

THE STACKING EFFECT AND METHODS OF CORRECTING FOR IT

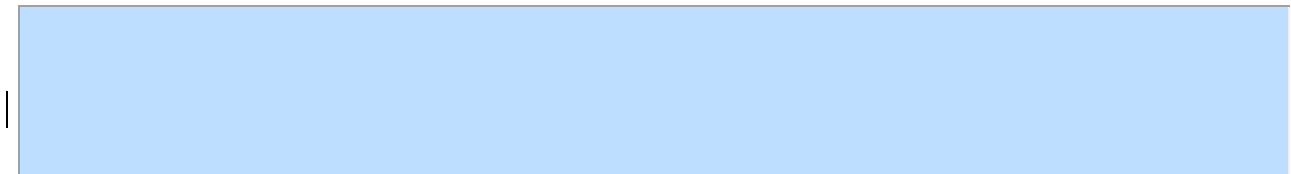
By ROBERT H. THOULESS AND ROBERT M. BRIER^{1 [1]}

Journal of Parapsychologie, 34 – 2 June 1970, pp. 126-128

traduction française par Dr Bernard Auriol :

Le Stacking-effet et les méthodes pour le corriger

Abstract: Occasionally it is practicable to use an ESP test in which all the subjects make their calls for the same set of targets instead of for individual sets. When a single set is used for the group, it is necessary to make a statistical correction to compensate for the possibility that the subjects may share the same calling bias and that this may coincide with (or deviate from) the target order and give spuriously high (or low) results (the "stacking effect"). One of the methods of correction was worked out by T. N. E. Greville. Other experimenters have allowed for the stacking effect by merely reducing their critical ratios by 10%, a simplification which lacks theoretical justification. Two other methods pointed out by the authors to evaluate these multiple-calling experiments are the majority vote technique and the use of the "index of preference."-Ed.



Résumé: Il est parfois pratique d'utiliser un test ESP dans lequel tous les sujets font leurs réponses pour un même ensemble de cibles au lieu d'ensembles individuels. Quand un ensemble unique est employé pour le groupe, il est nécessaire de faire une correction statistique afin de compenser la possibilité que les sujets puissent partager un même biais de réponse et que ceci puisse coïncider avec (ou dévier de) l'ordre des cible et donner faussement des résultats élevés ou diminués (par un effet d'*empilement* : « [stacking effect](#) »). Un des méthodes de correction a été établie par [T. N. E. Greville](#). D'autres expérimentateurs ont tenu compte de l'effet d'empilement en réduisant simplement leurs raisons critiques de 10%, simplification qui manque de justification théorique. Deux autres méthodes sont proposées par les auteurs pour évaluer ces expériences à réponses-multiples : la technique de vote majoritaire et l'utilisation du l' « index de préférence ».

<p>From time to time, experiments are carried out which involve what is called the "stacking effect." In such experiments, it is necessary to compensate for this effect by making a statistical correction in the evaluation of the results. It is the purpose of this paper to describe the stacking effect and to discuss the various methods of correcting for it.</p>	<p>De temps en temps, on effectue des expériences qui impliquent ce qu'on a appelé un « stacking effect » (effet d'empilement). Dans de telles expériences, il est nécessaire de compenser cet effet en faisant une correction statistique dans l'évaluation des résultats. Le but de cet article est de décrire ce « stacking effect » et de discuter les diverses méthodes pour le corriger.</p>
<p>The most efficient design for an ESP test is usually one in which each subject has an individual set of targets. However, there may sometimes be justifiable reasons for adopting the design of an experiment involving a number of subjects guessing at the same target order. For instance, in testing a class of young students, the experimenter may judge that a single target order for all</p>	<p>La conception la plus appropriée pour un test ESP est habituellement celle dans laquelle chaque sujet a un ensemble individuel de cibles. Cependant, il peut parfois y avoir des raisons justifiables d'adopter la conception d'une expérience impliquant un certain nombre de sujets devinant une même succession de cibles. Par exemple, en examinant une classe de jeunes étudiants, l'expérimentateur peut</p>

<p>subjects is better motivationally because the experimenter can provide duplicate record sheets, and, when the guesses are completed, collect the carbon copies, after which he can read out the targets so that the subject can make an immediate check-up. Another situation in which the single target order is likely to be preferred is when very large numbers of subjects are involved in an experiment.</p>	<p>juger qu'un seul ordre des cible pour tous les sujets est plus motivant parce que l'expérimentateur peut fournir des feuilles avec double carbone pour noter les résultats, et, quand les tests sont terminés, il rassemble les copies carbonées, après quoi il peut donner lecture des cibles de sorte que chaque sujet puisse faire un contrôle immédiat. Une autre situation dans laquelle un ordre unique des cibles sera préféré concerne les expérience impliquant très un grand nombre de sujets.</p>
<p>In a group ESP test in which all subjects have the same target order, the problem of assessing the significance of an observed deviation from mean chance expectation in the scores is more difficult than in an ESP test situation in which each subject has a different target order, but the advantages in economizing on experimental time may be held to outweigh the disadvantages of increased statistical complexity in evaluation.</p>	<p>Dans un test ESP de groupe, dans lequel tous les sujets ont le même ordre des cibles, le problème d'évaluer la signification d'une différence observée par rapport à l'espérance mathématique dans les scores, est plus difficile que dans un test ESP dans lequel chaque sujet a un ordre différent des cibles, mais l'avantages en économie de temps expérimental peut être tenu pour supérieur à l'inconvénient que représente une complexité statistique accrue.</p>
<p>The ordinary methods used for evaluating results from experiments in which there are individual target orders are relatively easy to carry out, since they depend on the assumption that the result of one run (of a 25-card ESP deck, for instance) is independent of the result of another run. This cannot be assumed, however, when a number of runs of guesses are made for the same target order, for the lack of independence between the results of the runs made by different subjects might produce a stacking effect; that is, a spurious</p>	<p>Les méthodes ordinaires employées pour évaluer les résultats à partir d'expériences dans lesquelles il y a un ordre individuel des cibles, sont relativement facile à mettre en oeuvre, puisqu'elles dépendent de l'hypothèse que le résultat d'une salve (sur un paquet de 25-cartes, par exemple) est indépendant du résultat d'une autre salve. Ceci ne peut pas être garanti, cependant, quand un certain nombre de salves de réponses sont faites pour la même succession de cibles, parce que le manque d'indépendance entre les</p>

<p>appearance of psi-hitting (or of psi-missing). It might for instance, be the case that an "open" target order (i.e., one in which the frequency of targets in a 25trial run is randomly determined) might contain more than five of a target which also happened to be a favourite guess. Suppose, for example, that the run contained eight stars and that there was a general tendency for subjects to guess star more often than the other symbols. This would probably result in more hits per run than the mean chance expectation of 5, and this spurious excess over mean chance expectation would, by ordinary methods of calculation, appear more significant as the number of subjects was increased. A spurious deficiency of hits could, of course, equally well be obtained if there were a general tendency among subjects to guess star less frequently than other symbols.</p>	<p>résultats des salves dues à différents sujets pourrait produire un <u>effet d'empilement</u>; c'est-à-dire, un faux aspect de psi-hitting (ou de psi-missing). Il se pourrait par exemple, qu'une succession de cibles « ouvert » (avec remise : c.-à-d., un dans lequel la fréquence des cibles dans une salve de 25 essais est déterminée aléatoirement) pourrait contenir plus de cinq fois comme cible une carte qui serait également être une réponse favorite. Supposons, par exemple, que la salve contienne huit étoiles et qu'il y avait une tendance générale pour les sujets de choisir « étoile » comme réponse plus souvent que les autres symboles. Ceci aurait probablement comme conséquence plus de succès par salve que l'espérance moyenne qui est de 5 chance sur 25, et cette surestimation, par les méthodes de calcul ordinaires, semblera d'autant plus significative que le nombre de sujets augmentera. Une fausse insuffisance des succès pourrait, naturellement, également bien être obtenue s'il y avait une tendance générale parmi les sujets à deviner « étoile » moins fréquemment que les autres symboles.</p>
<p>It is not sufficient to avoid this difficulty by substituting a "closed" deck (with five of each kind of target) for the open one. Difficulty would then arise from the possibility that there would be a correspondence between the serial order of the targets and preferred serial orders in the subjects' guessing. For example,</p>	<p>Il n'est pas suffisant d'éviter cette difficulté en substituant un paquet "fermé" (comportant cinq fois chaque cible) au paquet « ouvert » (tirage aléatoire avec remise). La difficulté résulterait alors de la possibilité qu'il y ait une correspondance entre l'ordre sériel des cibles et des ordres sériels</p>

<p>suppose that the first two cards in the pack happened to be a square and a circle and that the sequence of square-circle was more commonly chosen for the first two guesses than any of the 24 other pairs that might be chosen. Again, if such preferred sequences, here or elsewhere in the deck, coincided with actual sequences in the target cards, the result would be a spurious tendency for the subjects to score an excess of hits, which might even be magnified to produce an illusory appearance of significance if enough subjects agreed in their preferred orders of guessing. Also, of course, preferences in guessing sequence may produce an equally illusory appearance of psi-missing.</p>	<p>préférés par les sujets devinant. Par exemple, supposez que les deux premières cartes dans le paquet se soient avérées être un carré et un cercle et que la séquence carré-cercle ait été généralement choisie pour les deux premières conjectures que l'une quelconque des 24 autres paires qui pourraient être choisies. De plus, si de telles séquences préférés, ici ou ailleurs dans le paquet, coïncidaient avec des séquences réelles dans les cartes cibles, le résultat présenterait une fausse tendance à ce que les sujets semblent avoir un excès de tirs au but, qui pourraient même être magnifiés jusqu'à produire une signification statistique illusoire si assez de sujets convergeaient quant à leurs ordres préférés de divination. En outre, naturellement, les préférences pour des séquences de divination peuvent produire un aspect également illusoire de psi-missing.</p>
<p>What is needed to avoid this difficulty is to use a method of evaluating the significance of a set of runs by different subjects on the same target order which is free from the possibility that accidental correspondences between target frequencies and preferred calls or between target sequences and preferred guessing-sequences can produce spurious indications of psi-hitting (or of psi-missing).</p>	<p>Pour éviter cette difficulté on doit employer une méthode d'évaluation de la signification d'un ensemble de salves par différents sujets sur une même séquence de cibles qui soit exempte de la possibilité que des correspondances accidentelles entre fréquence des cibles et réponses préférées ou entre séquences des cibles et séquences préférées de réponses puissent produire de fausses indications de Psi-Hitting (ou de Psi-Missing).</p>
<p>The standard way of doing this is by the</p>	<p>La manière standard d'y parvenir</p>

<p>use of the "Greville correction" of variance for multiple calling developed by T. N. E. Greville and published by him in the <i>Annals of Mathematical Statistics</i>^{2[2]} (2). Since this publication is not readily available to parapsychologists, Greville's method was presented in simplified form by J. G. Pratt in the <i>Journal of Parapsychologie</i>^{3[3]} (3). It is this form that has generally been used in the past by experimental parapsychologists to deal with the problem of evaluation of results obtained by multiple calling on one deck.</p>	<p>consiste à utiliser la "correction de Greville" de la variance pour réponses multiples. Cette méthode a été développée par T. N. E. Greville et publiée dans les <i>Annals of Mathematical Statistics</i>^{4[4]}. Mais comme cette publication n'est pas aisément disponible aux parapsychologues, la méthode de Greville a été présentée sous une forme simplifiée par J. G. Pratt dans le journal de Parapsychologie (3). C'est cette forme qui a été généralement employée dans le passé par les parapsychologues expérimentaux pour traiter le problème de l'évaluation des résultats obtenus par les réponses multiples sur un seul paquet de cibles possibles.</p>
<p>Greville's method involves tabulating the calls made on each target by all the subjects. From this tabulation one can then compute the variance for the calls made on each of the targets. The computation is somewhat different for the open deck and the closed deck; the method for each is given in Pratt's article. For the open deck, when the 25 variances of calls for each card in the target run have been calculated, these are added to obtain the variance for the complete run. The square root of this variance is then taken as the standard</p>	<p>La méthode de Greville implique de tabuler les réponses faites sur chaque cible par tous les sujets. De cette tabulation on peut alors calculer la variance pour les réponses faites sur chacune des cibles. Le calcul est quelque peu différent pour le « paquet ouvert » et le « paquet fermé »; les méthodes respectives sont données dans l'article de Pratt. Pour le « paquet ouvert », quand les 25 variances des appels pour chaque carte dans la succession des cibles ont été calculées, celles-ci sont ajoutées pour obtenir la variance globale</p>

<p>deviation to be used in calculating the critical ratio. For the closed deck, the variance for the entire set of 25 targets as the unit must be calculated. Then the square root of this is taken as the standard deviation to be used.</p>	<p>de la salve complète. La racine carrée de cette variance est alors prise comme écart type à employer dans le calcul de la « raison critique ». Pour le paquet fermé, la variance pour l'ensemble entier de 25 cibles comme unité doit être calculé. Alors la racine carrée de ceci est prise comme écart type à employer.</p>
<p>This procedure normally gives a higher variance than the commonly used theoretical variance appropriate for experiments in which only one run of guesses has been made for each target order. Therefore it has the effect of reducing the significance of an observed deviation in either direction. This, of course, is to be expected, since a likely effect of multiple calling is to magnify accidental correspondence or lack of correspondence between target and call.</p>	<p>Ce procédé donne normalement une variance plus élevée que la variance théorique généralement utilisé, appropriée pour les expériences dans lesquelles une seule série de réponses a été faite pour chaque suite de cibles. Par conséquent il a l'effet de réduire la signification d'une déviation observée dans l'une ou l'autre direction. Naturellement, on devait s'y attendre, puisqu'un effet probable de réponses multiples est de magnifier une correspondance accidentelle ou un manque de correspondance entre la cible et la réponse.</p>
<p>The tabulation required for the application of the Greville correction is laborious, and some experimenters have preferred to use simpler methods of arriving at the same result. The simplest was one suggested by Van Busschbach^{5[5]} (5). In some of his earlier work, he had applied the Greville correction and had found that the effect was to reduce the critical ratios by less</p>	<p>La tabulation exigée pour l'application de la correction de Greville est laborieuse, et quelques expérimentateurs ont préféré employer des méthodes plus simples d'arrivée au même résultat. Le plus simple était un suggéré en Van Busschbach (5). Dans certains de ses premiers travaux, il avait appliqué la correction de Greville et avait constaté que l'effet devait réduire les rapports</p>

<p>than 10%. This led him to suggest that a 10% reduction of the critical ratios could be used to avoid the complicated calculations required by the Greville method. This simplification lacks theoretical justification; the correction required is not constant for all card distributions or for all numbers of guessing subjects. Osis^{6[6]} has discussed cases in which the Greville correction has reduced CRs by considerably more than 10%. Estimates of significance based on this 10% reduction, such those made by Van Busschbach^{7[7]} (5) and Brier^{8[8]} (1), must be considered to rest on somewhat dubious foundations.</p>	<p>critiques de moins de 10%. Ceci l'a mené à suggérer qu'une réduction de 10% des rapports critiques pourrait être employée pour éviter les calculs compliqués exigés par la méthode de Greville. Cette simplification manque de justification théorique; la correction exigée n'est pas constante pour toutes les distributions de carte ou pour tous les nombres de sujets percipients. Osis^{9[9]} a discuté les cas dans lesquels la correction de Greville a réduit les régions critiques de nettement davantage que 10%. Les estimations de signification basées sur cette réduction de 10%, tels ceux de Van Busschbach^{10[10]} (5) et Brier^{11[11]} (1), doivent être considérées comme reposant sur des bases plutôt douteuses</p>
<p>The simplest and most direct alternative to Greville's method would seem to, be the repeated-guessing, or the majority-vote, technique in which each subject's guess is treated as a vote for the symbol on the corresponding target card^{12[12]} (4).</p>	<p>L'alternative la plus simple et la plus directe à la méthode de Greville semblerait être la présentation réitérée de la même cible, ou le vote majoritaire, technique dans laquelle la conjecture de chaque sujet est traitée comme un vote</p>

<p>The whole system of guesses by all subjects is thus reduced to a single run of 25 guesses. One way of doing this is to take as the final guess on any target card that target symbol which has been guessed by the largest number of experimental subjects. This process gives the experimenter a single final order for the guesses, which he can then compare with the order of the target deck.</p>	<p>pour un symbole qui fait partie des cibles possibles^{13[13]} (4). Le système entier des conjectures par tous les sujets est ainsi réduit à une seule série de 25 conjectures. Une façon de faire est de prendre comme conjecture finale sur n'importe quelle cible, le symbole qui a été choisi par le plus grand nombre des sujets d'expérience. Ce processus donne à l'expérimentateur une seule série finale pour les conjectures, qu'il peut alors comparer à l'ordre du paquet cible.</p>
<p>A lengthier but more sensitive method of getting this final guess order is by the use of the "index of preference"^{14[14]} (4). In this method the votes are weighted in a manner determined by the frequencies of different guesses and of different targets. If a subject or group of subjects, has a bias for or against calling a particular symbol, this is taken into account when determining which symbol is to be considered the final guess for the group; for example, if one symbol is hardly ever called by the subject and if the subject breaks his habit and makes many calls for this symbol, it might be reasonable to count this symbol as the final guess for the target, even though another symbol might have been called slightly more</p>	<p>Une méthode plus longue mais plus sensible d'obtenir cet ordre final des réponses est l'utilisation d'un "index des préférences"^{15[15]} (4). Dans cette méthode les votes sont pondérés d'après la fréquences des différentes réponses et des différentes cibles. Si un sujet ou un groupe de sujets montre une polarisation pour ou contre le choix d'un symbole particulier, cela est pris en considération pour déterminer quel symbole doit être considéré comme réponse finale pour le groupe; par exemple, si un symbole n'est presque jamais choisi par le sujet et si le sujet contre son habitude choisit ce symbole à plusieurs reprises, il pourrait être raisonnable de compter ce symbole comme réponse finale pour la cible, même si un symbole différent était</p>

often.	choisi légèrement plus souvent.
<p>By this method, the preferred guess for each target card is determined by the use of the formula n_{xy}/g_y where n_{xy} is the number of times a certain target card x has been guessed as bearing the symbol y, and g_y is the total number of guesses of y over all target cards. The preferred guess, then, is that one of the guess symbols which gives the maximum value of this index for each particular target card.</p>	<p>Par cette méthode, le choix préféré pour chaque carte cible est déterminé par l'utilisation de la formule n_{xy}/g_y où n_{xy} est le nombre de fois où une certaine carte cible x a été devinée comme portant le symbole y, et g_y est le nombre total de réponses de y sur toutes les cartes cibles. La réponse préférée, alors, est celui des symboles choisis qui donne la valeur maximale de cet index pour chaque carte cible particulière.</p>
<p>Whichever of the alternatives to the Greville method is used, the result is that, if ESP is affecting a considerable number of the subjects' guesses even to a small extent, the final guess order so obtained will correspond more closely with the target deck than will an individual run of guesses. The evaluation, either by the majority vote method or the index of preference, is very simple, based as it is on a comparison between a single target order and a single set of guesses. It gives an estimate of the significance of the correspondence between target order and subjects' guesses which is entirely free from any effect of magnification by multiple guessing due either to commonly preferred targets or commonly preferred sequences. It achieves the end of the Greville correction (for either an open or closed deck situation) in a simpler and somewhat less laborious way.</p>	<p>Quelle que soit la solution de rechange à la méthode de Greville employée, le résultat est que, si l'ESP affecte un nombre considérable des réponses des sujets même un faible degré, l'ordre final des réponses ainsi obtenu correspondra plus étroitement au paquet cible que ne le fera la série correspondante des réponses d'un individu. L'évaluation, soit par la méthode du vote majoritaire ou par l'index de préférence, est très simple, car basé sur une comparaison entre une séquence unique des cibles et une séquence unique de réponses. Elle donne une évaluation de l'importance de la correspondance entre la succession des cibles et les réponses des sujets, qui est entièrement exempte d'une forme quelconque d'exagération par un protocole de réponses multiples qui serait due soit à des préférences collectives pour certaines cibles soit à des préférences collectives pour certaines séquences de cibles. Cela</p>

	atteint le but de la correction de Greville (pour une situation en paquet ouvert ou fermé) d'une manière plus simple et légèrement moins laborieuse.
<p>One obvious suggestion that may follow from the above discussion is that experiments should be designed, when at all possible, so that an individual set of targets is presented for each subject. This would eliminate the need for considerations such as those discussed in this paper.</p>	<p>Une suggestion évidente qui peut suivre de la discussion ci-dessus est que des expériences devraient être conçues, pour autant que ce soit possible, de sorte qu'un ensemble individuel de cibles soit présenté à chaque sujet. Ceci éliminerait la nécessité des considérations présentées dans cet article.</p>
<p>REFERENCES</p> <ol style="list-style-type: none"> 1. 1. BRIER, R. M. A correspondence ESP experiment with high-I.Q. subjects. <i>J. Parapsychol.</i>, 1967, 31, 143-48. 2. 2. GREVILLE, T. N. E. <u>On multiple matching with one variance deck.</u> <i>Annals of Mathematical Statistics</i>, 1944, 15, 432-34 3. 3. PRATT, J. G. <u>The variance for multiple-calling ESP data.</u> <i>J. Parapsychol.</i>, 1954, 18, 37-40. 4. 4. THOULESS, R. H. <u>The repeated-guessing technique.</u> <i>Int. J. Parapsychol.</i>, 1960, 2 (No. 3), 21-36. 5. 5. VAN BUSSCHBACH, J. G. A further report on an investigation of ESP in school children. <i>J. Parapsychol.</i>, 1955, 19, 73-81 	
<p>2 Leys Rd. Cambridge CB4 2AU England</p>	<p><i>Institute for Parapsychologie College Station Durham, N. C. 27708</i></p>

chercher sur ce site



© Copyright Bernard AURIOL (email : auriol@free.fr)

dernière mise à jour le

15 Janvier 2004