

El anonimato, la respuesta aleatoria e Internet como control de la deseabilidad social en contenidos sexuales

Anonymity, Randomized Response and Internet as a control for social desirability in sexual matters

*Javier Suárez, Ignacio Pedrosa, Blanca Pérez y Eduardo García-Cueto
Universidad de Oviedo*

El anonimato, la respuesta aleatoria e Internet como control de la deseabilidad social en contenidos sexuales

El objetivo de la presente investigación es estudiar la influencia que el método de aplicación puede tener en la deseabilidad social, cuando los ítems tienen una fuerte carga social negativa. Más concretamente, se quiere estudiar si el sesgo producido por la deseabilidad social varía en función del modo de aplicación del cuestionario. Para ello, se aplica un test sobre conducta sexual habitual bajo tres métodos distintos, uno de recogida de información anónima, otro utilizando la técnica de respuesta aleatoria (condicionando la respuesta a un suceso de probabilidad conocida a priori) y un tercero, utilizando la World Wide Web. La muestra total se reparte entre los tres métodos utilizados y está formada por 699 participantes con una media de edad 24,91 y una desviación típica de 8,30. En todos los casos, los participantes han sido varones mayores de edad y han accedido de forma voluntaria a rellenar el cuestionario. Los resultados obtenidos muestran que a pesar de existir diferencias en función del método, no aparece una ganancia significativa a favor de ninguno de ellos en cuanto a reducción de la deseabilidad social.

Palabras claves: deseabilidad social, respuesta aleatoria, Internet, sexualidad. Estudio descriptivo mediante encuestas

Anonymity, the random answer and Internet as a control of social desirability in sexual matters

This research was carrying on in order to study the influence that the method of application has on social desirability, particularly when items have a strongly negative social load. Specifically, the object of study is whether the bias produced by social desirability varies depending on the method of application of the questionnaire. For this reason, a test is applied to usual sexual conduct using three different methods; one is the collection of anonymous information, another is to use the random answer technique (determining the answer to a previously known event of probability) and a third one, using the World Wide Web. The final sample is divided among the three methods used. It consists of 699 participants with an average age of 24.91 and a standard deviation of 8.30. In every case, the participants are male adults and have agreed voluntarily to fill in the questionnaire. The results obtained show that despite the fact of differences depending on the method, used there are no significant benefits in favour of any of them (as for as a reduction of social desirability goes).

Key words: Social desirability, random answer, Internet, sexuality. Descriptive study by means surveys.

El anonimato, la respuesta aleatoria e Internet como control de la deseabilidad social en contenidos sexuales

A pesar de haber trabajos anteriores (Bernreuter, 1933), no es hasta los años 50 del siglo XX, con los trabajos de A.L Edwards, cuando el estudio de la deseabilidad social adquiere su máximo apogeo (Edwards, 1953,1957,1959). Este constructo se basa en la distorsión que las personas dan en las respuestas a un cuestionario con la intención de dar una imagen de sí mismas mejor de lo que en realidad es. De esta forma, puede dar lugar a que determinados sujetos seleccionen las alternativas más aceptables socialmente en vez de responder con sinceridad (Edwards, 1953,1957,1959; Morales, 2000). Ésta puede ser una explicación plausible a que se encuentren peores propiedades psicométricas en test de personalidad que en pruebas de capacidad. En este sentido, variables como neuroticismo, paranoidismo o estabilidad emocional reflejan características deseables o indeseables con bastante claridad mientras que la deseabilidad social en subescalas de capacidad se traduce en motivación por conseguir la máxima puntuación (Anastasi, 1973; Cronbach, 1985; Ellingson, Columbus, Smith y Sackett, 2001).

A consecuencia del interés suscitado por el sesgo que la deseabilidad social puede tener sobre los resultados de un test, numerosos autores han trabajado en instrumentos de medida que consigan minimizar el influjo de dicha variable contaminante. De entre los más relevantes, se encuentran: las escalas K y Lie del MMPI (Hataway y McKinley, 1942), la escala de deseabilidad social de Edwards (1957), la escala de deseabilidad social de Crowne-Marlowe (1960) y la escala de sinceridad de Eysenck (Eysenck y Eysenck, 1964). En trabajos posteriores, se observa como la deseabilidad social ya no es considerada sólo como una tendencia a dar respuestas distorsionadas, sino que se trata de un rasgo relativamente estable y consistente de personalidad (Eysenck y Eysenck, 1976). El interés por todos estos aspectos continúan hasta la actualidad, un claro ejemplo de ello son los trabajos de Ferrando y Chico (2000), Salgado (2005) y, Schermer y Vernon (2010).

Además de estos clásicos trabajos, existen otro importante número de ellos con el mismo objetivo, detectar y minimizar el influjo de la deseabilidad social. Así por ejemplo, casi todos los cuestionarios incluyen escalas de sinceridad (Dunnett, Koun y Barber, 1981; Ellingson et al., 1999). El funcionamiento de dichas escalas se basa en la presentación de ítems cuya probabilidad de aparición en la población es cercana a 1 ó 0, en función de la redacción. Cuando la tasa de respuesta dada difiere en exceso de lo esperado se invalidan todas las respuestas de la persona en ese test. Como principales inconvenientes a este método se encuentran: la pérdida de casos por personas que tratan de distorsionar la respuesta, que este tipo de escalas suele ser relativamente fácil de identificar y, que el sujeto responda adecuadamente a las preguntas “control” no quiere decir que el resto de respuestas no las esté distorsionando (es sensato pensar que una persona admita ciertas cuestiones más o menos triviales, como pueden ser los ítems control, pero los aspectos delicados o fundamentales del test los distorsione con el fin de ajustarse al perfil buscado o imagen deseada).

En cualquier caso, las escalas de sinceridad no son las únicas estrategias que se utilizan en la actualidad, de hecho, existen un gran número de ellas. Algunas de estas estrategias son: la utilización de ítems de elección forzada o test ipsativos (Cronbach, 1960; Scott, 1968; Braun, 1969); el método “bogus pipeline” basado en hacer creer al encuestado que mediante un sofisticado aparataje se puede comprobar la veracidad de sus respuestas (Aguinis, Pierce, y Quigley, 1995; Tourangeau, Simith y Rasinski, 1997; Alexander y Fisher, 2003); el método de preguntas mutuamente excluyentes (Warner,

1965) y la utilización de la técnica de análisis factorial (Ferrando, 2005). Finalmente, existen dos métodos, que por su extensa bibliografía merecen especial atención. Estos son el método de respuesta aleatoria (Miller, 1981; Bourke y Moran, 1988; Scheers y Dayton, 1988; Hosseini y Armacost, 1990; Zimmerman y Langer, 1995; Berrens, Bohara y Kerkvliet, 1997; Larkins y Hume, 1997; Wimbush y Dalton, 1997; García-Cueto, Cabañeros y Lozano, 2003) y las respuesta a través de Internet (Joinson, 1999; Pettit, 2002; Hancock y Flowers, 2001; Mühlensfeld, 2004; Duffy y Smith, 2005; Kreuter, Presser y Tourangeau, 2008; Jong, Pieters y Fox, 2010). Una amplia redacción de estos métodos puede verse en Nederhof (1985), Morales (2000), Myagkov (2002) y Salgado (2005).

La técnica de respuestas aleatorias está basada en el método de preguntas mutuamente excluyentes de Warner (1965) y se utiliza especialmente cuando se tocan cuestiones delicadas en las que una respuesta afirmativa (o negativa) ponga en riesgo la imagen social de la persona. La lógica subyacente es que un sujeto tenderá a decir la verdad, por muy dura que sea, si el encuestador no puede llegar a averiguar la respuesta de esa persona (Morales, 2000). Para ello, se pueden elaborar dos preguntas contradictorias o una relevante y otra intrascendente, condicionando la respuesta del sujeto a un suceso de probabilidad conocida a priori. Por ejemplo, Wise (1980), da a cada sujeto 20 cartas; en 14 de ellas pone *Responda la A* y las otras 6 *Responda la B*. El participante baraja y selecciona una carta, de esta forma, se asegura totalmente el anonimato y, el investigador, sabe que en el 70% de las respuestas lo serán al ítem delicado. Otro ejemplo de este método lo desarrolla García-Cueto et al. (2003) en un estudio donde lo comparan con la administración anónima y la entrevista a cara a cara en un cuestionario de verdadero y falso sobre diferentes temas comprometidos (agresividad, drogas, honradez, etc). En este caso, la respuesta aleatoria consiste en lanzar una moneda al aire, si sale cara se responde verdadero independientemente de la pregunta y, si sale cruz, se responde la verdad. Estos autores concluyen que existe un mejor control de la deseabilidad social cuando se utiliza el método de la respuesta aleatoria.

La expansión de Internet en los años 90, provoca que la investigación científica en general, y la Psicología en particular, se interese en su utilización dando lugar a la actual masificación de cuestionarios a través del World Wide Web (WWW.). Las ventajas que esta gran red de comunicación proporciona son numerosas respecto a las clásicas de papel y lápiz, sin embargo, a la par de estas ventajas se enumeran sus inconvenientes. De entre las ventajas, Best y Krueger (2004) proponen las nuevas posibilidades de muestreo (usuarios de ciertos servicios), flexibilidad en la administración (multimedia, aleatorización de preguntas, etc) y alternativas de recopilación (recordatorios, incentivos, etc). Sin embargo, también consideran ciertas limitaciones como puede ser la pobre accesibilidad a determinadas muestras (zona geográfica, nivel económico bajo, etc.), la posible variabilidad en la visualización del cuestionario o el riesgo que conlleva no poder supervisar la administración. Skitka y Sargis (2005) también proponen como ventajoso la eficiencia que estos cuestionarios vía Web tienen en cuanto a facilidad, rapidez y bajo coste. A pesar de ello, añaden inconvenientes como puedan ser las posibles diferencias existentes entre usuarios de Internet y no usuarios (inteligencia, edad, cultura...), dificultades para extraer muestras probabilísticas o menor probabilidad de participación (por un menor compromiso que el que se origina por métodos tradicionales). Para realizar un análisis más profundo véase Duffy y Smith (2005) y Orosa, Pinto y Sales (2008).

La discusión está abierta, se encuentran diversidad de opiniones en una u otra dirección acerca del uso de cuestionarios por Internet. Los que aconsejan su uso,

postulan una reducción de la ansiedad y la deseabilidad social, así como, una mayor tasa de respuesta (respuestas por unidad de tiempo) respecto a las clásicas aplicaciones de papel y lápiz (Moorman y Podsakoff, 1992; Joinson, 1999; Pettit, 2002). En el otro lado están trabajos como el de Hancock y Flowers (2001) en el que no encuentran diferencias significativas en deseabilidad social entre la WWW y los cuestionarios de papel y lápiz, ni tampoco entre condiciones de anonimato versus no anonimato. De forma similar, Mühlenfeld (2004) plantea un estudio comparativo entre un cuestionario Web (anónimo) y una entrevista personal a través de Internet (no anónimo). Sus resultados muestran que no existen diferencias entre los modos de entrevista en función del anonimato. En la misma línea de trabajo, Kreuter et al. (2008), estudian el sesgo de la deseabilidad social entre encuestas aplicadas telefónicamente y las realizadas por Internet. Los resultados de estos autores muestran que la administración Web incrementa el nivel y la precisión sobre los informes de datos delicados y, que estas diferencias aumentan en aquellos ítems con mayor carga social negativa. Finalmente concluyen que no hay un modelo superior a otro, sino que, la encuesta telefónica favorece a las tasas de respuesta y a evitar casos perdidos y, la administración Web mejora la precisión de los informes.

El objetivo de la presente investigación es estudiar si el método de aplicación de un cuestionario influye en la reducción de la deseabilidad social, sobre todo cuando los ítems tienen una fuerte carga social negativa. Más concretamente, se quiere estudiar si el sesgo producido por la deseabilidad social varía en función de la aplicación del cuestionario. El aporte novedoso de este trabajo es que se han utilizado aquellos métodos más estudiados y, también más discutidos, y se han comparado en un cuestionario con una alta intromisión personal. Para ello, se aplica un test de conducta sexual habitual bajo recogida de información anónima, utilizando la técnica de respuesta aleatoria (condicionando la respuesta a un suceso de probabilidad conocida a priori) y a través de Internet.

Si alguno de estos métodos influye para eliminar la deseabilidad social, en las preguntas con mayor carga social negativa se encontrará un porcentaje mayor de respuestas afirmativas y, en las preguntas neutras o sin carga social negativa, serán iguales en todos los tipos de aplicación.

Además, parece sensato pensar que Internet es el mejor de los métodos. Esto se explica en la medida en que puede llegar a ser más creíble para el participante que su identidad y sus respuestas dadas se mantengan en el anonimato cuando ni tan siquiera ven al investigador que sí, a pesar de asegurar el anonimato con respuesta aleatoria, el cuestionario se cubre cara a cara. Sin embargo, esto solo es en apariencia porque, en realidad, la moneda garantiza absolutamente el anonimato y en último caso, el rastreo de IP permite averiguar con poco margen de error quién ha cubierto el test por Internet.

Por otro lado, al disponer de un amplio cuestionario sobre conducta sexual habitual se han seleccionado aquellos ítems que, a priori, pudieran tener relación con conductas heterosexuales, homosexuales y sadomasoquistas para el estudio de la dimensionalidad. Si el cuestionario mide estas tres dimensiones los resultados se ajustarán a una estructura trifactorial.

Método

Participantes

Se ha contado con la colaboración de 699 varones, comprendidos en un rango de edad entre 18 y 77 años, con una media de 24,91 y una desviación típica de 8,30. Se han utilizado exclusivamente varones para no incluir el efecto de las diferencias que pudieran deberse al sexo, pues el principal objetivo es estudiar el efecto que el tipo de test puede ejercer sobre la deseabilidad social y no las conductas derivadas del test. Para los test de papel y lápiz se cuenta con: 130 participantes para la administración anónima, con una media de edad de 25,87 y una desviación típica de 12,027 y, con otros 130 para el método de la respuesta aleatoria, con una media de 25,03 y una desviación típica de 9,54. Por otro lado, se ha contado con 439 participantes para la administración Web del cuestionario, con una media de 24,98 y una desviación típica de 8,496. Todos los participantes son mayores de edad, han accedido de forma voluntaria a la realización del cuestionario y no han recibido compensación alguna.

Instrumentos

Se ha elaborado un cuestionario dicotómico (tipo verdadero y falso) sobre conducta sexual habitual con 44 ítems, 28 de los cuales son conductas con una fuerte deseabilidad social y dos preguntas control (cuya probabilidad de respuesta es conocida). Así por ejemplo, ítems como *Alguna vez me he masturbado en algún lugar público donde podían verme*, *He tenido relaciones sexuales con familiares cercanos* o *En alguna ocasión he besado a un hombre en la boca*, son considerados con una alta carga social negativa. Frente a ítems como *En alguna ocasión he besado a una mujer en la boca* o *Me excita ver a mujeres desnudas* que, en un principio, carecen de negatividad (debe tenerse en cuenta que la muestra está compuesta exclusivamente por varones). Finalmente, los ítems control son: *Habitualmente me desplazo montado a caballo y tengo cuatro Rolls Royce*.

Procedimiento

El mismo cuestionario sobre conducta sexual habitual se aplicó según tres métodos distintos. Por un lado, se realizó una aplicación grupal del cuestionario asegurando el absoluto respeto del anonimato. Por otro lado, se utiliza el procedimiento de respuesta aleatoria condicionando el resultado obtenido al lanzar una moneda al aire. Una vez planteada la pregunta, si el sujeto obtenía cara en la moneda, debía responder verdadero independientemente de lo que se preguntase y, si obtenía cruz, debía responder la verdad. De esta forma, se garantiza absolutamente el anonimato en la medida en que una respuesta afirmativa (lo que en la mayoría de ítems suponía una imagen social negativa) no se podría saber si es debida a que el sujeto obtuvo cara en la moneda o si realmente realiza esa conducta habitualmente.

Paralelamente se crea un dominio online donde se muestra el cuestionario y se difunde la dirección Web a través de e-mail y de redes sociales (Facebook, etc). En ningún caso se pide el nombre a los participante, únicamente información sociodemográfica (edad, sexo, etc). Además, se incluye al final del test un pequeño incentivo en forma de “feedback” a modo de agradecimiento respecto al tipo de conducta sexual derivado del conjunto de respuestas obtenidas.

Naturalmente, para poder comparar los resultados obtenidos mediante los métodos utilizados resulta imprescindible establecer algunos supuestos teóricos que permitan tomar decisiones en función del número de respuestas afirmativas a las cuestiones planteadas que no se deben al azar. Es decir, hay que tener en cuenta que las respuestas obtenidas con el método de la respuesta aleatoria (moneda) no se pueden comparar directamente con los otros métodos si no se establecen previamente unos supuestos teóricos.

Supuestos

1. La frecuencia de caras y cruces obtenidas al lanzar una moneda será aproximadamente igual a la frecuencia de caras y cruces esperadas. Siendo p la proporción de cruces obtenidas y π la probabilidad de cruces esperadas, $p = \pi$.
2. El resultado obtenido al lanzar la moneda es independiente de la conducta real del encuestado
3. El resultado obtenido al lanzar la moneda es independiente de la pregunta realizada.
4. Por consiguiente, si denominamos A al número de respuestas negativas observadas, y X el número de conductas estimadas con connotación social negativa, entonces, siendo n el tamaño muestral y π la probabilidad esperada, como en este caso, $\pi = 0,5$, entonces $X = n - 2 \sum_{i=1}^n A_i$

Análisis de los datos

Para comprobar el primero de los supuestos se incluyen dos preguntas control y se realiza una diferencia de proporciones. Se verificará el supuesto si la probabilidad asociada a la diferencia, $p - \pi > 0,025$ (Amón, 1978). Para estudiar las diferencias en función del método se realiza, al no cumplirse el supuesto de homocedasticidad y contar con grupos no balanceados, una prueba de análisis de varianza no paramétrico, en concreto, la prueba de Kruskal-Wallis. Finalmente, para el estudio de la dimensionalidad del cuestionario se realiza un Análisis Factorial Confirmatorio con el método de Máxima Verosimilitud. Todos los análisis se han realizado a un nivel de confianza del 95%.

Resultados

Se lleva a cabo una prueba de Z para comprobar la igualdad de proporciones en las preguntas de control. Como puede observarse en la *Tabla 1* el valor de la probabilidad asociada al estadístico Z no permite rechazar la hipótesis nula.

Tabla 1

Significación de diferencias acerca de las proporciones observadas y esperadas

Ítems Control	p	Proporción
Tengo cuatro Rolls Royce	0,60	0,48
Habitualmente me desplazo montado a caballo	0,16	0,44

Para estudiar las diferencias en función del método de aplicación se utiliza la prueba de Kruskal-Wallis. Se ha utilizado como variable de agrupamiento la forma de administración del cuestionario: anónimo, respuesta aleatoria (moneda) e Internet. Se han eliminado del análisis los ítems control y no se ha tenido en cuenta el método de la moneda en aquellos ítems donde el porcentaje de respuestas afirmativas es inferior al esperado. En concreto los ítems: 4, 6, 11, 12, 14, 17, 21, 23, 25, 33, 34, 35, 39 y 41. Los gráficos adjuntos muestran las diferencias encontradas en cada uno de los ítems en función del método de aplicación. En abscisas se sitúan los ítems y en ordenadas el valor medio del rango. Como se muestra en la *Gráfica 1*. (correspondiente a los ítems neutros) en 7 de 14 conductas neutras existen diferencias estadísticamente significativas en función del método (marcadas con *). Por otro lado, cuando los ítems tienen una fuerte carga social negativa se encuentran diferencias significativas en función del

método en 18 de las 28 conductas con alta deseabilidad social del cuestionario (ver *Gráfico 2*).

Gráfico 1.
Diferencias en función del método: Ítems neutros
(*Dif. Sig al NC=95%)

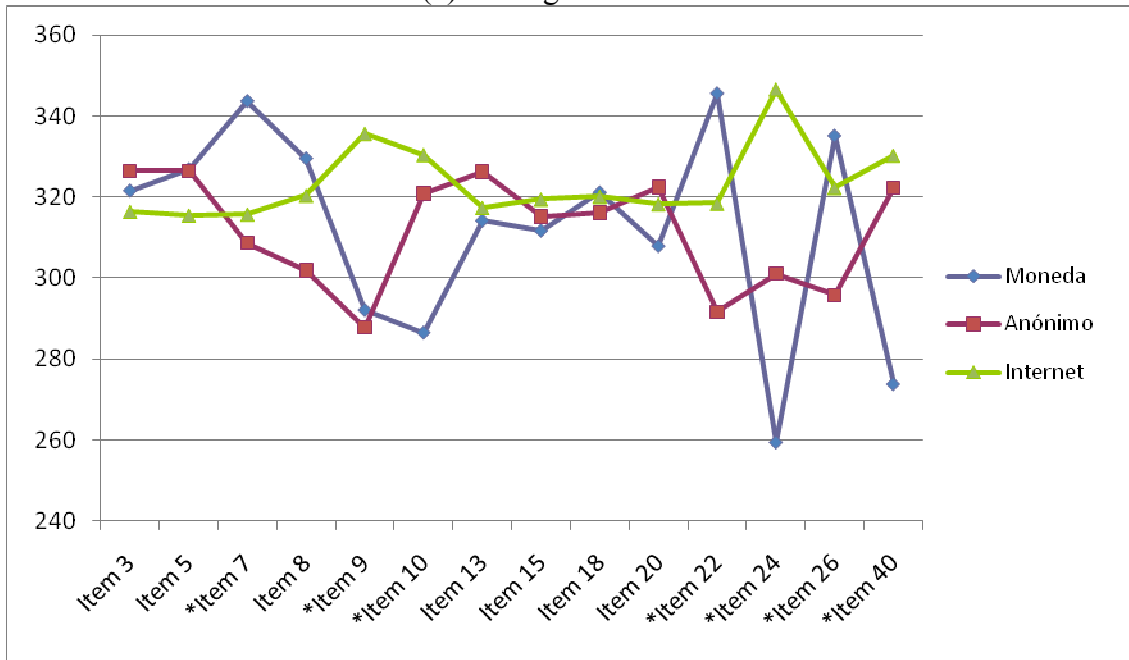
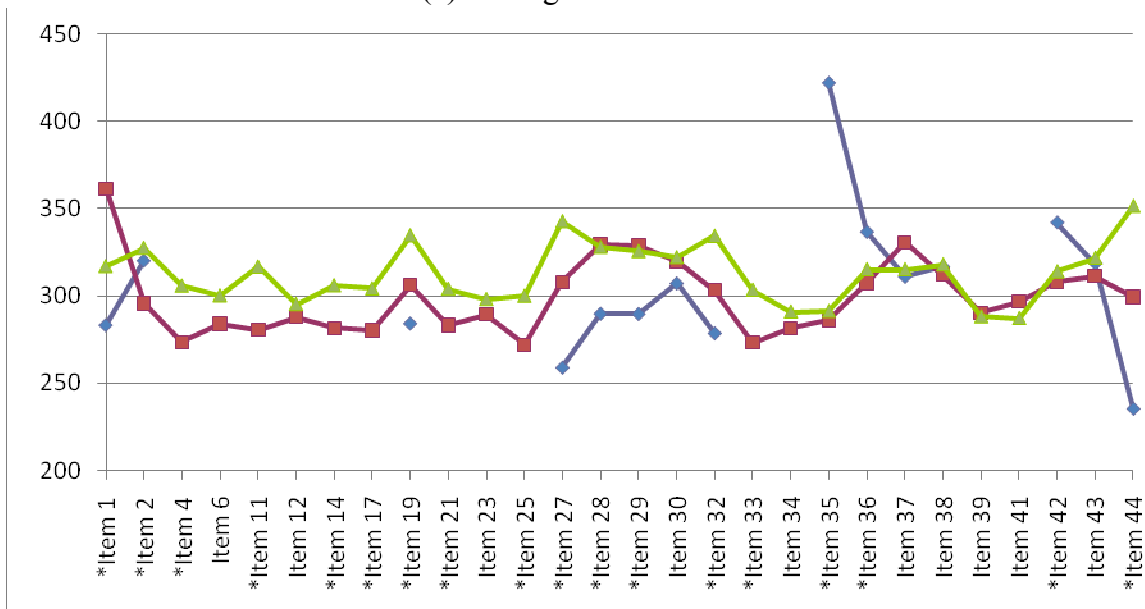


Gráfico 2.
Diferencias en función del método: Ítems con carga social negativa
(*Dif. Sig al NC=95%)



Para el estudio de la dimensionalidad del cuestionario se han eliminado los ítems control y aquellos ítems que puedan ser comunes a todos los factores (p.e. me masturbo). Finalmente se han utilizado 21 ítems y, como muestra la *Tabla 2.*, el modelo

parece mostrar un buen ajuste a una estructura trifactorial. Los coeficientes alpha de cada uno de los factores Conducta Heterosexual, Homosexual y Sadomasoquista son, respectivamente: 0,86; 0,92 y 0,85.

Tabla 2.
Índices de ajuste del AF Confirmatorio

CMIN/DF	NFI	RFI	IFI	CFI	RMSEA
1,982	,935	,917	,967	,966	,041

Discusión

El análisis de los ítems control (*Tabla 1*) para el método de respuesta aleatoria revela que la frecuencia de caras y cruces al lanzar la moneda no difiere de forma significativa de lo esperado (una proporción de 0,5). Según las hipótesis del estudio, las respuestas a los ítems neutros o sin carga social negativa deberían ser iguales en los tres métodos en la medida en que no tiene sentido que la deseabilidad social haga distorsionar las respuestas en estos casos. Sin embargo, los datos muestran que existen diferencias estadísticamente significativas (NC=95%) en 7 de los 14 ítems neutros. Para comprobar la ventaja, si la hubiera, de alguno de los métodos sobre otro, se deberían encontrar en las respuestas a los ítems negativos un porcentaje significativamente superior a favor del mejor de los métodos. Sin embargo, los datos muestran que solamente existen diferencias significativas en 18 de los 28 ítems con carga social negativa y, además, no siempre a favor del mismo método.

Por otro lado, es importante tener en cuenta que en 14 ítems, bajo el método de la moneda, el porcentaje de respuestas afirmativas ha sido inferior al esperado (una proporción de 0,5). Se sospecha que los participantes han mentido al responderlos ignorando las instrucciones y contestando falso cuando les salió cara en la moneda (curiosamente todos ellos con una fuerte carga social negativa). Sin embargo, el porcentaje de respuesta sólo difiere significativamente de lo esperado en uno de los ítems (*En alguna ocasión he lamido los genitales de un hombre*). Por lo tanto, sólo en este ítem se puede mantener la sospecha de que los participantes han respondido falso cuando les salió cara en la moneda, en el resto de casos pueden ser debidos al azar.

A la vista de los gráficos, cuando se trata de ítems con carga social negativa, se observa que existen más ítems donde la diferencia es significativa a favor de Internet que para los métodos de la moneda y el anónimo. Sin embargo, este resultado no es estable para todos los ítems negativos y que, además, existen casos favorables a cada uno de los métodos, hace insensato mantener la mejoría de alguno de ellos sobre el resto. Lo que parece indicar que no existe un método mejor que otro para el control de la deseabilidad social.

Por último, se han seleccionado aquellos ítems que, a priori, pudieran tener relación con conductas de tipo heterosexual, homosexual y sadomasoquista con el fin de obtener un cuestionario que mida estos aspectos. El análisis factorial confirmatorio indica que se dispone de un cuestionario válido para la medida de estos aspectos de la conducta sexual. Los valores del coeficiente alfa de Cronbach encontrados, indican que el cuestionario tiene una adecuada fiabilidad.

Como conclusión parece que ninguno de estos métodos provoca una clara mejoría en la reducción de la deseabilidad social, tal y como algún otro autor apoya (Hancock y Flowers, 2001 y Mühlenfeld, 2004). Sin embargo, es preciso tener en cuenta a la hora de contradecir trabajos como el de Moorman y Podsakoff (1992),

Joinson (1999), Pettit (2002) o García-Cueto et al. (2003) que se trata de un cuestionario exclusivamente sobre conducta sexual donde el contenido de los ítems quizá tenga una mayor connotación negativa. Se podría decir que en ítems con una alta intromisión en la vida privada todo el mundo miente.

Referencias

- Aguinis, H., Pierce C.A. y Quigley, B.M. (1995). Enhancing the validity of self-reporter alcohol and marijuana consumption using a bogus pipeline procedure: A meta-analytic review. *Basic and Applied Social Psychology*, 16(4), 512-527.
- Alexander, M.G. y Fisher, T.D. (2003). Truth and consequences: using the bogus pipeline to examine sex differences in self-reported sexuality. *Journal of Sex Research*, 40(1), 27-35.
- Amón, J. (1978). *Estadística Inferencial*. Madrid: Pirámide.
- Anastasi, A. (1973). *Tests psicológicos*. Madrid: Aguilar.
- Bernreuter, R.G. (1933). Validity of the personality inventory. *Personality Journal*, 11, 383-386.
- Berrens, R.P; Bohara, A.K. y Kerkvliet, J. (1997). A randomized response approach to dichotomous Choice contingent valuation. *American Journal of Agricultural Economic*, 79, 252-266.
- Best, S.J. y Krueger, B.S. (2004). *Internet Data Collection*. London: Sage.
- Bourke, J.R. y Moran, M.A. (1988). Estimating proportions from randomized response data using the EM algorithm. *Journal of the American Statistical Association*, 83(404), 964-968.
- Braun, J.R. (1969). Comment of Fricke's approach to controlling social desirability in the forced-choice format. *Psychological Reports*, 25(1), 93-84.
- Cronbach, L.J. (1960). *Essentials of Psychological Testing*. New York: Harper and Row.
- Cronbach, L.J. (1985). *Fundamentos de la exploración psicológica*. Madrid: Biblioteca Nueva.
- Crowne, D.P. y Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Duffy, B. y Smith, K. (2005). Comparing data from online and face-to-face surveys. *International Journal of Market Research*, 47, 6.
- Dunnett, S; Koun, S. y Barber, P. J. (1981). Social desirability in the Eysenck Personality Inventory. *British Journal of Psychology*, 72, 19-26.
- Edwards, A.L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, 37, 90-93.
- Edwards, A.L. (1957a). *Techniques of Attitude Scale Construction*. New York: Applenton Century-Crofts.
- Edwards, A.L. (1957b). *The social desirability variable in personality assessment and research*. New York: Dryden.
- Edwards, A.L. (1959). Social desirability variable and the description of others. *Journal of Social and Abnormal Psychology*, 59, 434-436.
- Ellingson, J. E; Sackett, P. R. y Hough, L.M. (1999). Social desirability corrections in personality measurement: Issues of applicants comparison and construct validity. *Journal of Applied Psychology*, 84, 155-166.

- Ellingson, J. E; Columbus, O. H; Smith, D. B y Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology*, 86(1), 122-133.
- Eysenck, H.J. y Eysenck, S.B. (1964). *The Eysenck Personality Inventory manual*. London: Routledge.
- Eysenck, H.J. y Eysenck, S.B.G. (1976). *Psychoticism as a dimension of personality*. New York: Crane, Russak & Company.
- Ferrando, P.J. y Chico, E. (2000). Adaptación y análisis psicométrico de la escala de deseabilidad social de Marlowe y Crowne. *Psicothema*, 12(3), 383-389.
- Ferrando, P.J. (2005). Factor Analytic Procedures for Assessing Social Desirability in Binary Items. *Multivariate Behavioral Research*, 40(3), 331-349.
- García-Cueto, E.; Cabañeros, J y Lozano, L.M. (2003). Control de la deseabilidad social en encuestas.
- Hancock, D.R. y Flowers, C.P. (2001). Comparing social desirability responding on world wide web and paper-administered surveys. *Educational Technology Research and Development*, 49, 5-13.
- Hataway, S.R. y McKinley, J.C. (1942). *The Minnesota Multiphasic Personality Inventory*. Minneapolis: University of Minnesota Press.
- Hosseini, J.C. y Armacost, R.L. (1990). Randomized responses: A better way to obtain sensitive information. *Business Horizons*, 82-86.
- Joinson, A. (1999). Social desirability, anonymity and Internet-based questionnaires. *Behavior Research Methods*, 31(3), 433-438.
- Jong, M.G.; Pieters, R. y Fox, J.P.(2010). Reducing social desirability bias through item randomized response: An application to measure underreported desires. *Journal of Marketing Research*, 47, 14-27.
- Kreuter, F.; Presser, S. y Tourangeau, R. (2008). Social desirability bias in CATI, IVR and Web Surveys. *Public Opinion Quarterly*, 72(5), 847-865.
- Larkins, E.R. y Hume, E.C. (1997). The validity of the randomised response method in tax ethics research. *Journal of Applied Business Research*, 13(3), -2532.
- Miller, J.D. (1981). Complexities of the randomised response test solution. *American Sociological Review*, 46(6), 928-930.
- Moorman, R.H. y Podsakoff, P.M. (1992). A meta-analytic review and empirical test of the potential confounding effects of social desirability response sets in organizational behaviour research. *Journal of Occupational and Organizational Psychology*, 65(3), 131-150.
- Morales, P. (2000). *Medición de actitudes en psicología y educación*. Madrid: Universidad Pontificia de Comillas.
- Mühlenfed, H.U. (2004). Differences between “talking about” and “admitting” sensitive behaviour in anonymous and non-anonymous web-based interviews. *Computers in Human Behaviour*, 21, 993-1003.
- Myagkov, A.Y. (2002). Statistical strategies for Collection of data about sensitive topics. *Sotsiologicheskie Issledovaniya*, 28(1), 111-121.
- Nederhof, A.J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, 15(3), 263-280.
- Orosa, F.J.; Pinto, I.F. y Sales, P.P. (2008). Cuestionarios psicológicos e investigación en Internet: Una revisión de la literatura. *Anales de Psicología*, 24, 150-157.
- Petit, F.A. (2002). A comparison of World-wide-web and paper-and-pencil personality questionnaires. *Behavior Research Methods*, 34(1), 50-54.
- Salgado, F. (2005). Personalidad y deseabilidad social en contextos organizacionales: implicaciones para la práctica de la psicología del trabajo y las organizaciones.

- Papeles del Psicólogo*, 26, 115-128.
- Scheers, N.J y Dayton, M. (1988). Covariate randomised response models. *Journal of the American Statistical Association*, 86(404), 969-974.
- Schermer, J.A. y Vernon, P.A. (2010). The correlation between general intelligence (g), a general factor of personality (GFP), and social desirability. *Personality and Individual Differences*, 48, 187-189.
- Scott, W.A. (1968). Comparative validities of forced-choice and single-stimulus tests. *Psychological Bulletin*, 70, 231-244.
- Skitka, L. Y Sargis, E. (2005). Social psychological research and the Internet: the promise and peril of a new methodological frontier. En Y. Amichai-Hamburger (Ed), *The social net: Human behavior in cyberspace*. New York: Oxford University Press.
- Torangeau, R., Smith, T.W. y Rasinski, K.A. (1997). Motivation to report sensitive behaviours on surveys: Evidence from bogus pipeline experiment. *Journal of Applied Social Psychology*, 27(3), 209-222.
- Warner, S.L. (1965). Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, 63-69.
- Wimbush, J.C. y Dalton, D.R. (1997). Base rate for theft: Convergence of multiple methods. *Journal of Applied Psychology*, 82(5), 576-583.
- Wise, S.L. (1980). *An approach to sensitive-questions: the randomized response technique*. Alexandria, Virginia: Eric Reports.
- Zimmanman, R.S y Langer, L.M. (1995). Improving estimates of prevalence rates of sensitive behaviours: The randomised lists technique and considerations of self-reported honesty. *Journal of Sex Research*, 32(2), 107-117.