595
E	.298	-.424	.282	.213	-.272	.271	-.308	.294	.257	-.183	.399	.632
A	-.194	-.008	.393	.096	.003	-.154	-.001	.394	.152	.031	.200	.447
N	-.266	-.206	-.131	.171	.055	-.348	-.303	-.177	.173	.093	.213	.462

Dans ce groupe dans lequel ne subsiste qu'un petit pourcentage de sujets falsificateurs les validités sont quasi identiques au groupe « sincère », très légèrement atténuée pour Conscience, Extraversion et Amabilité, un peu plus pour Névrosisme et légèrement améliorée pour Ouverture. On peut donc penser que les falsificateurs de ce groupe sont de « petits » falsificateurs.

PROSE supérieur ou égal à 50. N = 290

	Validités					Bétas					R2	R
	DSG	CONS	AMA	ACTI	IND	DSG	CONS	AMA	ACTI	IND		
O	.153	-.085	.064	-.185	-.005	.210	-.051	.074	-.199	.045	.078	.279
C	.086	-.034	-.022	.005	.027	.101	-.021	.010	-.006	.060	.011	NS
E	.205	-.305	.043	.004	-.320	.106	-.221	.091	-.025	-.234	.168	.410
A	-.127	.015	.243	-.055	-.041	-.081	-.016	.226	.023	-.057	.066	.257
N	-.049	.110	.009	.040	.130	.002	.080	.020	.063	.116	.027	NS

Dans ce groupe dans lequel subsiste quelques sujets sincères, les validités sont nulles pour Conscience et Névrosisme, les deux dimensions les plus affectées par la falsification. Pour les autres, Ouverture et Amabilité, elles sont très modestes et pour Extraversion la validité est assez forte.

PROSE inférieur à 40. N = 204

	Validités					Bétas					R2	R
	DSG	CONS	AMA	ACTI	IND	DSG	CONS	AMA	ACTI	IND		
O	-.009	-.432	.157	-.257	.039	-.054	-.501	.076	-.277	.241	.310	.557
C	.279	.415	-.215	.366	.221	.335	.467	-.009	.322	.040	.416	.645
E	.272	-.451	.300	.256	-.273	.260	-.302	.324	.293	-.161	.423	.650
A	-.246	-.026	.430	.100	.009	-.159	-.010	.421	.160	.066	.237	.487
N	-.241	-.198	-.118	.178	.045	-.381	-.328	-.230	.181	.064	.219	.468

Dans ce groupe ou ne subsiste qu'un faible pourcentage de « falsificateurs » probablement « modestes » les validités s'approchent des validités des données sincères

PROSE entre 40 et 60. N = 150

	Validités					Bétas					R2	R
	DSG	CONS	AMA	ACTI	IND	DSG	CONS	AMA	ACTI	IND		
O	.237	-.313	.025	-.209	-.017	.102	-.306	.008	-.230	.105	.166	.407
C	.245	.030	-.099	.228	.012	.385	.256	.045	.242	.021	.155	.394
E	.347	-.458	.061	.071	-.424	.158	-.273	.143	.079	-.304	.323	.568
A	-.054	.059	.376	-.024	.038	.110	.102	.415	.073	.015	.155	.394
N	-.056	.024	-.013	.155	.163	-.044	-.045	.006	.156	.164	.053	NS

Dans ce groupe qui comprend un peu plus de sujets falsificateurs que de sujets sincères les validités sont toutes inférieures à celles des sujets sincères et supérieures à celles des sujets falsificateurs. Elles sont très proches des validités de tous les sujets mélangés (sincères et falsificateurs et qui sont dans l'ordre : .407, .380, .510, .355 et .313. Elles sont égale pour O, supérieures pour C, E et A et inférieure (.230) pour N. Cela suggère que les « falsificateurs de ce groupe sont de s falsificateurs assez modestes sauf en ce qui concerna la dimension Stabilité.

PROSE supérieur à 60. N = 200

	Validités					Bétas					R2	R
	DSG	CONS	AMA	ACTI	IND	DSG	CONS	AMA	ACTI	IND		
O	.141	-.065	.092	-.240	-.007	.202	.005	.090	-.218	.033	.088	.297
C	.049	-.030	.003	-.043	.035	.068	-.018	.016	-.037	.055	.007	NS
E	.140	-.237	.078	-.055	-.221	.076	-.174	.079	-.036	-.105	.096	.310
A	-.077	.030	.196	-.028	.126	-.019	.064	.203	.035	.137	.059	.243
N	-.095	.154	.042	-.010	.134	-.003	.135	.058	.014	.108	.038	NS

Dans ce groupe qui comprend quelques sujets sincères, les validités sont non significatives pour Conscience et Névrosisme les dimensions les plus sensibles à la falsification, comme pour les sujets falsificateurs. Pour les autres dimensions Ouverture, Extraversion et Amabilité, elles sont supérieures aux validités des données falsifiées tout en restant très modestes. Ceci est le signe de la présence de quelques « perles rares » dans ce groupe, que nous allons essayer de détecter maintenant.

Identification des sujets sincères et falsificateurs à l'intérieur de chacun des groupes de PROSE à l'aide des probabilités conditionnelles (PROSEC) et effet sur les validités.

Pour pouvoir identifier des sujets sincères ou falsificateurs dans l'un ou l'autre des 3 groupes il faut que la proportion des sujets recherchés ne soit pas en nombre trop faible. Ceci se présente dans les groupes extrêmes. Pour pallier à cette difficulté nous avons scindé ces deux groupes en deux : pour le groupe 1 (PROSE inférieur à 40) nous préconisons de faire 2 groupes : un groupe extrême ou PROSE est inférieur à 25 (N =92), et qui ne comprend que 3% de sujets «falsificateurs » . On peut déclarer tous ces sujets Sincères. Pour le groupe 3 (PROSE supérieur ou égal à 60), nous créons un groupe extrême ou PROSE est égal ou supérieur à 75 (N= 28) et qui ne comprend aucun sujet « sincère ». On peut déclarer tous ces sujets Falsificateurs.

Pour chacun des 3 groupes retenus nous avons calculé la validité des échelles fondamentales et des indices IF et CONVAP, la corrélation multiple qui en découle et ses poids béta et sur cette base déterminé un modèle réduit qui est celui que nous préconisons, et pour lequel nous indiquons la table de vérité (Classification) et les coefficients de l'équation logistique.

Groupe 1 (PROSE supérieur ou égal à 25 et inférieur à 40), N =112

	DSG	CONS	AMA	ACTI	IND	IF	CONVAP	R
Validité (condition)	.098	.084	.247	.001	.065	-.074	-.140	
Poids béta	.649	.511	.797	-.036	.171	.243	-.208	
Coef logist standard	1.224	1.023	1.902	-.139	.395	.668	-.606	.571
Coef log st abrégé	1.147	1.087	1.677			.499	-.479	.550

	1	2	TOT	%
1	91	2	93	98
2	10	9	19	47
TOT	101	11	112	90

Coefficients PROSEC : Cte -5.320 12.345 DSG 9.719 CONS 17.294 AMA 25.464 IF et -5.397

Prosec permet de détecter 9 des 19 falsificateurs et décline 2 sincères.....Il améliore donc l'interprétation mais de manière assez modeste et ceci pour un groupe dont les sujets qu'ils soient sincère ou falsificateur n'ont pas un profil particulièrement désirable.

Effets sur la validité

Nous n'indiquons que les valeurs supérieures ou égales à .15

PROSE inférieur ou égal à 24 (N =92).

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O		-.40	.25	-.17			.44
Néo C	.31	.41	-.22	.40	.26		.68
Néo E	.30	-.45	.29	.34	-.30		.66
Néo A	-.24		.46				.43
Néo C	-.15			.23			-.36

PROSE supérieur ou égal à 25 (N=112)

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O		-.39	.26	-.34			.50
Néo C	.19	.40	-.33	.33	.17		.62
Néo E	.25	-.51	.34	.18	-.24		.61

Néo A	-.32		.41			.47
Néo N	-.18			.16		-.32
PROSEC inférieur à 50 (supposé sincères). N = 102						
	DSG	CONS	AMA	ACTI	IND	OCEAN SOCRATE
Néo O		-.42	.31	-.37		.54
Néo C	.21	.41	-.37	.31	.17	.64
Néo E	.24	-.54	.35	.15	-.25	.61
Néo A	-.34		.41			.47
Néo N	-.21			.18		-.37

PROSEC supérieur à 50 (supposé falsificateurs). N = 10

Trop peu nombreux pour avoir des valeurs stables.

Cependant le constat que les valeurs pour PROSEC inférieur à 50 sont légèrement supérieures à PROSE supérieur à 25 indiquent que cette augmentation est due à l'élimination des détectés falsificateurs et qu'ils ont donc une validité inférieurs à ceux détectés sincères.

TOUS les sujets du groupe 1 (PROSE inférieur à 40). N = 204

	DSG	CONS	AMA	ACTI	IND	OCEAN SOCRATE
Néo O		-.43	.16	-.26		.50
Néo C	.28	.42	-.22	.37	.22	.64
Néo E	.27	.45	.30	.26	-.27	.63
Néo A	-.25	.43				.45
Néo N	-.24	-.20		.18		-.46

Conclusion. Les différences constatées entre les différentes situations envisagées même si elles montrent que les meilleures validités sont obtenues dans le groupe extrême qui ne comprend qu'une très faible proportion de falsificateurs (2%) et que PROSEC permet d'améliorer la validité du groupe détecté sincère par rapport au groupe pour lequel il est proposé (PROSE sup à 25), ne diffèrent guère des validités constatées pour le groupe 1 total, que ce soit par rapport aux échelles fondamentales ou par rapport aux échelles OCEAN de Socrate vis à vis du NEO PI.

En pratique, pour ce groupe majoritairement constitué de sujets sincères et de falsificateurs « modestes » **on peut considérer tous ces sujets comme sincère sans trop de crainte de se tromper.**

Si on veut « raffiner » rien n'interdit cependant d'expérimenter PROSEC.

Groupe 2 (PROSE supérieur ou égal à 40 et inférieur ou égal à 60), N =150

	DSG	CONS	AMA	ACTI	IND	IF	CONVAP	R
Validité (condition)	.221	.163	.415	.033	.160	-.116	.190	
Poids béta	.644	.484	.648	.174	.155	.069	-.049	
Coef logist standard	1.431	.944	1.383	.440	.350	.033	-.164	.709
Coef log st abrégé	1.360	.921	1.285	.428	.313			

	1	2	TOT	%
1	56	13	69	81
2	11	70	81	86
TOT	67	83	150	

Prosec permet détecter 81% des individus sincères et 86% des individus falsificateurs. Ceci représente une grande amélioration.

Coefficient PROSEC : Cte -10.702 16.290 DSG 10.492 CONS 15.332 AMA 5.795 ACT et 5.044 IND.

Effet sur la validité

Tous les sujets du groupe 2 (PROSE ENTRE 40 ET 59)

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O	.24	-.31		-.21		.39	
Néo C	.25			.23		.32	
Néo E	.35	-.46			-.43	.52	
Néo A			.38			.32	
Néo N				.16	.16	-.14	

PROSE inférieur à 50 (supposé sincères). N = 60

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O	.39	-.48		-.32		.50	
Néo C	.25		-.18	.34		.45	
Néo E	.38	-.46	.16		-.31	.56	
Néo A			.32			.37	
Néo N				.20	.22	-.20	

PROSEC inférieur à 50 (supposé sincères). N = 67

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O	.33	-.37		-.21	-.18	.41	
Néo C	.34		-.28	.26		.44	
Néo E	.40	-.60		.19	-.50	.67	
Néo A	-.19		.40			.43	
Néo N				.33		-.20	

PROSE supérieur à 50 (supposé falsificateur). N = 90

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O		-.19				.28	
Néo C	.29					.16	
Néo E	.38	-.45			-.50	.50	
Néo A			.41			.32	
Néo N						-.11	

PROSEC supérieur à 50 (supposé falsificateurs). N = 83

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O	.18	-.31		-.22	.18	.43	
Néo C				.22		.18	
Néo E	.38	-.33			-.38	.39	
Néo A						-.07	
Néo N					.22	-.21	

Conclusions. Les validités relatives à l'utilisation de PROSEC montrent l'avantage sur la simple utilisation de PROSE tant en ce qui concerne les sujets détectés sincères, dont les validités sont globalement meilleures, qu'en ce qui concerne les sujets détectés falsificateurs dont les validités sont globalement plus faibles.

Groupe 3 (PROSE supérieur à 60 et inférieur à 75), N = 172

	DSG	CONS	AMA	ACTI	IND	IF	CONVAP	R
Validité (condition)	.070	.380	.233	-.076	.096	-.021	.176	

Poids béta	-100	.637	.641	.063	.143	.742	-.025
Coef logist standard							.611
Coef log st abrégé	1.143	1.310		.196	1.317		.606

	1	2	TOT	%
1	17	9	26	65
2	4	142	146	97
TOT	21	151	172	92

Prosec permet de détecter 65% des individus sincères et 97% des individus falsificateurs. Ceci est une amélioration considérable pour ce groupe « stratégique » ou les sujets qu'ils soient sincères ou falsificateurs ont des profils désirables.

Coefficient PROSEC : Cte -14.483 17.111 CONS 21.937 AMA 4.364 IND et 46.292 IF

Effet sur la validité

Tous les sujets. N = 200

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O					-.24		.17
Néo C							.00
Néo E		-.24					-.22
Néo A			.20				.01
Néo N		.15					.01

PROSE supérieur à 75, falsificateurs. N = 28

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O	.39	-.20					-.19
Néo C	-.15	-.35	.42				-.53
Néo E	.15		.36				-.39
Néo A	-.18		.47				.42
Néo N		.24			.18		-.12

Nous remarquons les validités négatives vis à vis de la dimension C, la plus désirable pour toutes les professions, indiquant que ce sous groupe est constitué de « gros » falsificateurs.

PROSE inférieur à 75. N= 172

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O	.17						-.25
Néo C							.16
Néo E	.16	-.27					.06
Néo A							-.20
Néo N		.17					-.15

PROSEC inférieur à 50 (supposés sincères) et PROSE inf à 75. N =21

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O	.43	-.38	.30				-.40
Néo C		.20					.33
Néo E	.38	-.44					-.18
Néo A	-.34	.30	.46				-.15
Néo N		.44	-.51				.38

PROSEC supérieur à 50 (supposés falsificateurs) et prose inf à 75. N = 151

	DSG	CONS	AMA	ACTI	IND	OCEAN	SOCRATE
Néo O							-.21
Néo C							.20
Néo E		-.20					.02
Néo A							-.19
Néo N							.17

Néo A	-0.16	0.07
Néo N		-0.05

Conclusions. L'utilisation de PROSEC permet d'isoler dans ce groupe de forts, falsificateurs, dont les protocoles ont des validités pratiquement nulles, un ilot de 21 sujets dont les validités sont satisfaisantes et dans lequel se trouvent les « perles rares » tant recherchées, avec un risque d'environ 50%.... il appartient au psychologue de faire le tri à l'aide de méthodes annexes comme l'entretien et l'analyse du curriculum vitae.

SYNTHESE ET CONCLUSIONS

ANALYSES FACTORIELLES

Coordonnées factorielles après rotation Varimax : TOUS MELANGES N= 554

	D1	D2	D3	D4	D5	D6
S_DSG	0,895	0,032	-0,298	0,056	0,192	0,028
S_CONS vs INNOV	0,081	-0,920	0,069	-0,142	0,097	0,157
S_AMA vs TENS	0,542	0,056	0,752	-0,216	-0,116	0,056
S_ACTI vs IMAGI	-0,029	-0,161	0,045	0,946	0,009	-0,074
S_IND vs COLL	-0,249	-0,263	-0,047	0,071	0,018	0,742
S_IF	0,443	0,165	-0,492	0,158	-0,117	0,051
S_CONVAP	0,275	0,220	0,167	-0,144	-0,008	0,606
S_bigo	-0,198	0,861	-0,068	-0,333	-0,136	0,032
S_bigc	0,603	-0,650	-0,119	0,325	0,222	0,091
S_bige	0,756	0,396	0,149	0,344	0,035	-0,201
S_biga	-0,233	-0,136	0,872	0,170	-0,201	-0,045
S_bigs	0,889	-0,221	-0,018	-0,268	0,174	-0,028
PROSE_S	0,928	-0,242	0,017	-0,065	0,026	0,046
neo-o	0,064	0,525	-0,011	-0,135	-0,026	0,353
neo-c	0,037	-0,202	-0,053	0,278	0,718	0,204
neo-e	0,286	0,495	0,221	0,250	0,356	-0,212
neo-a	-0,043	0,136	0,590	0,087	0,274	0,122
neo-n	-0,203	0,051	-0,072	0,287	-0,729	0,192

Le premier facteur est un facteur PROSE de falsification incluant DSG, AMA, IF, C et S, dans lequel les échelles du NEOPI ne sont pas concernées.

Le second facteur est un facteur d'Ouverture incluant Innovation, O, C-, ainsi que les échelles O et E du Néo PI.

Le troisième facteur est un facteur d'Amabilité, incluant Amabilité, A et IF négatif ainsi que l'échelle A du Néo PI.

Le quatrième facteur est un facteur d'Activité ne comprenant que la dimension Activité.

Le cinquième facteur est un facteur du Néo PI incluant C et N-.

Le sixième facteur est un facteur d'Individualisme, incluant Individualisme et Convergence Adjectifs Verbes.

Coordonnées factorielles après rotation Varimax : Tous Falsificateurs N =277

	D1	D2	D3	D4	D5	D6	D7
S_DSG	0,961	0,149	-0,038	0,079	0,077	-0,038	0,03
S_CONS vs INNOV	-0,045	-0,979	0,010	-0,041	0,020	0,089	-0,07
S_AMA vs TENS	-0,058	0,144	0,914	-0,225	-0,063	-0,021	0,12
S_ACTI vs IMAGI	-0,079	-0,049	-0,157	0,975	0,011	0,013	-0,04
S_IND vs COLL	-0,285	-0,130	-0,011	0,024	-0,053	0,826	-0,08
S_IF	0,789	0,220	-0,422	0,113	0,087	-0,030	0,03
S_CONVAP	0,269	0,086	0,543	-0,082	0,262	0,539	0,02
S_bigo	-0,058	0,801	0,124	-0,543	-0,060	0,093	0,07
S_bigc	0,525	-0,562	-0,176	0,574	0,073	0,018	-0,07
S_bige	0,529	0,593	0,243	0,414	-0,005	-0,308	0,07
S_biga	-0,768	-0,046	0,556	0,202	-0,110	-0,036	0,02
S_bigs	0,848	-0,111	0,191	-0,281	0,035	-0,285	0,03
PROSE_S	0,869	-0,197	0,217	0,107	-0,047	-0,030	-0,00
neo-o	0,170	0,135	-0,120	-0,247	-0,382	0,296	0,54
neo-c	0,002	0,016	-0,037	-0,003	0,796	0,203	0,00
neo-e	0,159	0,165	0,016	-0,004	0,170	-0,203	0,67
neo-a	-0,164	-0,055	0,144	0,010	0,092	0,011	0,74
neo-n	-0,127	0,067	-0,001	-0,023	-0,697	0,207	-0,17

Le premier facteur est un facteur PROSE de Falsification /Désirabilité.

Le second facteur est un facteur d'Ouverture, n'incluant aucune écleees du Néo PI

Le troisième facteur est un facteur d4amabilité associé à IF négatif et CONVAP

Le quatrième est un facteur d'Activité, de C , de E et de O-

Le cinquième est un facteur du Neo PI (sincère) de C et N-

Le sixième est un facteur d'Individualisme associé à Convap

Le septième est un facteur du Néo PI (sincère) incluant O, E et A.

Coordonnées factorielles après rotation Varimax : TOUS SINCERES

	D1	D2	D3	D4	D5	D6
S_DSG	0,651	0,546	-0,453	0,050	0,129	0,062
S_CONS vs INNOV	0,200	-0,641	0,082	0,663	0,180	-0,110
S_AMA vs TENS	0,361	0,152	0,821	-0,174	-0,210	0,011
S_ACTI vs IMAGI	-0,433	0,558	0,073	0,620	0,199	0,035
S_IND vs COLL	-0,126	-0,331	-0,030	0,052	0,729	0,012
S_IF	-0,095	0,078	-0,260	0,033	-0,169	0,817
S_CONVAP	0,105	-0,052	0,245	-0,354	0,305	0,584
S_bigo	-0,090	0,186	-0,091	-0,922	-0,145	0,071
S_bigc	0,405	0,115	-0,155	0,803	0,334	-0,026
S_bige	0,280	0,901	0,097	-0,084	-0,132	0,077
S_biga	-0,258	-0,022	0,886	0,176	-0,124	-0,109

S_bigs	0,939	0,126	-0,101	0,123	-0,030	-0,022
PROSE_S	0,825	0,210	0,063	0,289	-0,092	0,268
neo-o	-0,034	0,190	0,024	-0,663	0,293	0,013
neo-c	0,241	0,216	-0,068	0,493	0,564	-0,133
neo-e	0,173	0,781	0,205	-0,118	-0,075	-0,048
neo-a	-0,019	0,132	0,702	-0,014	0,175	-0,033
neo-n	-0,761	-0,003	-0,108	-0,016	-0,050	0,190

Le premier facteur est un facteur de Positivité incluant DSG, IMAG, C, S, Prose et l'échelle N- du Néo PI.

Le second facteur est un facteur d'Extraversion incluant DSG, Innov, Acti, E et l'échelle E du Néo Pi.

Le troisième facteur est un facteur d'Amabilité incluant DSG-, AMA, A et l'échelle A du Néo PI

Le quatrième facteur est un facteur de Conscience opposé à Ouverture incluant CONS, Acti, O-, C ainsi que les échelles O- et C du Néo PI

Le cinquième facteur est un d'Autonomie regroupant IND et l'échelle C du Néo PI

Le sixième facteur est un facteur regroupant IF et CONVAP.

CONCLUSIONS : Ces analyses factorielles illustrent clairement le problème de la falsification. En situation « sincère » la structure des résultats est claire : elle regroupe les échelles de Socrate et du Néo PI conformément à la théorie des Big Fives, le sixième facteur étant un facteur de « contrôle » unissant IF et CONVAP. Par contre en situation « falsification » la structure des résultats met en évidence des facteurs séparés pour SOCRATE et pour le Néo PI, découlant de l'invalidité des dimensions de SOCRATE falsifié vis-à-vis du Néo PI, même si pour SOCRATE on retrouve une structure analogue à celle « sincère », démontrant que la falsification n'est pas « aléatoire » mais bien structurée d'une manière analogue aux réponses sincères, ce qui en rend la détection plus difficile.

ANALYSES EN COMPOSANTES PRINCIPALES DES INDICATEURS DE LA FALSIFICATION

Coordonnées factorielles : Données sincères (S) et falsifiées" (F); N = 277

	F1	F2
S_COTOT	0,000	0,016
S_MOY	0,468	-0,328
S_DSG	0,693	-0,510
S_IF	0,015	0,024
S_CONVAP	0,102	0,178
S_bigc	0,436	-0,600
S_bigs	0,683	-0,583
PROSE_S	0,598	-0,537

F_COTOT	0,446	0,286
F_MOY	0,456	-0,066
F_DSG	0,841	0,423
F_IF	0,635	0,338
F_CONVAP	0,324	0,129
F_bigc	0,415	0,334
F_bigs	0,770	0,358
PROSE_F	0,706	0,548
CONV NEOS	-0,134	0,200
CONV NEOF	-0,416	0,031

Les deux premiers facteurs sont très intéressants.

Le premier facteur est un facteur de Désirabilité des données sincères ET falsifiées. Il regroupe Prose sincère et Prose Falsifié, DSG sincère et DSG falsifiés, Moyenne sincère et Moyenne falsifié, et aussi des indices qui sont propres à la situation falsifié : Cotot, IF, et Convergence des échelles OCEAN de Socrate et celles du Néo PI, négativement.....

Ce facteur identifie donc des sujets dont le profil aussi bien « sincère » que « falsifié » est désirable.....ce sont les perles rares que nous avons essayé d'identifier grâce à PROSEC appliqué au groupe de sujets ayant PROSE supérieur à 60.

Le second facteur oppose les indicateurs de falsification dans les données falsifiées à savoir DSG, PROSE et d'une manière plus faible IF, aux indicateurs de falsification dans les données sincères : DSG, C, S et PROSE, en d'autre termes ce sont les falsificateurs « ordinaires » qui ayant un profil sincère peu « désirable » vont le transformer en profil « désirable ». Pour ces sujets, les méthodes de détection PROSE, DSG et aussi IF sont efficaces.

PREDICTION DE LA FALSIFICATION.

Dans un premier temps nous avons étudié la possibilité d'ajouter à PROSE, ou probabilité de falsification, un second indice de probabilité conditionnelle, permettant d'affiner la détection des falsificateurs et des sincères à l'intérieur de groupe déterminés par des niveaux de PROSE et de voir l'incidence de cette détection sur la validité des protocoles.

Nous avons défini 5 groupes de sujets en fonction de PROSE : ceux dont PROSE est inférieur à 25 et qui peuvent être déclarés « sincères » avec une grande certitude, ceux dont PROSE est compris entre 25 et 39 et qui contiennent environ 12% de falsificateurs que l'on peut qualifier de modeste et qui se différencient des sincères surtout par des scores plus élevés à l'échelle AMA. Prosec dans ces conditions est peu efficace et n'améliore que marginalement la prédiction et la validité. On peut certes l'expérimenter mais cela n'améliore guère la décision de considérer tous les sujets de ce groupe comme sincères. Ensuite le groupe « moyen » dont PROSE est supérieur à 40 et inférieur à 60, et qui comprend un nombre à peu près égal de sincères et de falsificateurs. Ces derniers se différencient des sincères comme pour le groupe précédent par une dimension AMA plus élevée (mais beaucoup plus fortement) ainsi qu'une dimension DSG plus élevée. Prosec est alors efficace et permet d'améliorer la validité des sujets détectés « sincères », par rapport à la règle consistant à considérer ceux dont PROSE est inférieur à 50 comme « sincère ». Dans ces conditions nous recommandons d'expérimenter PROSEC en pratique.

Le groupe suivant est le groupe des sujets ayant PROSE supérieur à 75 et qui ne comprend pratiquement que des falsificateurs. Les sujets de ce groupe peuvent être déclarés « falsificateurs » avec une grande certitude et donc invalider leurs protocoles. Le dernier groupe est le plus intéressant, celui des sujets ayant PROSE compris entre 60 et 74, et qui contiennent environ 20% de sincères, les perles rares si recherchées. Ces « perles » se caractérisent par des dimensions CONS et AMA plus faibles que les falsificateurs. Leur détection est moyennement efficace puisqu'elle permet d'en identifier la moitié mais avec une certitude de 50%, puisqu'ils sont mélangés à environ un même nombre de falsificateurs. Il appartient au psychologue de démêler le vrai du faux. Nous pensons donc utile d'expérimenter PROSEC en pratique.

FIN

Ces recherches ont été effectuées entre 2010 et 2017, en collaboration avec Roland CAPEL et sous sa direction, avec la participation de Cinzia Zanetti, , Cynthia Vaudroz, Emmanuele Meier et Marc Dupuis.