

Good guys or potential cooperators: an experimental and Bayesian account for two explanations

Said E. Jiménez

July 2019

1 Introducción

Los seres humanos somos animales sociales y obtenemos ventajas de serlo. Historicamente, pertenecer a un grupo social ha brindado protección y garantizado el acceso a recursos como el alimento, actualmente se ha mostrado evidencia de que pertenecer a un círculo social tiene un efecto positivo en la longevidad, la salud física y mental (Holt-Lunstad, 2018; Tough, Siegrist, & Fekete, 2017). La interrelación entre la cercanía social y la cooperación es clave para la conformación de grandes grupos de individuos sin una relación genética, dichos grupos conforman la base de comunidades, sociedades y naciones, así como probablemente constituyen una de las condiciones más fundamentales para la supervivencia humana (Fehr & Fischbacher, 2003; Fehr & Schurtenberger, 2018).

Un requisito importante para que haya cooperación entre individuos es la comunicación, un metaanálisis de 45 estudios reportó un efecto positivo grande ($d = 1.01$) de la comunicación en la cooperación independientemente del medio comunicativo (Balliet, 2010). De manera paralela, se sabe que posteriormente a sesiones de comunicación entre individuos completamente desconocidos hay un incremento de la cercanía social subjetiva (Aron, Melinat, Aron, Vallone, & Bator, 1997). Sugiriendo una relación positiva entre la comunicación, la cooperación y la cercanía social.

Una de las formas de comunicación que ha recibido mayor atención en la teoría de juegos psicológica son las promesas, en una interacción entre quien la envía y la recibe, se cree que las promesas influyen en las creencias del receptor generando confianza y cooperación (Charness & Dufwenberg, 2006; Vanberg, 2008). Sin embargo, existen situaciones en que se traiciona la confianza, se rompen las promesas o se engaña, por ejemplo, aproximadamente el 11% de las personas encuestadas por la General Social Survey (GSS) en 2018 respondieron que sí han tenido sexo con alguien diferente a su pareja mientras están casados (Smith, Davern, Freese, & Hout, 2018). Otro ejemplo, de acuerdo con el Washington Post el presidente Trump ha dicho 10,796 afirmaciones falsas o engañosas en 869 días en el cargo (Kessler, Rizzo, & Kelly, 2019).

En el laboratorio se han estudiado explícitamente estas transgresiones a la confianza en experimentos donde las personas tienen incentivos para mentir o no cumplir sus promesas. Sin embargo, la inmensa mayoría de estos estudios en economía, psicología y neurociencias,

se han realizado en personas que NO se conocen entre sí (por mencionar algunos, Gneezy, 2005; Baumgartner, Fischbacher, Feierabend, Lutz, & Fehr, 2009; Baumgartner, Gianotti, & Knoch, 2013; Charness & Dufwenberg, 2006; Fischbacher & Föllmi-Heusi, 2013; Gneezy, Rockenbach, & Serra-Garcia, 2013; Mazar, Amir, & Ariely, 2008). En psicología social existen estudios que abordan el engaño en las relaciones interpersonales cercanas, sin embargo utilizan metodologías de autorreporte (DePaulo, Ansfield, Kirkendol, & Boden, 2004; DePaulo & Kashy, 1998) o más bien exploran el desarrollo de habilidades de detección del engaño en amigos del mismo sexo (Anderson, DePaulo, & Ansfield, 2002).

2 Métodos

2.1 Sujetos

Se incluyeron 45 sujetos (15 hombres), reclutados en la Universidad Nacional Autónoma de México, en un rango de edad de 19 a 33 años y su nivel educativo mínimo fue licenciatura. Los participantes acudieron al estudio con un amigo considerado *cercano* por ellos mismos, quien cumplía las siguientes características: era una persona del mismo sexo que el participante, no tenía un vínculo familiar y no se trataba de una persona con una relación de tipo sentimental y/o sexual.

2.2 Tarea y procedimiento

Todos los participantes realizaron una adaptación de la tarea de confianza y promesas presentada por Baumgartner (2009), que consiste en 24 ensayos entre dos jugadores **A** y **B**, el jugador **A** tiene originalmente 0 pesos mexicanos y el jugador **B** tiene 2 pesos, al jugador **B** se le presenta la oportunidad de invertir su dinero al jugador **A** o conservarlo, si invierte su dinero se multiplica por 5 pesos, de modo que el jugador **A** recibe 10 pesos. Finalmente, el jugador **A** toma la decisión de pagar la mitad al jugador **B** o quedarse los 10 pesos. La estructura descrita se repite en 24 ensayos, sin embargo, en 4 de los ensayos el jugador **A** puede enviar una promesa a **B**, las promesas son que *siempre*, *casi siempre*, *algunas veces* o *nunca* pagaran la mitad. Cada promesa es válida por tres ensayos, de modo que 12 de los ensayos tienen el efecto de la promesa y los otros 12 no lo tienen.

La variación respecto a la tarea original es que en nuestro experimento se presentaron tres compañeros de juego con tres niveles de cercanía social en el papel del jugador **B**: **computadora** (cercanía nula), **confederado** (cercanía baja) y **amigo** (cercanía alta), mientras que el participante principal siempre fungió el papel del jugador **A**. Cada compañero de juego realizó el rol de **B** en 8 de los 24 ensayos totales, sin embargo, tanto la condición de promesas como la de cercanía social fueron presentadas al sujeto **A** en orden pseudoaleatorio, lo que significa que se realizó la aleatorización de los ensayos una vez y se mantuvo el mismo orden para todos los sujetos. Con el fin de simplificar la tarea y homogeneizar la experiencia de todos nuestros participantes, se programaron *a priori* las decisiones de los compañeros **B** sobre si invertían o no su monto inicial, se programó que los tres compañeros decidieran invertir en 6 ensayos y no invertir en 2, los cuales también se presentaron al sujeto **A** en orden pseudoaleatorio. La historia encubierta para todos nuestros sujetos en el papel de **A**

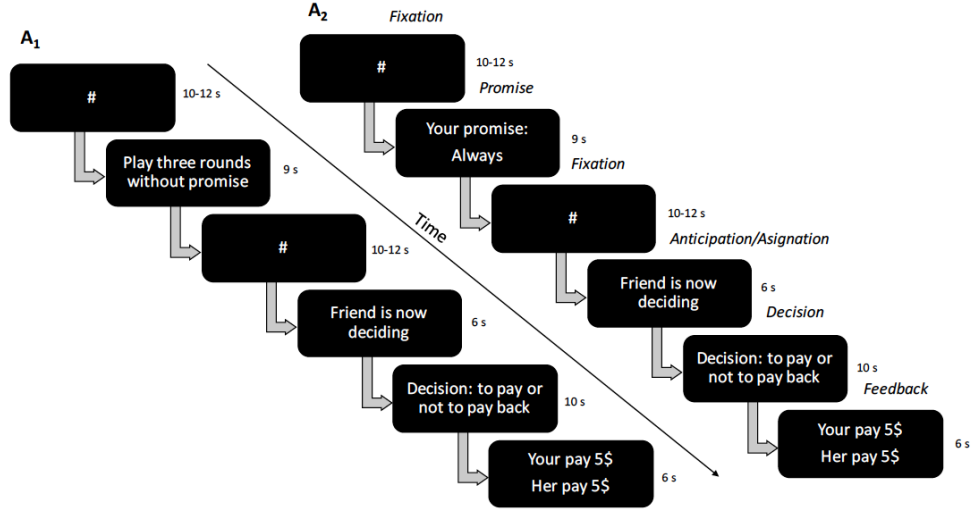


Figure 1: Task of promises and social closeness

fue que estarían jugando en tiempo real con su amigo, el confederado (que se le dijo que sería otra persona desconocida de su mismo sexo) y la computadora.

En la Figura 1 se muestra un diagrama de la cronología de la tarea experimental con las duraciones de cada fase en segundos, cada caja de izquierda a derecha representa una pantalla que se le mostró a los sujetos de manera secuencial. La parte A₁ corresponde a un ejemplo de los ensayos sin promesas y la A₂ de los ensayos con promesas. La fase de fijación sólo consistió en un período en el que el sujeto prestaba atención sin realizar ninguna conducta en particular, la **fase de promesas** en la parte A₁ le indicó al sujeto que los siguientes tres ensayos se podían decidir sin el efecto de la promesa, mientras que en A₂ se le pidió al sujeto que decidiera entre *siempre*, *casi siempre*, *algunas veces* o *nunca* pagar la mitad. A lo anterior, le siguió otro período de fijación, posteriormente los ensayos tanto en A₁ como en A₂ continuaron aproximadamente de la misma forma, en la **fase de anticipación/asignación** se le indicó a los sujetos quién era su compañero de juego para ese ensayo (computadora, confederado o amigo) y se le dió el mensaje de que su compañero estaba tomando su decisión, posteriormente en la **fase de decisión** se le dió al participante la información sobre cuál había sido la decisión de su compañero (si invirtió sus 2 pesos o no), se le recordaba su nivel de promesas (si se trataba de ensayos con promesas) y, en el caso de que su compañero hubiera invertido, se le pedía que decidiera si pagaba de regreso o no. Finalmente, en la **fase de feedback** se mostraban los pagos para ese ensayo y se repetía la secuencia. La tarea se programó en PsychoPy2 versión 1.84.2, que registraba las respuestas de los sujetos y los tiempos de reacción en segundos.

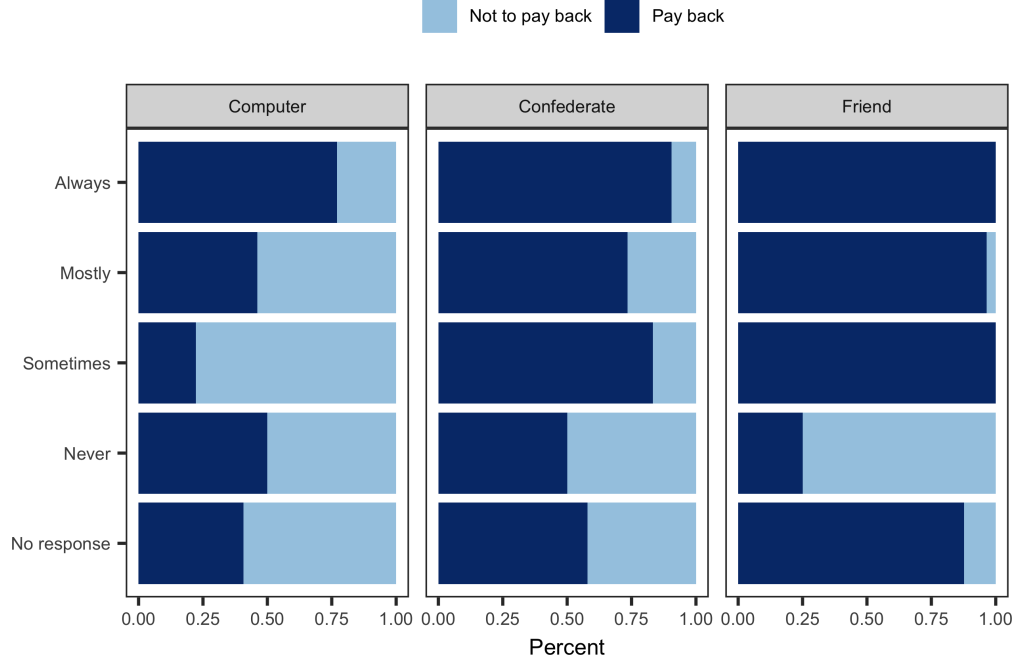


Figure 2: Pay back rate by partners and promises levels

3 Resultados

3.1 Descriptivos

Se analizaron 756 decisiones sobre pagar o no la mitad (18 por sujeto), así como 168 promesas (4 por sujeto), 69.6% de las decisiones fueron pagar la mitad, mientras que la proporción de las promesas seleccionadas fueron 4.17%, 14.9%, 41.1%, 37.5%, respectivamente para las promesas, nunca, algunas veces, casi siempre y siempre. En la Figura 2 se muestra la proporción de pagos dependiendo de los niveles de promesas y los compañeros de juego. Romper promesas, es decir, que el sujeto haya prometido siempre y posteriormente decida no pagar la mitad ocurrió solamente el 11.1% de los ensayos, de los cuales, 7.63% fueron ensayos con la computadora, 3.47% con el desconocido y 0% con el amigo.

3.2 Modelamiento Bayesiano Jerárquico

Todos los modelos presentados a continuación se programaron en R por medio del paquete **brms** (Bürkner, 2017; R Core Team, 2019), que realiza la inferencia mediante el muestreo por Markov chain Monte Carlo a través de **Stan** (Stan Development Team, 2018). Para cada modelo, las distribuciones posteriores de todos los parámetros fueron aproximadas con cuatro cadenas de 2000 iteraciones cada una, las primeras 1000 iteraciones de cada cadena fueron descartadas (periodo de calentamiento), para un total de 4000 muestras post-calentamiento. La convergencia de los modelos fue evaluada a través de la inspección visual de las cadenas y el cálculo del estadístico \hat{R} , que para todos los parámetros fue de 1, lo que se interpreta como convergencia.

3.2.1 Efecto de la cercanía social en el engaño

Para evaluar la hipótesis de que la cercanía social reduciría el engaño, se filtraron todos los ensayos en los que los sujetos promietieron que *siempre* pagarían y después decidieron no pagar. Posteriormente, se utilizó inferencia bayesiana para evaluar el efecto de los compañeros de juego en la decisión de romper la promesa a nivel individual y poblacional. Para tal propósito, se realizó un modelo bayesiano jerárquico que asume que la incertidumbre en la estimación del efecto de los compañeros sobre la decisión de romper la promesa varía dependiendo de cada individuo, sin embargo, también asume que dichas variaciones pertenecen a distribuciones poblacionales comunes para el efecto de cada compañero sobre la decisión (Gelman & Hill, 2006).

$$y_i \sim \text{Bernoulli}(\theta_i) \\ \text{logit}(\theta_i) = \mathbf{X}\beta + \mathbf{Z}u$$

En el modelo anterior, la decisión de romper la promesa y_i proviene de una distribución Bernoulli con probabilidad de pagar θ_i , el objetivo del modelo jerárquico es predecir cada decisión a través de la combinación lineal de los efectos de cada compañero de juego, transformados por su función ligadora inversa logit (Bürkner, 2017). En este modelo, β y u son coeficientes a nivel poblacional y nivel individual respectivamente, mientras que \mathbf{X} , \mathbf{Z} son sus correspondientes matrices del diseño. En este caso, los coeficientes poblacionales corresponden con la presencia de los compañeros con cercanía social nula β^{computer} , baja $\beta^{\text{confederate}}$ y alta β^{friend} .

En la Figura 3, se muestra la probabilidad posterior de romper la promesa dependiendo del compañero de juego, los círculos corresponden con las medianas de las distribuciones de las estimaciones posteriores de los efectos, la barra gruesa y la barra delgada corresponden con el intervalo del 50% y 95% de probabilidad posterior. Es claro que la probabilidad de romper la promesa va disminuyendo conforme va aumentando la cercanía social, sin embargo, calculamos las razones de evidencia para las siguientes hipótesis:

1. Hay mayor probabilidad de que se rompa la promesa a un compañero con cercanía nula (computadora) en comparación con los compañeros con cercanía social (desconocido y amigo).
2. Hay mayor probabilidad de que se rompa la promesa a un compañero con cercanía social baja (desconocido) en comparación con el compañero de cercanía social alta (amigo).

La razón entre la probabilidad posterior de las hipótesis mencionadas y sus correspondientes hipótesis alternativas fue de 10.33 y 5.58 respectivamente para las hipótesis 1 y 2. Cabe mencionar que la hipótesis de que la cercanía social reduciría el engaño fue expresada *a priori* como parte de un proyecto de investigación más amplio y se puede consultar aquí: <https://osf.io/re7kp/>.

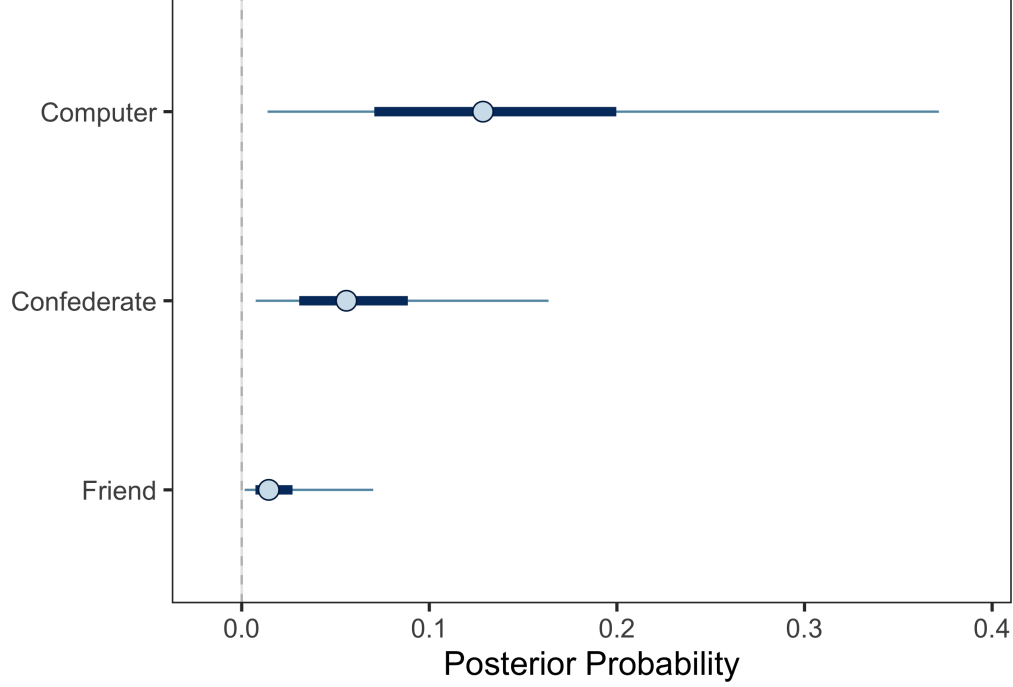


Figure 3: Posterior probability of Breaking the Promise

3.2.2 Efectos de las promesas y la cercanía social en la decisión

Para evaluar el efecto de las condiciones experimentales en la decisión de pagar la mitad, se utilizaron todos los ensayos en los que los compañeros decidieron invertir. Nuevamente, se realizó un modelo bayesiano jerárquico, que asume que las promesas y los compañeros de juego tienen un efecto que varía para cada individuo, sin embargo, también asume que dichas variaciones pertenecen a distribuciones poblacionales comunes. El modelo estima que la probabilidad θ_i de que en cada ensayo se decida pagar la mitad está en función del efecto de la presencia de promesas $\beta^{promise}$, así como de los compañeros con cercanía social baja $\beta^{confederate}$ y alta β^{friend} .

En la Figura 4 se presentan las distribuciones posteriores de los coeficientes β en escala logit, la línea central representa la mediana de la distribución y el área sombreada corresponde con el intervalo de 50% de probabilidad posterior. Como se puede observar, más del 95% de la densidad posterior de los coeficientes poblacionales está superior al 0, lo que indica fuerte evidencia del efecto de las condiciones experimentales en la decisión de pagar, de modo que la presencia de promesas y del confederado aumentan claramente la probabilidad de pagar, pero la presencia del amigo es la condición que más efecto tiene en esta conducta.

Asimismo, la Figura 5 presenta la distribución posterior predictiva en contraste con la proporción de las decisiones a cada compañero, cada panel corresponde con uno de los primeros doce sujetos en la fase de promesas. Las respuestas sobre la línea punteada indicarían que el sujeto pago la mitad de los ensayos en que ese compañero estuvo presente. La distribución posterior predictiva simula observaciones del modelo y las compara con los datos reales, nos sirve para identificar si el modelo es suficientemente cercano al proceso que

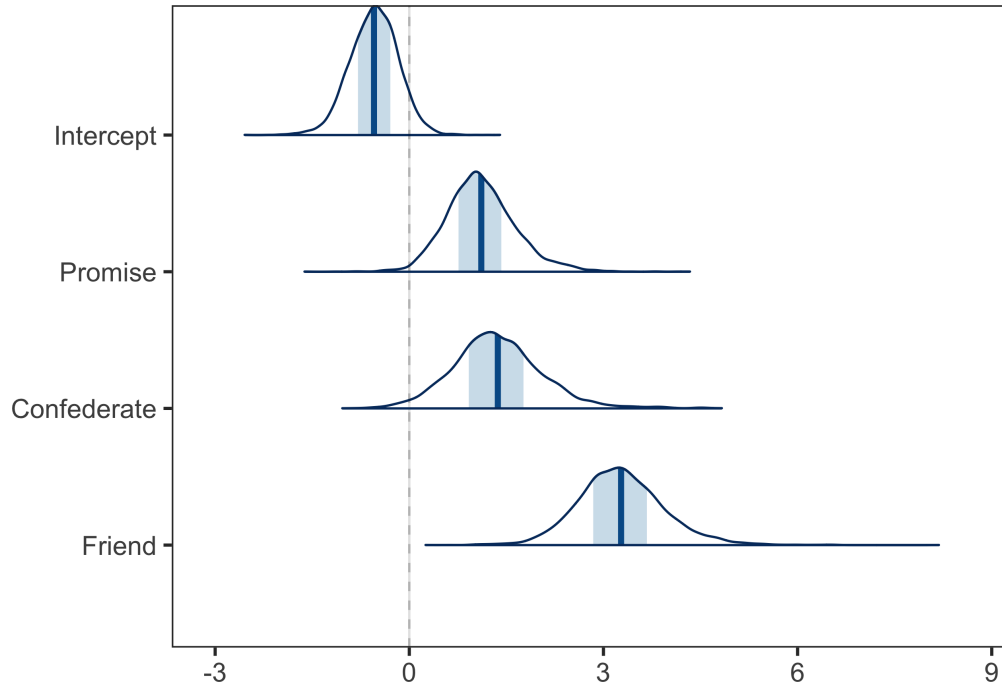


Figure 4: Posterior estimates in logit scale

ha generado los datos (Lee & Wagenmakers, 2014; Schad, Betancourt, & Vasishth, 2019). Se puede observar que existe gran correspondencia entre la respuesta de los sujetos y las predicciones del modelo, aún en los casos en los que el modelo se “equivoca” (por ejemplo el sujeto 1) la respuesta observada se sitúa en el margen de error de la predicción, lo que da credibilidad a las estimaciones.

3.2.3 Efectos de la cercanía social variantes por promesas

En este modelo se asume que la cercanía social tiene un efecto que varía para cada nivel de promesa, lo que implica que hay niveles de promesa que son más sensibles que otros al efecto de los compañeros de juego sobre la decisión de pagar la mitad. Nuevamente se realizó un modelo bayesiano jerárquico para estimar el efecto de los compañeros de juego a nivel poblacional y las variaciones dependiendo de los niveles de promesa. La Tabla 1 resume las distribuciones posteriores de los coeficientes del modelo en escala logit, incluye los estimadores puntuales, sus errores estándar y los intervalos de credibilidad bayesianos del 95%. Se puede observar una estimación similar a los modelos anteriores, con evidencia fuerte del efecto de la cercanía social sobre la decisión de pagar la mitad. Por otra parte, en la Figura 6 se muestra la distribución posterior predictiva, comparada con las tasas de pago de todos los individuos a los diferentes compañeros y su variación por el nivel de promesa. A excepción de con la promesa *never*, se observa un efecto positivo monótonico de los compañeros de juego en todos los niveles de promesa, sin embargo se puede observar cómo el efecto de la cercanía social varía dependiendo de la fuerza de la promesa.

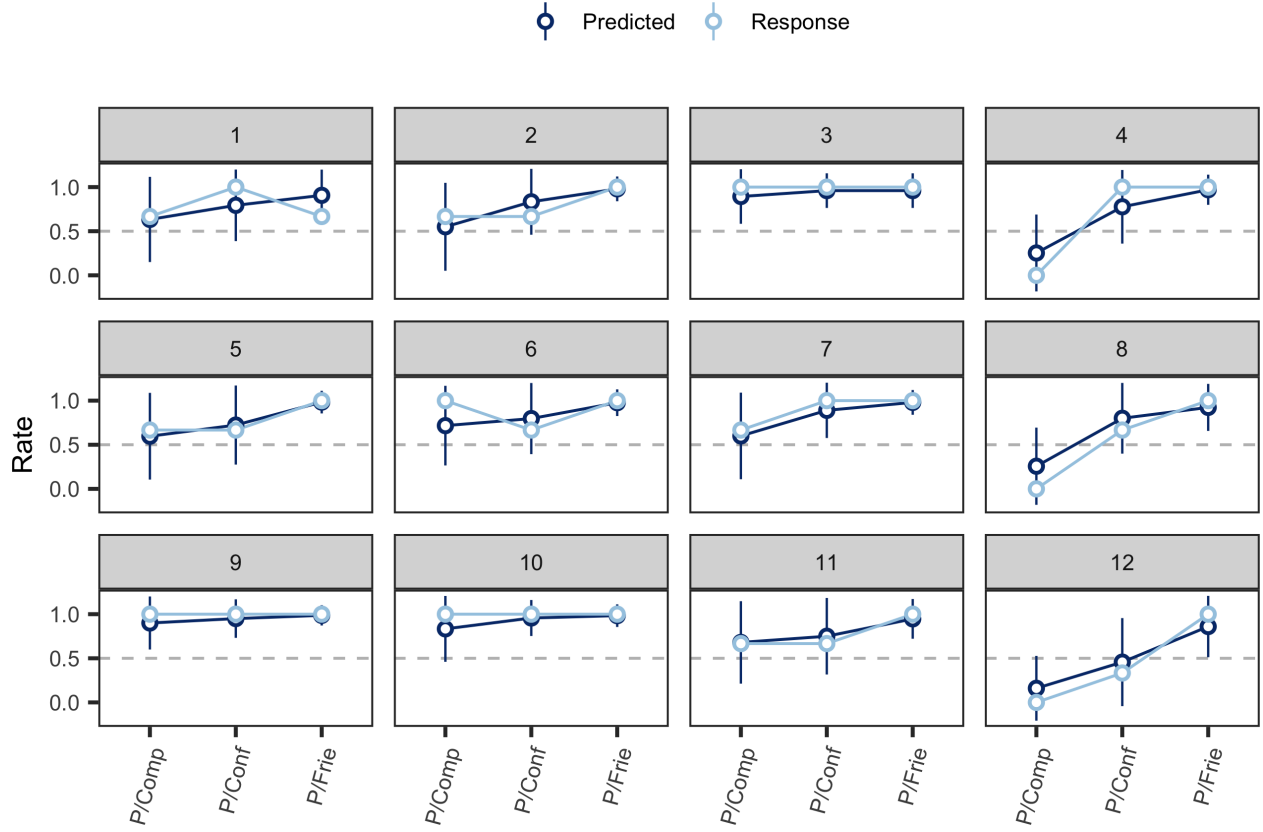


Figure 5: Posterior predictive over pay back rate

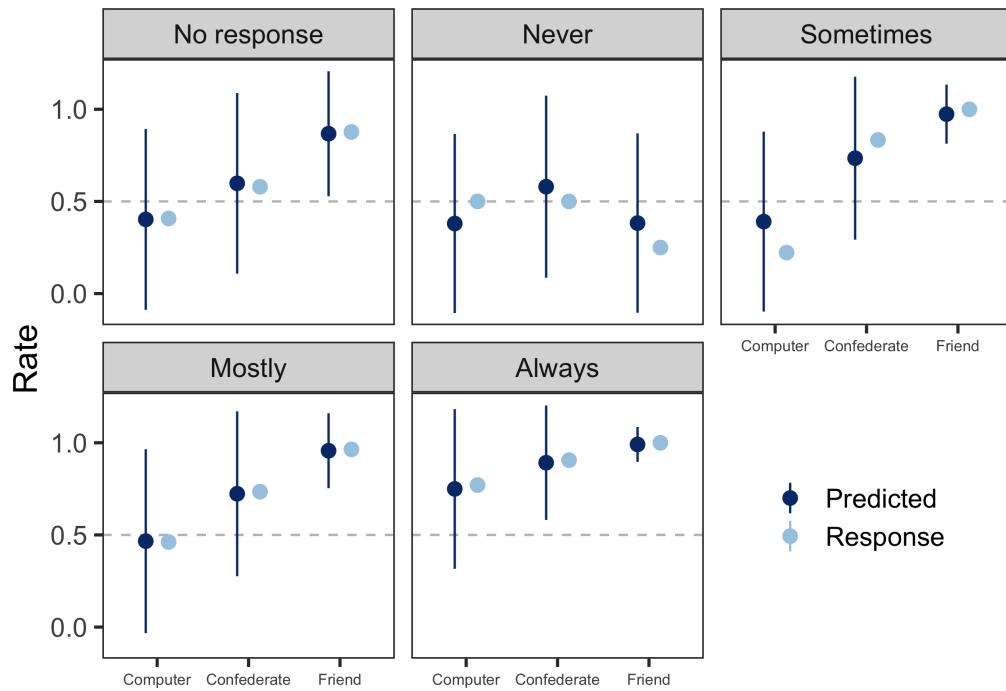


Figure 6: Posterior predictive over pay back rate, Social Closeness Varying effects by Promise levels

Table 1: Posterior coefficients estimates

Term	Estimate	Est.Error	95 % CI	
			Lower	Upper
Intercept	-0.130	0.577	-1.062	0.774
Confederate	1.000	0.430	0.325	1.669
Friend	2.268	1.196	0.097	4.005

3.2.4 Compromiso expresado en las promesas

En el estudio original de promesas (Baumgartner et al., 2009), los autores optaron por dividir su muestra en dos de acuerdo a la técnica de agrupamiento jerárquico con el método de Ward. De esta manera obtuvieron dos conjuntos de participantes que eran diferentes en sus tasas de pago, a pesar de que ambos hacían promesas muy altas, por tal motivo los autores denominaron al grupo que pagaba poco *deshonestos* y al grupo que pagaba mucho *honestos*. En un ejercicio similar, en el presente estudio realizamos la técnica de agrupamiento jerárquico para obtener una solución de 2 grupos y encontramos dos conjuntos similares en n que denominamos grupo “Low” y “High” (Low = 23, High = 19), con proporciones de pago bastante diferentes (Low = 58%, High = 84%), con una diferencia en un intervalo de confianza del 95% de 19% a 32% en favor del grupo High.

Aunque pareciera un resultado similar al reportado en el trabajo de Baumgartner (2009), exploramos el patrón de promesas de ambos grupos para determinar si era posible clasificar a nuestros sujetos como honestos y deshonestos. Si hipotetizamos que el compromiso de pagar sería reflejado en el nivel de promesas seleccionado, un grupo de deshonestos podría generar en sus compañeros la creencia de que pagará al seleccionar un nivel alto de promesas (siempre o casi siempre) y, posteriormente, traicionar esa confianza al decidir no pagar la mitad.

Con el objetivo de explorar si los grupos representan a poblaciones que no difieren en el nivel de compromiso expresado en las promesas, utilizamos un modelo bayesiano acumulativo, que asume que los niveles de promesas son una variable ordinal observada Y que se origina de la categorización de una variable latente continua \tilde{Y} , para este caso, *el compromiso expresado de pagar la mitad* (Bürkner & Vuorre, 2019). El grado en que los sujetos del grupo High difieren del grupo Low, en desviaciones estándar normales (z -values), en la escala latente de \tilde{Y} , tiene un estimador puntual de 0.95, lo que implica que el grupo de High tiene 0.95 z -values mayor compromiso de pagar la mitad que el grupo Low.

El intervalo de credibilidad bayesiano del 95% indica que el grupo High está entre 0.35 a 1.58 z -values de diferencia con respecto a Low. Por lo que podemos concluir con al menos 95% de probabilidad posterior que las personas que pertenecen al grupo High expresaron en sus promesas mayor compromiso de pagar la mitad de lo que hicieron los sujetos del grupo Low. Si observamos la Figura 7, la probabilidad de seleccionar los diferentes niveles de promesas varía dependiendo del grupo, el que tuvo mayor porcentaje de decisiones de pagar la mitad tiene más probabilidad de seleccionar *Always* (grupo High), mientras que el grupo que tuvo menor porcentaje de decisiones de pagar la mitad tiene más probabilidad de

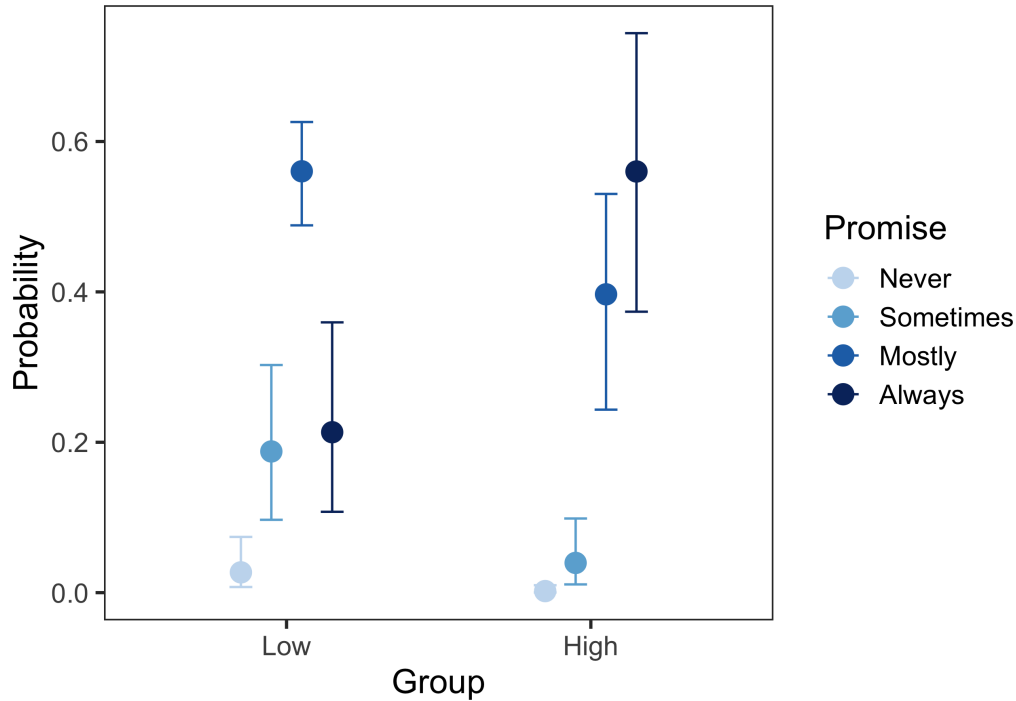


Figure 7: Promise choices by group

seleccionar *Mostly* (grupo Low).

Según nuestros datos, no podríamos justificar la clasificación de nuestros grupos de acuerdo con su honestidad o deshonestidad, al menos no con la técnica empleada en Baumgartner (2009). Ya que las personas fueron bastante consistentes con cumplir lo que prometían pagar.

4 Discusión

Referencias

- Anderson, D. E., DePaulo, B. M., & Ansfield, M. E. (2002). The development of deception detection skill: A longitudinal study of same-sex friends. *Personality and Social Psychology Bulletin*, 28(4), 536–545. <https://doi.org/10.1177/0146167202287010>
- Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin*, 23(4), 363–377. <https://doi.org/10.1177/0146167297234003>
- Balliet, D. (2010). Communication and cooperation in social dilemmas: A meta-analytic review. *Journal of Conflict Resolution*, 54(1), 39–57. <https://doi.org/10.1177/0022002709352443>
- Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., & Fehr, E. (2009). The Neural Circuitry of a Broken Promise. *Neuron*, 64(5), 756–770. <https://doi.org/10.1016/j.neuron.>

2009.11.017

Baumgartner, T., Gianotti, L. R. R., & Knoch, D. (2013). Who is honest and why: Baseline activation in anterior insula predicts inter-individual differences in deceptive behavior. *Biological Psychology*, 94(1), 192–197. <https://doi.org/10.1016/j.biopsycho.2013.05.018>

Bürkner, P.-C. (2017). brms : An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1). <https://doi.org/10.18637/jss.v080.i01>

Bürkner, P.-C., & Vuorre, M. (2019). Ordinal Regression Models in Psychology: A Tutorial. *Advances in Methods and Practices in Psychological Science*, 2(1), 77–101. <https://doi.org/10.1177/2515245918823199>

Charness, G., & Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, 74(6), 1579–1601. <https://doi.org/10.1111/j.1468-0262.2006.00719.x>

DePaulo, B. M., Ansfield, M. E., Kirkendol, S. E., & Boden, J. M. (2004). Serious Lies. *Basic & Applied Social Psychology*, 26(2/3), 147–167. https://doi.org/10.1207/s15324834basp2602&3_4

DePaulo, B. M., & Kashy, D. A. (1998). Everyday Lies in Close and Casual Relationships. *Journal of Personality and Social Psychology*, 74(1), 63–79. <https://doi.org/10.1037/0022-3514.74.1.63>

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960), 785–791. <https://doi.org/10.1038/nature02043>

Fehr, E., & Schurtenberger, I. (2018). Normative foundations of human cooperation review-article. *Nature Human Behaviour*, 2(7), 458–468. <https://doi.org/10.1038/s41562-018-0385-5>

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, 11(3), 525–547. <https://doi.org/10.1111/jeea.12014>

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. (p. 651). <https://doi.org/10.2277/0521867061>

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, 95(1), 384–394. <https://doi.org/10.1257/0002828053828662>

Gneezy, U., Rockenbach, B., & Serra-Garcia, M. (2013). Measuring lying aversion. *Journal of Economic Behavior and Organization*, 93, 293–300. <https://doi.org/10.1016/j.jebo.2013.03.025>

Holt-Lunstad, J. (2018). Why Social Relationships are Important for Physical Health: A Systems Approach to Understanding and Modifying Risk and Protection. *Ssrn*, (October 2017), 1–22. <https://doi.org/10.1146/annurev-psych-122216-011902>

Kessler, G., Rizzo, S., & Kelly, M. (2019). President Trump has made more than 10,000 false or misleading claims. Retrieved from https://www.washingtonpost.com/politics/2019/06/10/president-trump-has-made-false-or-misleading-claims-over-days/?utm{_}term=.8565a16aa5e8

- Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling*. <https://doi.org/10.1017/cbo9781139087759>
- Mazar, N., Amir, O., & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research*, 45(6), 633–644. <https://doi.org/10.1509/jmkr.45.6.633>
- R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Schad, D., Betancourt, M., & Vasishth, S. (2019). Toward a principled Bayesian workflow in cognitive science. Retrieved from osf.io/b2vx9
- Smith, T., Davern, M., Freese, J., & Hout, M. (2018). GSS Data Explorer | NORC at the University of Chicago. Retrieved from <https://gssdataexplorer.norc.umd.edu/variables/5067/vshow>
- Stan Development Team. (2018). RStan: The R interface to Stan. Retrieved from <http://mc-stan.org/>
- Tough, H., Siegrist, J., & Fekete, C. (2017). Social relationships, mental health and wellbeing in physical disability: A systematic review. *BMC Public Health*, 17(1), 1–18. <https://doi.org/10.1186/s12889-017-4308-6>
- Vanberg, C. (2008). Why Do People Keep Their Promises? An Experimental Test of Two Explanations¹. *Econometrica*, 76(6), 1467–1480. <https://doi.org/10.3982/ECTA7673>