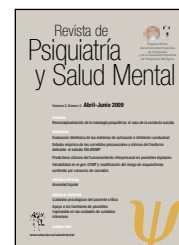


# Revista de Psiquiatría y Salud Mental

www.elsevier.es/saludmental



## ORIGINAL

## Evaluation of behavioural activation and inhibition systems by telephone

Alejandra Pinto-Meza,<sup>a,b,\*</sup> David Suárez,<sup>a</sup> Xavier Caseras,<sup>b</sup> Josep Maria Haro,<sup>a</sup> Antoni Serrano-Blanco<sup>a</sup> and Rafael Torrubia<sup>b</sup>

<sup>a</sup>Sant Joan de Déu Foundation, Research and Development Unit, Sant Joan de Déu-SSM, Barcelona, Spain

<sup>b</sup>Medical Psychology Unit and Neuroscience Institute, Autonomous University of Barcelona, Barcelona, Spain

Received on 13 November 2008; accepted on 20 February 2009

### KEYWORDS

Telephone assessment;  
Behavioral activation  
system;  
Behavioral inhibition  
system;  
SPSRQ

### Abstract

**Introduction and objectives:** The activity of behavioral inhibition and activation systems (BIS and BAS) has been related to several mental disorders. The availability of a validated questionnaire to assess BIS and BAS over the telephone could aid research. The objective of the present study was to evaluate the procedural validity of telephone administration of the sensitivity to punishment/ sensitivity to reward questionnaire (SPSRQ) as a measure of BIS and BAS activity (by means of its two scales: the sensitivity to punishment [SP] scale and the sensitivity to reward [SR] scale).

**Methods:** A total of 231 participants were assessed with the SPSRQ twice (over the telephone and in a self-administered format). Intraclass correlation coefficients and kappa indices were calculated. Additionally, possible differences between the mean scores obtained with the two assessment procedures and internal consistency were explored.

**Results:** Telephone and self-assessment of BIS and BAS, by means of the SPSRQ, provided similar results, supporting the procedural validity of telephone administration. A slight but statistically significant tendency (estimated betha = 0.62; 95% CI, 0.33-0.92;  $p < 0.001$ ) toward lower scores was observed for the telephone-administered SP scale and for the interaction "first or second assessment/telephone or self-administered SPSRQ" for the SR scale ( $p = 0.023$ ).

**Conclusions:** The results of the present study support the reliability of telephone assessment of BIS and BAS by means of the SPSRQ.

© 2008 Sociedad Española de Psiquiatría y Sociedad Española de Psiquiatría Biológica. Published by Elsevier España, S.L. All rights reserved.

\*Corresponding author.

E-mail: apinto@sjd-ssm.com (A. Pinto-Meza).

**PALABRAS CLAVE**

Evaluación telefónica;  
Sistema de activación  
conductual;  
Sistema de inhibición  
conductual;  
CSCSR

**Evaluación telefónica de los sistemas de activación e inhibición conductual****Resumen**

**Introducción y objetivos:** La actividad de los sistemas de activación e inhibición conductual (SAC y SIC) se ha relacionado con diversos trastornos mentales. Tener un instrumento que permita evaluar telefónicamente la actividad del SAC y el SIC podría facilitar la investigación. El objetivo del presente estudio fue evaluar la validez del procedimiento de administración telefónica del cuestionario sensibilidad al castigo y sensibilidad a la recompensa (CSCSR) como medida de la actividad del SAC y el SIC, mediante sus dos escalas: la escala sensibilidad al castigo (SC) y la escala sensibilidad a la recompensa (SR). **Métodos:** Un total de 231 sujetos fueron evaluados en dos oportunidades con el CSCSR (una vez por teléfono y otra en formato autoadministrado). Se calcularon el coeficiente de correlación intraclass y el índice kappa. Además, se exploraron posibles diferencias entre la media de las puntuaciones obtenidas en ambas modalidades y la consistencia interna.

**Resultados:** Los resultados permiten establecer que la evaluación telefónica y en formato autoadministrado del SIC y el SAC, mediante el CSCSR, ofrece resultados similares, lo que avala una adecuada validez del procedimiento de la administración telefónica del cuestionario. Hay una tendencia pequeña pero estadísticamente significativa (beta estimada = 0,62; intervalo de confianza del 95%, 0,33-0,92;  $p < 0,001$ ) a puntuaciones menores para la administración telefónica de la escala SC y para la interacción "primera o segunda evaluación/CSCSR aplicado por teléfono o en formato autoadministrado" para la escala SR ( $p = 0,023$ ).

**Conclusiones:** Los resultados de este estudio avalan la fiabilidad de la evaluación telefónica del SIC y el SAC mediante el CSCSR.

© 2008 Sociedad Española de Psiquiatría y Sociedad Española de Psiquiatría Biológica. Publicado por Elsevier España, S.L. Todos los derechos reservados.

**Introduction**

From a neurobehavioural perspective, personality traits arise from individual differences in the activity of certain brain systems or circuits. According to Gray,<sup>1</sup> the behavioural inhibition system (BIS) and the behavioural activation system (BAS) are two (out of a total of three) general systems, which underlie two orthogonal personality dimensions. The third system proposed by the author, the fight/flight system, mediates emotions such as anger or panic. It has been proposed that different constellations of BIS and BAS activity might be related to various mental disorders. For example, it has been suggested that BIS hyperactivity is related to anxiety disorders,<sup>1</sup> BIS hypoactivity to psychopathy<sup>2,3</sup>; BIS hypoactivity and BAS hyperactivity to problems resulting from substance abuse, and histrionic and anti-social personality disorders,<sup>4</sup> and BAS hypoactivity and BIS hyperactivity to depression.<sup>5-7</sup>

Traditionally, BAS and BIS activity have been evaluated using self-administered questionnaires. However, when epidemiological studies or studies involving large samples distributed over extensive geographical areas are conducted, the administration of the instruments which are used by telephone is a great advantage, in particular because it reduces geographical barriers, which facilitates contact with interview candidates. In fact, several structured interviews and symptom scales have shown the validity of their procedures to be adequate in versions administered by telephone, given that a reasonable level of similitude

has been established between face-to-face evaluations and evaluations conducted by telephone (for example, the Personality Disorder Examination,<sup>8</sup> the Structured Clinical Interview for DSM-III axis I Disorders SCID<sup>9</sup> or the Patient Health Questionnaire 9-item Depression Module PHQ-9<sup>10</sup>). The concept of procedure validity refers to the extent to which a new procedure (for example, the administration of a specific scale or questionnaire by telephone) offers similar results to those for an established procedure which is used as a criterion (for example, the self-administered application). The validity of the procedure refers exclusively to the question of the validity of the evaluation procedure and not to the validity of the instrument as such.<sup>11</sup>

The principal aim of this study was to determine the validity of the procedure for the administration of the sensitivity to punishment/ sensitivity to reward questionnaire (SPSRQ) by telephone in a primary care (PC) subject sample. The SPSRQ is a self-administered questionnaire designed to evaluate individual differences in BAS and BIS activity (a detailed description of this instrument can be consulted in other articles).<sup>12,13</sup> In addition, two other aspects were evaluated: a) the extent to which differences in the score average obtained between the two evaluation procedures might be related to the difference in the mode of administration of the instrument or to the fact that the subjects responded to the same instrument on two occasions, and b) the internal consistency of the questionnaire administered by telephone compared to the internal consistency of the questionnaire applied in a self-administered format.

## Methods

### Subjects

The present study is part of a larger project, the principal objective of which was to evaluate the cost-effectiveness of the pharmacological treatment prescribed in PC to treat depression.<sup>14</sup> With the aim of evaluating the validity of the procedure for the administration of the SPSRQ by telephone, an additional sample was recruited. The subjects were selected by two PC doctors from amongst patients seeking medical care in two PC centres in the province of Barcelona (Spain). All the subjects were informed of the purpose of the study and they signed an informed consent form. A total of 231 individuals aged from 18 to 75 years were included.

### Material

The SPSRQ<sup>12</sup> specifically measures BIS and BAS activity, evaluating individual differences with respect to sensitivity to punishment (SP) and sensitivity to reward (SR) and the behavioural patterns which result, whether they correspond to behavioural inhibition or approximation. It is a self-administered questionnaire consisting of 48 items, 24 of which evaluate SP and 24 SR. The items are scored yes/no, 1 point being assigned to each positive response. The total score for each scale can range from 0 to 24. The SP scale, specifically designed to measure individual differences in BIS activity, measures: *a*) behavioural inhibition (passive avoidance) in general situations which involve novelty or possible adverse consequences, and *b*) fear or the cognitive processes elicited by the threat of punishment or failure. The SR scale, designed to evaluate individual differences with respect to BAS activity, evaluates the tendency to respond when people are confronted with situations in which they can perform actions in order to obtain positive reinforcement. These situations include aspects like money, sexual partner, social events, power, sensation seeking, etc. Both scales are orthogonal and have shown appropriate psychometric properties.<sup>12,13,15-18</sup>

### Procedure

All the subjects responded to the SPSRQ on two occasions (once by telephone and the other using the self-administered format) within a period of 2 to 7 days (mean  $\pm$  SD, 3.38  $\pm$  1.8 days). The sample was divided into two groups, depending on the order in which the questionnaires were administered: self-administered/telephone-administered (ST) and telephone-administered/self-administered (TS). With the intention of avoiding the possible effect of knowing the results of a previous evaluation, the final scores for both administration formats were calculated when all the data had been collected.

The subjects were not randomly assigned to groups so the ST group was completed first and then the TS group. The subjects in the ST group agreed voluntarily to respond to the self-administered SPSRQ (while they were waiting to be seen at their PC centre) and to be reinterviewed by telephone later on. First of all, the subjects in the TS group were contacted by their PC doctor in order to obtain permission for the assessor to telephone them. Again the

subjects agreed to participate voluntarily. With the aim of increasing the likelihood of response to the subsequent self-administration of the questionnaire, the subjects who were selected had a visit with their PC doctor scheduled some time during the 7 days following the telephone evaluation.

As the instructions for the SPSRQ were not entirely appropriate for their administration by telephone (the original instructions were: "Please respond to each question by encircling the word yes or the word no, depending on what you think or feel. There are no good or bad answers, or trick questions. Work quickly and do not give too much thought to the exact meaning of the question"), the instructions for the telephone administration of the SPSRQ were modified as follows: "I am going to ask you several questions. For each question the possible answers are yes or no, depending on what you think or feel. There are no good or bad answers, or trick questions, so do not give too much thought to the exact meaning of the question".

The study was approved by the ethics committee of San Juan de Dios-Mental Health Services and the San Juan de Dios Foundation.

### Statistical analysis

The degree of concordance between the responses obtained in the telephone and self-administered applications was analyzed by calculating the intraclass correlation coefficient (ICC) for individual scores (a bilateral mixed-effect ANOVA model with absolute concordance measures was used). In addition, the kappa index was calculated for each SPSRQ item. It has been established that ICC and kappa values lower than 0.4 indicate a poor degree of concordance, values between 0.41 and 0.75 moderate to good concordance and values higher than 0.75 excellent concordance.<sup>19</sup> As the interval between 0.41 and 0.75 might seem rather large, it has been proposed that values higher than 0.6 indicate a good degree of concordance.<sup>20</sup>

To evaluate the extent to which differences in the score averages for the two modes of evaluation could be related to the different procedures or to the fact that subjects responded twice to the same instrument, the averages for both modes were compared by means of the Student's *t*-test for repeated samples, the size of the effect was calculated using Cohen's kappa statistic and *linear generalized estimated equation (GEE) models* were constructed for the SP and SR scales, the order of administration (in other words, whether it was the first or second evaluation) and the administration mode (telephone- or self-administered SPSRQ), and sex, age, and educational level being regarded as covariables.<sup>21,22</sup>

The internal consistency of the telephone-administered SPSRQ was evaluated by calculating Cronbach's alpha and it was compared with the internal consistency of the instrument applied in the self-administered format. Finally, the socio-demographic characteristics of the groups were compared using the  $\chi^2$  or ANOVA tests.

## Results

The socio-demographic characteristics are shown in table 1. Statistically significant differences were not observed in

distribution according to sex, marital status or occupational status. Only educational level ( $p = 0.006$ ) and age ( $p = 0.016$ ) were significantly different for the 2 groups. The subjects in the ST group had spent more years in education and were younger than the subjects in the TS group.

The ICCs between the telephone and self-administered applications of the SPSRQ were high, ranging from 0.93 to 0.91. With respect to the SP scale, the ICC values were 0.92 (confidence interval [CI] 95%, 0.88-0.95) for the ST group, 0.93 for the TS group (CI 95%, 0.9-0.95) and 0.93 (CI 95%, 0.9-0.94) for the total sample (ST group plus TS group). In the case of the SR scale, the ICCs were 0.91 (CI 95%, 0.87-0.93) for the ST group; 0.92 for the TS group (CI 95%, 0.89-0.95) and 0.92 (CI 95%, 0.89-0.94) for the total sample (ST group plus TS group).

When concordance for the telephone or self-administered SPSRQ items was analyzed, most of the items on the SP scale exhibited values ranging between 0.41 and 0.75 (20 out of 24, which represents 83.33%). Of these a total of 15 (75%) showed kappa values higher than 0.6; 4 (16.76%) items had kappa indexes above 0.75 and none of the items presented values below 0.41. With respect to the SR scale, the majority of the items (15 out of 24, which represents 62.5%), presented kappa values between 0.41 and 0.75. Of these a total of 12 (80%) showed values above 0.6; 8 (33.33%) items had kappa values higher than 0.75 and only one item (number 24 on the questionnaire) showed a kappa value of less than 0.41. However, when analyzing the kappa index, it is important to mention that this item also obtained a very low percentage of positive responses (the majority of the subjects responded no, in other words, a response of 0 points), which ranged from 1 to 2.1%.

The score average obtained on the SPSRQ, the differences in the mean scores and the size of the Cohen effect between each pair of evaluations are summarized in table 2. The means comparison showed statistically significant differences on the SP scale for the ST and TS pairs. The subjects presented slightly lower scores for the telephone application of the SP scale compared to the self-administered scale. However, the size of the Cohen effect was minimal for both pairs (0.11 and 0.09 respectively).

Since differences emerged in the SPSRQ score averages between the groups, that variables such as sex and age appear to influence scores on the SP and SR scales<sup>12</sup> and that the groups displayed differences with respect to their educational level, the GEE models for the SP and SR scales were adjusted. When the SP scale was analyzed, a slight (0.62 points) but statistically significant tendency towards higher scores was observed for the self-administered format (estimated beta value = 0.62; CI 95%, 0.33-0.92;  $p < 0.001$ ). Moreover, female sex (estimated beta value = 2.98; CI 95%, 1.36-4.6;  $p < 0.001$ ) and age (estimated beta value = 0.05; CI 95%, 0.01-0.1;  $p = 0.027$ ) were significant in the model, higher scores being obtained on the SP scale for women and older people. On the other hand, no significant effect of the “first or second evaluation/SPSRQ applied by telephone or in self-administered format” interaction was recorded. When the SR scale was analyzed, a statistically significant effect was observed for the “first or second evaluation/SPSRQ applied by telephone or in self-administered format” interaction ( $p = 0.023$ ). In other words, if we take the subjects who responded to a second administration by telephone as our point of reference, the subjects who responded to a first evaluation by telephone

**Table 1** Socio-demographic characteristics of groups

	ST (n = 116)	TS (n = 115)	Total (n = 231))
Women	69.5%	63.7%	66.7%
Age (in years)*, average $\pm$ SD (mean)	48.01 $\pm$ 16.35 (47)	53.07 $\pm$ 15.32 (57)	50.48 $\pm$ 16.02 (51)
Marital status			
Married or live-in partner	71.2%	70.9%	71%
Never married	15.3%	10.6%	13%
Divorced, separated, widow	13.5%	18.5%	16%
Occupational status			
Working	33.1%	31%	32%
Not working (on sick leave, unemployed)	15.3%	13.3%	14.3%
Housewife	28%	41.5%	34.6%
Other (student/retired)	23.6%	14.2%	19.1%
Educational level (in years)*			
0-4	27.6%	28.3%	27.9%
5-8	19%	37.2%	27.9%
9-12	27.5%	21.2%	24.5%
Over 12	25.9%	13.3%	19.7%

ST: self-administered/ telephone administered questionnaire; SD: standard deviation; TS: telephone-administered/ self-administered questionnaire.

The  $\chi^2$  or ANOVA test was used for comparisons between groups.

\* $p < 0.05$ .

**Table 2** SPSRQ score averages, mean differences and size of the cohen effect

Groups formed according to the order of the administration procedure	SP	SR
ST (n=116)		
Self-administered	10.88 ± 6.19	6.07 ± 3.4
Telephone-administered	10.16 ± 6.31	6.1 ± 3.42
Mean differences: S-T	0.72 <sup>a</sup>	-0.03
Size of the Cohen effect	0.11	0.01
TS (n=115)		
Telephone-administered	10.15 ± 5.79	7.28 ± 3.96
Self-administered	10.68 ± 6.17	7.02 ± 4.18
Mean differences: T-S	-0.53 <sup>b</sup>	0.26
Size of the Cohen effect	0.09	0.06

ST: self-administered/ telephone-administered questionnaire; SPSRQ: sensitivity to punishment/ sensitivity to reward questionnaire; SP: sensitivity to punishment; SR: sensitivity to reward; TS: telephone-administered/ self-administered questionnaire.

The data indicates the mean ± standard deviation.

<sup>a</sup>p < 0.001.

<sup>b</sup>p < 0.05.

obtained an extra 1.16 points; the individuals who responded to a first self-administered format evaluation obtained 0.03 points less and subjects who responded to a second self-administered format evaluation obtained an extra 0.09 points. In addition to this, female sex was significant in the model (estimated beta value = -2.23; CI 95%, -3.27 to -1.19; p < 0.001). In other words, women achieved lower scores than men.

Cronbach's alpha for the telephone-administered SPSRQ was 0.89 (CI 95%, 0.87-0.91) for the SP scale and 0.76 (CI 95%, 0.69-0.8) for the SR scale. Cronbach's alpha for the self-administered format SPSRQ was 0.89 (IC del 95%, 0.87-0.91) for the SP scale and 0.76 (CI 95%, 0.71-0.81) for the SR scale.

## Discussion

The results of this study establish that BIS and BAS assessment by means of the telephone and self-administered format SPSRQ offers similar results, confirming that the validity of the procedure for the administration of the questionnaire by telephone is acceptable. There is a slight but statistically significant tendency in favour of higher scores for the self-administered form of the SP scale and for the "first or second evaluation/ SPSRQ applied by telephone or in self-administered format" interaction for the SR scale. Also, the internal consistency of the telephone-applied SPSRQ is good and very similar to that of the self-administered questionnaire.

The ICCs between the telephone and self-administered applications were good, irrespective of the order of the application procedure (ST or TS). Moreover, the concordance

for each item (calculated from the kappa index) was good or excellent in the case of most of the items, with the exception of item 24 on the SR scale. This item also showed a very low percentage of positive responses (most individuals responded no, a response of 0 points). According to Grove et al<sup>23</sup> and Spitznagel et al,<sup>24</sup> in these cases the interpretation of the kappa index is problematic, given that its values depend on the prevalence of the phenomenon which is being studied. In fact, the indications are that, when prevalence is very low (< 5%), the kappa index should not be calculated. For example, given constant specificity and sensitivity values of 0.95, a kappa value can decrease from 0.81 to 0.14 if prevalence decreases from 50 to 1%. Therefore, different kappa values may be solely due to differences in prevalence. If we take all these factors into consideration, we may assume that the telephone and self-administered application of the SPSRQ produce highly concordant results, confirming that the validity of the procedure for the telephone application of the instrument is adequate.

The comparison of the average SPSRQ scores revealed a significant tendency towards higher scores for the self-administered SP scale. However, the differences were minimal (0.62 points) and probably of very little relevance for the evaluation of individual differences. Although both procedures (telephone and self-administered application) employed the same questions (items), it might be the case that answering on an individual basis (in other words in the self-administered format) might favour the recognition of certain characteristics, which could inhibit subjects when they are face to face with another person asking questions (by telephone), either as a result of wariness on the part of subjects or a lack of privacy (if the interviewee is not alone when responding). As Evans et al<sup>25</sup> pointed out: "It is more difficult to ensure privacy during a telephone interview, as the interviewer does not know who else might be present, possibly inhibiting the openness of the subject". Even though the "first or second evaluation/SPSRQ applied by telephone or in self-administered format" interaction was statistically significantly for the SR scale, we believe that a decrease of 0.03 points when answering the self-administered format first may not be relevant when evaluating individual differences.

The internal consistency of the telephone-administered SP and SR scales was good and almost identical to the consistency of the self-administered format.

We need to recognize that there are certain limitations to this study and to telephone evaluation in general. Firstly, the subjects were not randomly assigned to their groups. This could explain why the groups displayed differences with respect to certain socio-demographic characteristics. Secondly, the sample was selected from amongst people who sought medical care in PC centres and, for the most part, was composed of women, which limits the generalization of our results. Thirdly, owing to problems of confidentiality (an assessor who was not authorized to access personal data, such as the telephone number of patients), the subjects were first of all contacted by their PC doctor in order to obtain authorization for them to be interviewed by telephone. Because of this, the results of this study could have limitations if they are generalized to studies in which

the first contact with individuals is by telephone. Finally, we have to bear in mind that, when we make evaluations by telephone, we are selectively excluding individuals without a telephone (for example, individuals from rural areas) and, therefore, there is a degree of bias in our results.

## Conclusions

The results of this study confirm the reliability of telephone evaluation of BIS and BAS by means of the SPSRQ questionnaire. We believe that this will enable more studies on BIS and BAS activity and their relationship to psychopathology to be conducted, with the aim of overcoming geographical barriers and fostering the inclusion of BIS and BAS measurement in longitudinal studies with frequent follow-up evaluations or in large epidemiological studies.

## Funding

This study was funded by the Agency for Technology Evaluation and Medical Research of Catalonia (Agència Catalana de Evaluació de Tecnologies e Investigació Mèdica) (063/ 26/ 2000). David Suárez has a grant from the health institute "Instituto de Salud Carlos III", Ministry of Science and Innovation, Spain (FIS CA05/ 0177).

## References

- Gray JA. The neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system. Oxford: Oxford University Press; 1982.
- Fowles DC. The three arousal model: Implications of Gray's two factor learning theory for heart rate, electrodermal activity, and psychopathy. *Psychophysiology*. 1980;17:87-104.
- Newman JP, MacCoon DG, Vaughn LJ, Sadeh, N. Validating a distinction between primary and secondary psychopathy with measures of Gray's BIS and BAS constructs. *J Abnorm Psychol*. 2005;114:319-23.
- Taylor J, Reeves M, James L, Bobadilla L. Disinhibitory trait profile and its relation to Cluster B personality disorder features and substance use problems. *Eur J Personality*. 2006;20:271-84.
- Johnson SL, Turner RJ, Iwata N. BIS/BAS levels and psychiatric disorder: An epidemiological study. *J Psychopathol Behav Assess*. 2003;25:25-36.
- Kasch KL, Arnow BA, Gotlib IH. Behavioral activation and inhibition systems and the severity and course of depression. *J Abnorm Psychol*. 2002;111:589-97.
- Pinto-Meza A, Caseras X, Soler J, Puigdemont D, Pérez V, Torrubia R. Behavioural inhibition and behavioural activation systems in current and recovered major depression participants. *Pers Individ Dif*. 2006;40:215-26.
- Rohde P, Lewinsohn PM, Seeley JR. Comparability of telephone and face-to-face interviews in assessing axis I and II disorders. *Am J Psychiatry*. 1997;154:1593-8.
- Cacciola JS, Alterman AI, Rutherford MJ, McKay JR, May DJ. Comparability of Telephone and In-Person Structured Clinical Interview for DSM-III-R (SCID) Diagnoses. *Assessment*. 1999;6:235-42.
- Pinto-Meza A, Serrano A, Peñarubia MT, Blanco E, Haro JM. Assessing depression in primary care with the PHQ-9: Can it be carried out over the telephone? *J Gen Intern Med*. 2005;20:738-42.
- Spitzer RL, Williams JB. Clasificación de los trastornos mentales. En: Kaplan HI, Sadock BJ, editores. *Tratado de psiquiatría*. 2.a ed. Barcelona: Salvat; 1989. p. 585-607.
- Torrubia R, Ávila C, Moltó J, Caseras X. The Sensitivity to Punishment and the Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Pers Individ Dif*. 2001;31:837-62.
- Caseras X, Ávila C, Torrubia R. The measurement of individual differences in Behavioural Inhibition and Behavioural Activation Systems: a comparison of personality scales. *Pers Individ Dif*. 2003;34:999-1013.
- Serrano A, Pinto-Meza A, Peñarubia MT, Suárez D, Vicens E, Fernández R, et al. Selective serotonin reuptake inhibitors: Are there differences in effectiveness in primary care? *Prim Care Comm Psychiatr*. 2006;11:113-20.
- Brebner JMT. Happiness and personality. *Pers Individ Dif*. 1998;25:279-96.
- Brebner JMT, Martin M. Testing for stress and happiness: The role of personality factors. En: Spielberger CD, Sarason IG, Brebner J, Greenglass E, Laungani P, O'Rourke AM, editores. *Stress and emotion: Anxiety, anger and curiosity*. Washington, DC: Taylor and Francis; 1995. p. 139-72.
- Zuckerman M. Incentive motivation: Just extraversion? *Behav Brain Sci*. 1999;22:539-40.
- Zuckerman M, Joireman J, Kraft M, Kuhlman DM. Where do motivational and emotional traits fit within three factor models of personality? *Pers Individ Dif*. 1999;26:4087-504.
- Flieiss JL. The design and analysis of clinical experiments. New York: John Wiley & Sons; 1986.
- Altman DG. Practical statistics for medical research. New York: Chapman and May; 1991.
- Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73:13-22.
- Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimated equations approach. *Biometrics*. 1988;44:1049-60.
- Grove WM, Andreasen NC, McDonald-Scott P, Keller MB, Shapiro RW. Reliability studies of psychiatric diagnosis. Theory and practice. *Arch Gen Psychiatry*. 1981;38:408-13.
- Spitznagel EL, Helzer JE. A proposed solution to the base rate problem in the Kappa statistic. *Arch Gen Psychiatry*. 1985;42:725-8.
- Evans M, Kessler D, Lewis G, Peters TJ, Sharp D. Assessing mental health in primary care research using standardized scales: can it be carried out over the telephone? *Psychol Med*. 2004;34:157-62.