

Nice guys or potential cooperators when keeping promises: 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). Así, los hallazgos sugieren 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 los sujetos 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 mediciones 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). Pese a la presencia generalizada de las interacciones entre seres humanos que pertenecen al mismo grupo, no se conoce cómo la cercanía social entre sujetos puede tener un efecto en el cumplimiento o rompimiento de las promesas.

En el presente manuscrito, se exploran los efectos de tres compañeros con diferentes niveles de cercanía social en el cumplimiento y rompimiento de promesas, así como en la cooperación en un juego de confianza. Lo anterior, lo analizamos a la luz de las dos motivaciones principales que se han señalado en la literatura respecto a cumplir promesas (Baumgartner et al., 2009):

1. La *instrumental*, sugiere que se cumplen las promesas para hacer más fácil la cooperación en el futuro.
2. La *intrínseca*, menciona que se cumplen las promesas por el propósito de hacer lo moralmente correcto.

En nuestro estudio, los sujetos realizan un juego de confianza estándar en parejas con tres fases: *primero*, el fideicomisario realiza una promesa sobre pagar la mitad de sus ganancias independientemente de quién sea su inversor; *segundo*, el inversor recibe la promesa y decide si invierte su presupuesto inicial; *tercero*, en caso de que el inversor haya dado el presupuesto, el fideicomisario se enfrenta a la decisión de pagar o no la mitad de sus ganancias. Con el propósito de evaluar el efecto de la cercanía social, nuestros sujetos participan en el papel de fideicomisario frente a tres compañeros con diferentes niveles de cercanía social (nula, baja y alta) en el rol del inversor: una *computadora*, un *extraño* y un *amigo*.

Las manipulaciones mencionadas nos permiten evaluar varias hipótesis, la primera fue propuesta *a priori* y es derivada de un proyecto de investigación más grande (se puede consultar el preregistro de la hipótesis aquí: <https://osf.io/u97fd>), mientras que las siguientes son de carácter exploratorio:

1. La cercanía social reducirá la conducta de romper la promesa. De acuerdo con la motivación *instrumental*, esperamos que los sujetos mantengan sus promesas con los amigos por el propósito de facilitar la cooperación en el futuro, la cual se puede extender inclusive más allá de los ensayos en el experimento. También esperamos, aunque en menor grado, que los sujetos cumplan las promesas a los extraños, con el propósito de facilitar la cooperación por lo menos durante los ensayos que dure el experimento. Finalmente, anticipamos que los participantes rompan las promesas a la computadora, por tratarse de un compañero sin cercanía social y del que no podrían asegurar cooperación en futuros ensayos. Cabe destacar que si los participantes cumplen las promesas a la computadora, se daría evidencia en favor de la motivación *intrínseca*.

2. Habrá un efecto de la cercanía social en la cooperación, independientemente de las promesas. Esperamos que haya más probabilidad de pagarle al *amigo* que al *extraño* y más probabilidad de pagarle al extraño que la *computadora*. De acuerdo con la motivación *instrumental*, la cooperación será mayor para compañeros con los que el sujeto anticipe mayor cooperación en el futuro ($\text{amigo} > \text{extraño} > \text{computadora}$).
3. Finalmente, obtendremos dos submuestras de acuerdo con la tasa de cooperación de los sujetos como se realizó en otro estudio (Baumgartner et al., 2009). Sin embargo, lo que en el estudio mencionado se clasificó como grupo de *honestos* y *deshonestos*, mostraremos que en nuestra muestra no se sostiene porque si los grupos difieren en su tasa de pago será por la diferencia en su nivel de compromiso expresado en las promesas.

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 sujeto, no tenía un vínculo familiar y no se trataba de una persona con una relación de tipo sentimental y/o sexual. 30 de los sujetos (15 hombres) realizaron la tarea dentro de un equipo de Imagen por Resonancia Magnética (MRI), sin embargo, sus datos de imagen no son analizados en el presente trabajo.

2.2 Tarea

Todos los sujetos realizaron una adaptación del juego de confianza y promesas con recompensas monetarias hipotéticas (Baumgartner et al., 2009). La tarea se programó en PsychoPy2 versión 1.84.2 (Peirce et al., 2019; Peirce, 2008) y consistió en 24 ensayos entre dos jugadores: fideicomisario e inversor. El fideicomisario tiene originalmente 0 pesos mexicanos y el inversor tiene 2 pesos, al inversor se le presenta la oportunidad de dar su dinero al fideicomisario o conservarlo, si invierte su dinero se multiplica por 5 pesos, de modo que el fideicomisario tenga 10 pesos. Finalmente, el fideicomisario toma la decisión de pagar la mitad al inversor o quedarse los 10 pesos.

La estructura descrita se repite en 24 ensayos, sin embargo, en 4 de los ensayos el fideicomisario puede enviar una promesa al inversor, 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. Cabe mencionar, que las promesas no se realizan hacia algún compañero en particular, si no que el sujeto expresa su nivel de compromiso de pagar la mitad independientemente de quiénes sean sus compañeros en los tres siguientes ensayos en que la promesa es válida.

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 inversor: *computadora*

(cercanía nula), *extraño* (cercanía baja) y *amigo* (cercanía alta), mientras que el sujeto funge el papel de fideicomisario. Cada compañero de juego realiza el rol de inversor en 8 de los 24 ensayos totales, sin embargo, tanto la condición de promesas como la de cercanía social fueron presentadas al fideicomisario en orden pseudoaleatorio.

Dado que nuestro interés principal fue la conducta del fideicomisario, se programaron *a priori* las decisiones de los inversores para que dieran su monto en 6 ensayos y en 2 no lo hicieran. La historia encubierta para todos nuestros sujetos fue que estarían jugando en tiempo real con su amigo, el extraño (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 continúan 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, extraño 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.

2.3 Procedimiento

Cada sujeto acudió con un amigo de su mismo sexo considerado cercano por él o ella, se les enfatizó que no tuvieran una relación romántica y/o familiar, para tratar de excluir el efecto en la cooperación por la relación consanguínea o por atracción sexual. Asimismo, para garantizar que los sujetos y sus acompañantes experimentaban un grado de cercanía social similar entre ellos, en el laboratorio ambos respondieron la escala de inclusión en el otro IOS (Aron et al., 1997) sin que su compañero observara sus respuestas. La escala consiste en siete pares de círculos que varían en el grado de traslape entre ellos, el respondiente debe seleccionar el par de círculos que mejor representa la cercanía subjetiva respecto a su compañero.

Los sujetos eran entrenados en la tarea por dos de los investigadores, primero se les explicó verbalmente el juego, y posteriormente se les presentó la interfaz computarizada en un equipo portátil. Se les indicó que tomarían sus decisiones con cuatro botones del teclado de la computadora, los cuales correspondían con los niveles de promesas y con sólo dos de estos

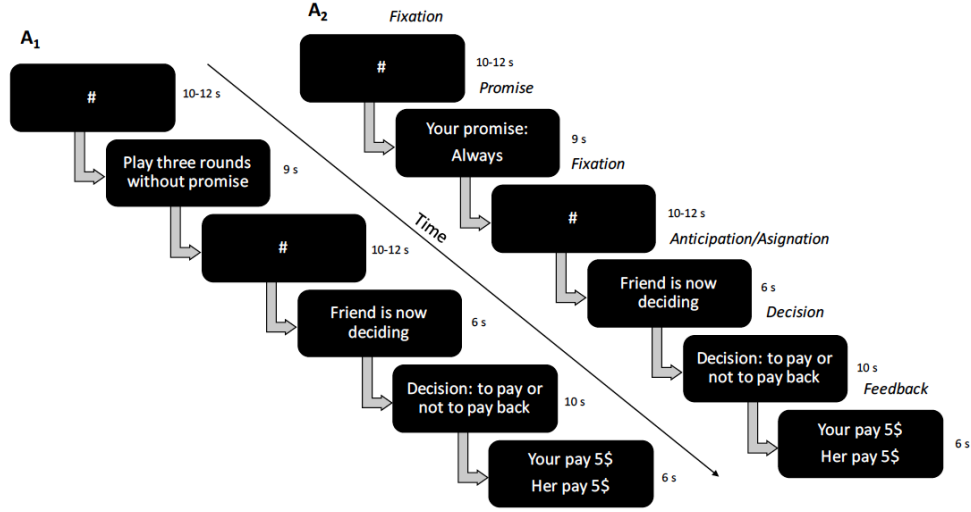


Figure 1: Task of promises and social closeness

para tomar su decisión de pagar o no la mitad. Los amigos escucharon la explicación, se les indicó que ellos tomarían sus decisiones en una habitación aparte en otra computadora. Los sujetos realizaban entre 3 y 6 ensayos de práctica junto con el investigador los cuales ejemplificaban el curso del juego cuando el inversor daba su presupuesto. En los ensayos de práctica se mostraban todas las fases señaladas en la Figura 1, sin embargo, la identidad de los inversores estaba reemplazada por signos de interrogación.

Cuando los sujetos no reportaban dudas procedíamos a llevarlos sin sus acompañantes a otra sala dónde tomarían sus decisiones en una computadora portátil, o bien, con el sistema de respuestas Lumina cedrus en el caso de que fuera un sujeto al que también realizamos MRI. Los sujetos iniciaban la tarea creyendo que jugarían realmente con los tres compañeros de diferente cercanía social. En la otra habitación o sala, otro de los investigadores realizaba el *debriefing* al acompañante, se les comentaba que el propósito de la investigación era saber si su presencia tenía un efecto en las decisiones del sujeto, pero que para simplificar el análisis sus respuestas fueron programadas *a priori* por lo que su participación concluía ahí.

Cuando el sujeto concluía la tarea también se realizaba el *debriefing*, se le comentaba que las decisiones de sus compañeros habían sido programadas. Aunque ningún participante ejerció su derecho, se les indicó tanto a los sujetos como sus acompañantes que podrían retirar su participación y consentimiento informado si consideraban estar en desacuerdo por el alguna de las manipulaciones realizadas por los investigadores.

3 Resultados

3.1 Descriptivos

Se analizaron 810 decisiones sobre pagar o no la mitad (18 por sujeto), así como 180 promesas (4 por sujeto), 69% de las decisiones fueron pagar la mitad, mientras que la proporción de las

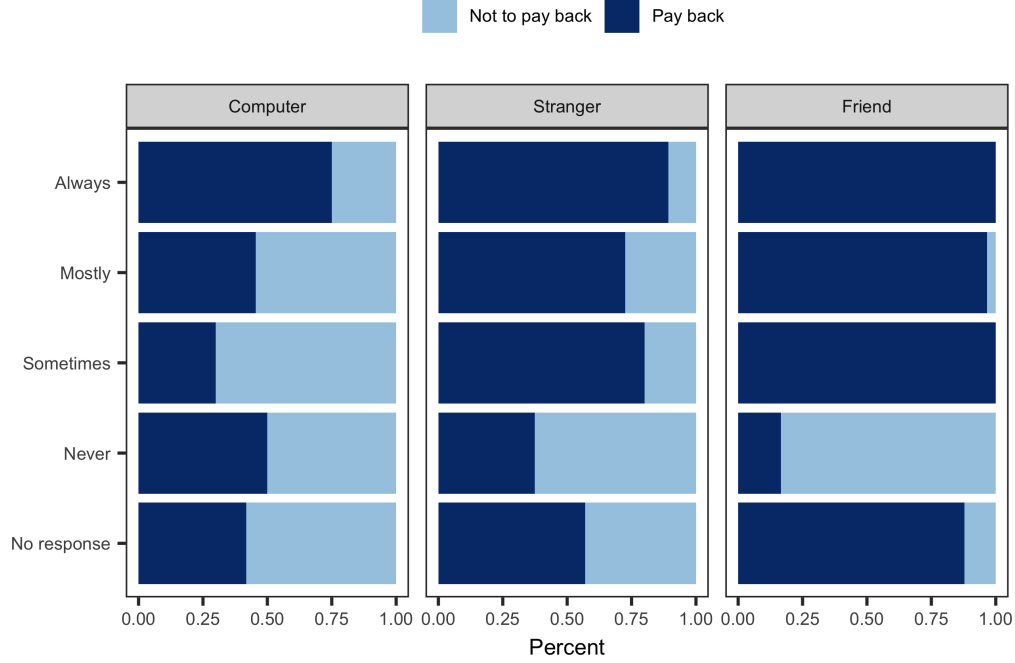


Figure 2: Pay back rate by partners and promises levels

promesas seleccionadas fueron 5.4%, 16.2% 43.8%, 40.2%, 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 12.4% de ensayos (19 veces no pagaron la mitad de 153 que prometieron *siempre*), de los cuales, 8.50% (13/153) fueron ensayos con la computadora, 3.92% (6/153) con el extraño y 0% (0/153) con el amigo. Es fundamental notar que romper la promesa no fue la decisión más frecuente, ya que aún cuando se trataba de la computadora, el 91.5% de los ensayos en que los sujetos prometieron *siempre* y tuvieron este compañero, decidieron pagar la mitad.

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. Se utilizaron distribuciones iniciales vagamente informativas para los parámetros de interés en los modelos, lo que permite que los datos dominen la inferencia, asimismo asume lo menos posible respecto a la naturaleza del fenómeno, lo que podría

ser adecuado para el estado de evidencia actual en el problema que estamos estudiando (McElreath, 2018).

3.2.1 Cercanía social entre parejas

Para evaluar la cercanía social, 30 (15 hombres) de nuestros sujetos y sus amigos respondieron la escala de inclusión en el otro IOS (Aron et al., 1997; Fareri & Delgado, 2014; Sip, Smith, Porcelli, Kar, & Delgado, 2015), que consisten en siete pares de círculos que varían en el grado de traslape y representan la cercanía social que percibe un individuo con respecto al otro. Un modelo Bayesiano jerárquico y acumulativo (Ver sección: Compromiso expresado en las promesas), apoya la hipótesis de que no existen diferencias entre la cercanía social subjetiva entre los sujetos y sus amigos, con una razón de evidencia posterior de 2.74 a su favor.

3.2.2 Efecto de la cercanía social sobre romper las promesas

Para evaluar la hipótesis de que la cercanía social reduciría la conducta de romper la promesa, 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 (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 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 β^{stranger} 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:

- Hay mayor probabilidad de que se rompa la promesa a un compañero con cercanía nula (computadora) en comparación con el compañero de cercanía social baja (extraño).

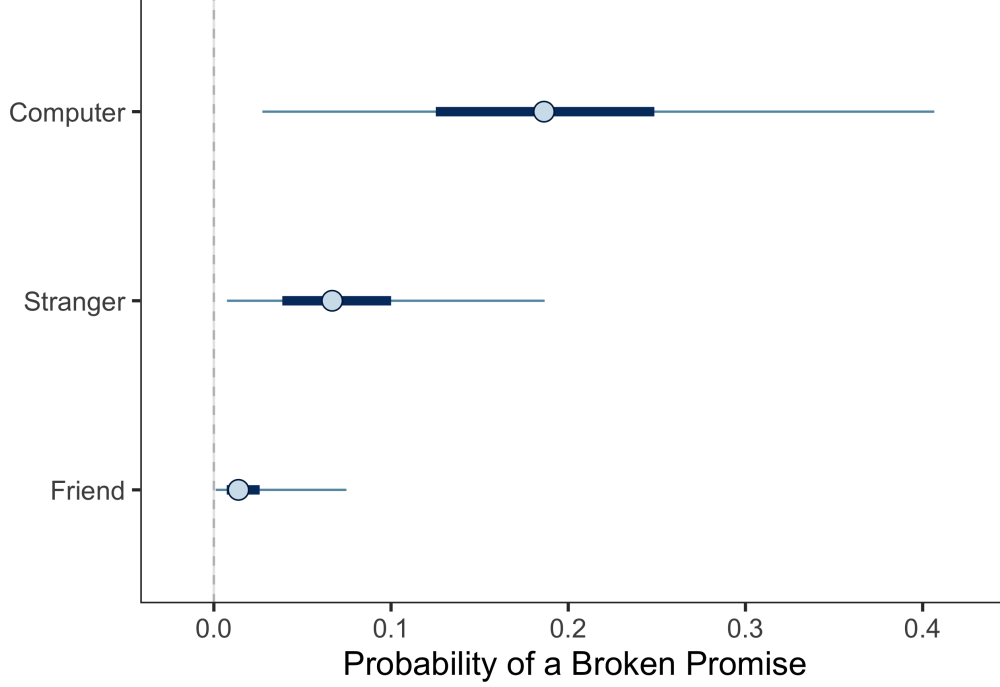


Figure 3: Posterior probability of Breaking the Promise

- Hay mayor probabilidad de que se rompa la promesa a un compañero con cercanía social baja (extraño) 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 7.47 y 8.64 respectivamente en favor de las hipótesis anteriores. Como se observa en la Figura 5 en la estimación para $\beta^{computer}$, cuando no hay cercanía social tenemos el 95% de probabilidad posterior de que la proporción de rompimiento de promesas poblacional esté entre 4-38%, que si bien es una proporción pequeña, claramente excluye la probabilidad de que no se rompan promesas.

3.2.3 Efectos de las promesas y la cercanía social en la cooperación

Para evaluar el efecto de las condiciones experimentales en la decisión de pagar la mitad, se utilizaron todos los ensayos en los que se programó que los compañeros invirtieran. Nuevamente, se realizó un modelo Bayesiano jerárquico, que asume que la fase de 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 $\beta^{stranger}$ 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

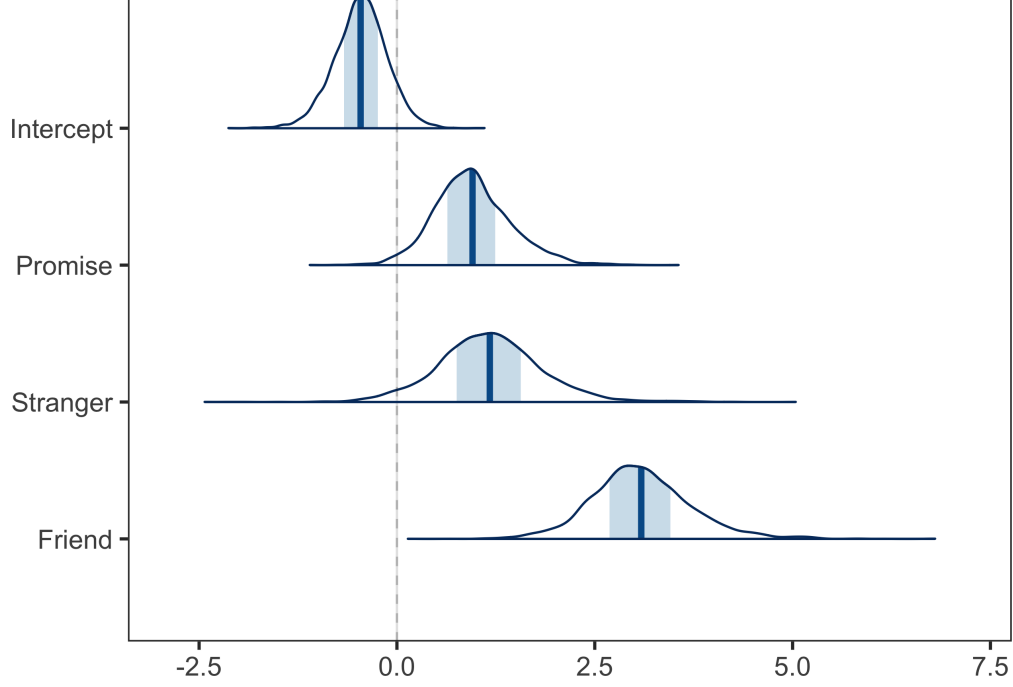


Figure 4: Posterior estimates in logit scale

evidencia del efecto de las condiciones experimentales en la decisión de pagar, de modo que la presencia de promesas y del extraño aumentan claramente la probabilidad de pagar, pero la presencia del amigo es la condición que más efecto tiene en esta conducta. Las distribuciones posteriores muestran claramente que la probabilidad de cooperación aumenta conforme lo hace también la cercanía social.

Asimismo, la Figura 5 presenta la distribución posterior predictiva en contraste con la proporción de las decisiones a cada compañero durante la fase de promesas, cada panel corresponde con uno de los primeros doce sujetos. 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 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 rango de una desviación estándar de la predicción, lo que da credibilidad a las estimaciones.

3.2.4 Efectos de la cercanía social en la cooperación 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

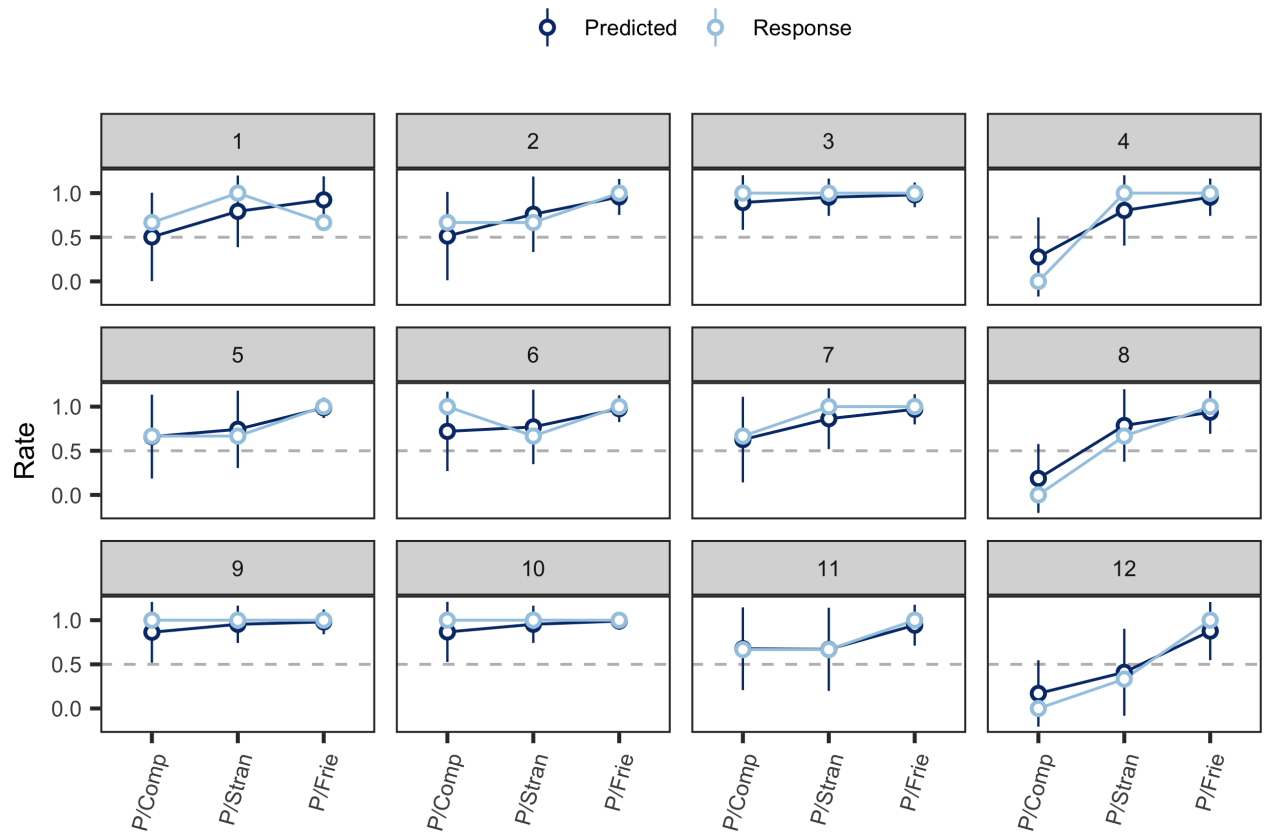


Figure 5: Posterior predictive over pay back rate

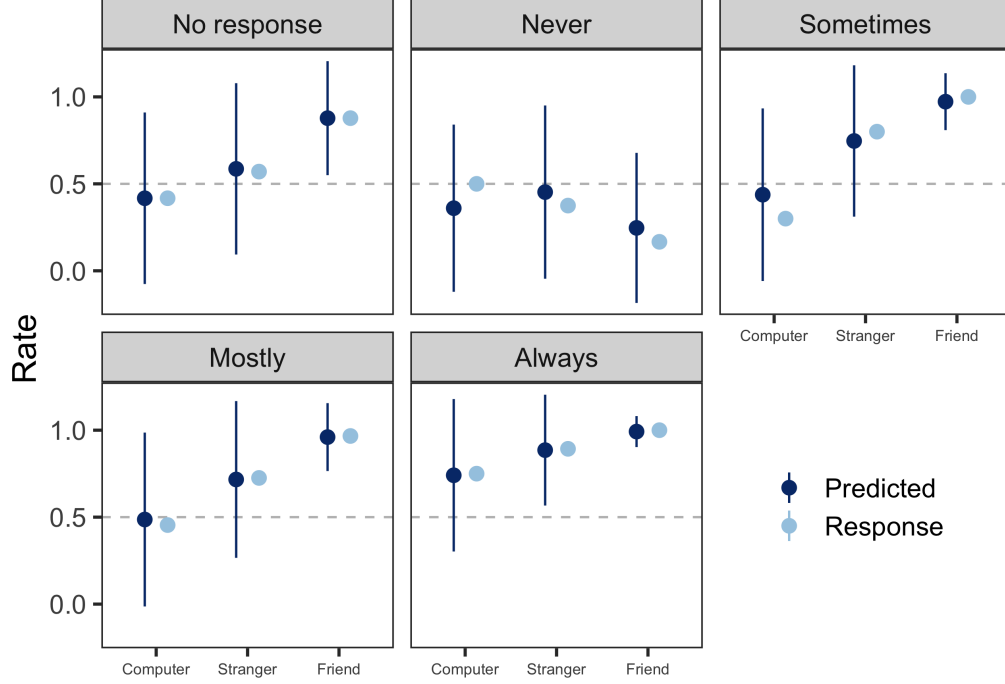


Figure 6: Posterior predictive over pay back rate, Social Closeness Varying effects by Promise levels

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. Aunque el intervalo de credibilidad para el efecto del amigo incluye el 0, la razón de evidencia respecto a que dicho efecto es mayor que cero es de 15.81 con 94% de probabilidad posterior.

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 monotónico 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, principalmente para las decisiones hacia la computadora.

Table 1: Posterior coefficients estimates

Term	Estimate	Est.Error	95 % CI	
			Lower	Upper
Intercept	-0.153	0.536	-1.005	0.617
Stranger	0.829	0.410	0.128	1.470
Friend	2.084	1.248	-0.151	3.975

3.2.5 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 = 25, High = 20), con proporciones de pago bastante diferentes (Low = 58%, High = 83%), con una diferencia en un intervalo de confianza del 95% de 20% 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 0.40 a 1.51 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 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 prometieron pagar.

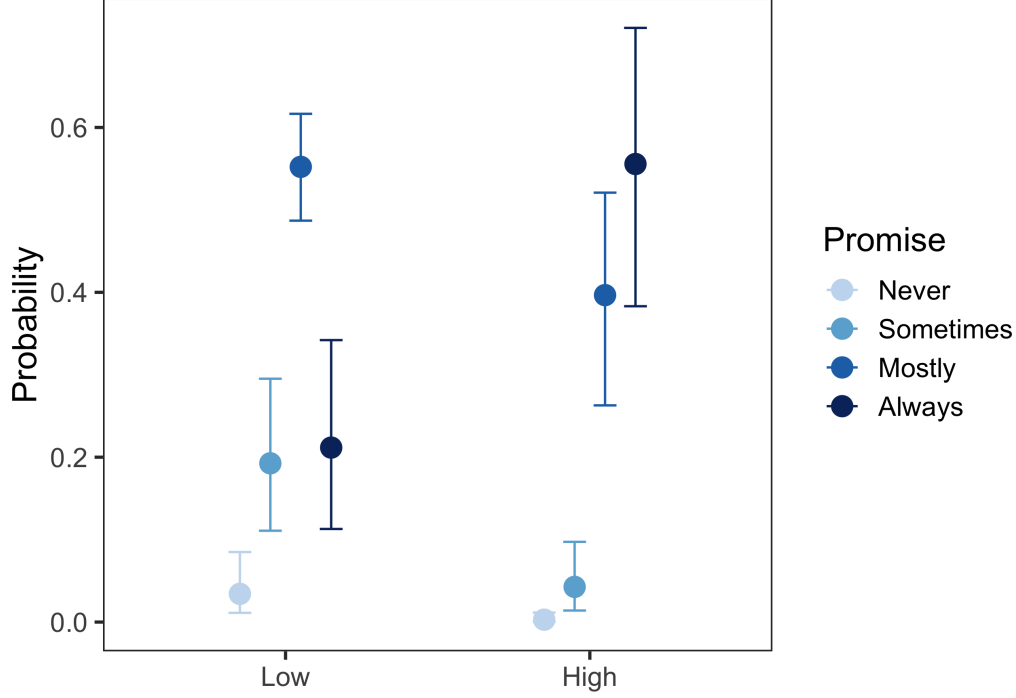


Figure 7: Promise choices by group

4 Discusión

El objetivo del presente estudio fue evaluar el efecto de la cercanía social en el cumplimiento y rompimiento de promesas, así como en la cooperación en un juego de confianza. Nuestros resultados dan evidencia de que la cercanía social nula incrementa la probabilidad de romper la promesa y que ésta va disminuyendo monotónicamente conforme la cercanía social va aumentando. Asimismo, a medida que la cercanía social se incrementa, también lo hace la cooperación; la cercanía social *alta* tiene un efecto en las decisiones que supera inclusive al que tienen todas las demás condiciones experimentales. Probablemente, el hallazgo más interesante fue que cuando los sujetos expresan alto compromiso de cooperar -a través de prometer que *siempre* pagaran la mitad-, la probabilidad de que cumplan es también bastante alta, sin embargo, sí existe una proporción de transgresiones que según nuestras inferencias está entre 4 - 38% con el 95% de certeza.

A nuestro conocimiento, se trata del primer estudio experimental que incorpora la cercanía social como predictor de la decisión de romper las promesas con compañeros socialmente relevantes. En los estudios anteriores los participantes permanecen anónimos durante el curso de las tareas (Baumgartner et al., 2009), las mediciones para evaluar transgresiones son de autorreporte (DePaulo et al., 2004; DePaulo & Kashy, 1998), no cuantifican directamente el rompimiento de las promesas (Charness & Dufwenberg, 2006; Vanberg, 2008), o como discutiremos más adelante, sus medidas de cercanía social pueden ser heterogéneas (Glaeser, Laibson, Szeinkman, & Soutter, 2000).

Nuestro experimento es probablemente el primero en incluir un compañero sin cercanía social

(*computadora*), nuestros datos dan evidencia de que el mero hecho de considerar que se juega con un humano disminuye la probabilidad de romper la promesa. Es aproximadamente 7.5 veces más probable romper la promesa a la computadora que al extraño, a pesar de que no se conoció a dicho compañero. En nuestro estudio, la cercanía social nula es importante porque permite explorar la motivación *intrínseca* para cumplir las promesas, si los humanos mantienen su palabra porque es lo moralmente correcto, anticiparíamos que la probabilidad de romper la promesa fuera muy baja en los tres compañeros y, particularmente, que la inferencia respecto a la probabilidad de romper la promesa a la computadora incluyera el cero.

La probabilidad de romper la promesa sin cercanía social, aunque es baja, es claramente mayor que cero, lo que proporciona evidencia que contradice -al menos en grado- la hipótesis de la motivación *intrínseca*. No podemos excluir la posibilidad que los sujetos cumplan las promesas principalmente motivados moralmente, porque aún tratándose de la computadora los sujetos cumplen en gran proporción, pero sí podemos excluir la posibilidad de que esa sea su única motivación.

Por otra parte, la cercanía social alta nos permite explorar la motivación *instrumental* para cumplir promesas. En nuestro estudio, es casi 9 veces más probable romper la promesa al extraño que al amigo; si los seres humanos mantienen su palabra para facilitar la cooperación en el futuro, anticiparíamos que ante un compañero de cercanía social alta -con el que tienen mucha posibilidad de cooperar, inclusive posterior al experimento- hubiera menos probabilidad de romper la promesa. De hecho, en nuestro estudio no se rompió la promesa a los amigos en ninguna ocasión, por lo que la inferencia en este caso resulta obvia.

Nuestros hallazgos podrían explicarse por la norma social denominada *cooperación condicional*, que indica que la creencia de que otras personas cooperan en altos niveles induce también altos niveles de cooperación (Fehr & Schurtenberger, 2018). Así, la falta de certeza respecto a las decisiones que tomaría la computadora, podría explicar la probabilidad de romper la promesa que existe hacia este compañero. De forma similar, la información con la que los sujetos cuentan respecto al comportamiento del amigo -inclusive antes del experimento-, puede explicar los altos niveles de cooperación y cumplimiento de promesas hacia este compañero.

Respecto a la cooperación, no es la primera investigación en que se evalúa el efecto de la cercanía social, sin embargo, la naturaleza de la manipulación experimental ha sido diversa. Por ejemplo, en un estudio se evaluó la contribución en un juego de bienes públicos durante varios ensayos en un grupo conformado por los mismos individuos (*partners*), en comparación con otro grupo de sujetos diferentes para cada ensayo (*strangers*) (Croson, 1996). Se suponía que en la condición de *partners* la cooperación sería mayor por tratarse de un grupo estable en comparación con el grupo de *strangers*. El estudio mencionado no da detalles respecto al reclutamiento de los participantes, por lo que podríamos asumir que aún en la condición de *partners* se trata de individuos que no se consideran cercanos socialmente.

En otro experimento se reportó un incremento de la cooperación en un juego de confianza en parejas cuando los individuos son cercanos socialmente (Glaeser et al., 2000), en dicho estudio los sujetos se conocen entre sí y los investigadores midieron cuidadosamente varias variables respecto a su conexión social, sin embargo se permitió que algunos individuos que

llegaban juntos realizaran la tarea entre ellos y otros fueron emparejados con un criterio poco claro. Lo anterior permite que la cercanía social entre las parejas sea heterogénea, permite que haya relaciones de tipo romántico y asume que dos individuos que llegan juntos a clase se consideran cercanos entre sí. También, por tratarse de un estudio de *one-shot*, excluye la posibilidad de evaluar cómo un mismo sujeto varía su comportamiento en función de diferentes niveles de cercanía social.

Para tratar de homogeneizar la cercanía social entre nuestros sujetos y sus acompañantes, los participantes realizaron la tarea con un amigo del mismo sexo considerado cercano por ellos mismos, en ningún caso se trató de un familiar y, aunque comprendemos que puede haber parejas del mismo sexo y que los sujetos pudieron haber ocultado su relación, se les enfatizó que para formar parte del estudio su acompañante no podría ser su pareja romántica y/o sexual. Además, el análisis Bayesiano de la escala IOS, mostró evidencia en favor de que el grado de cercanía social entre 30 de nuestros sujetos y sus amigos fue prácticamente el mismo. Si bien los resultados de los dos estudios mencionados están prácticamente en la misma dirección que los hallazgos de nuestro experimento, establecer la similitud en las magnitudes de los efectos no es posible por la diferencia en los diseños, las tareas realizadas y los procedimientos estadísticos empleados.

Un aporte metodológico de nuestro estudio es que, en la literatura de promesas, se trata del primero en utilizar herramientas de inferencia Bayesiana: *modelos jerárquicos* y *modelos acumulativos*. Los *modelos jerárquicos* (o multinivel), nos permitieron modelar el efecto de las condiciones experimentales en la respuesta individual, estos modelos asumen que las observaciones de nivel inferior (e.g. las decisiones de cada individuo), están anidadas en unidades de nivel superior (e.g. los sujetos individuales), los diseños intrasujetos tradicionalmente se han analizado con AN(C)OVA de medidas repetidas, sin embargo, los *modelos jerárquicos* conceden las ventajas de lidiar naturalmente con datos desbalanceados, incluir predictores de naturaleza categórica y/o numérica, incorporar de manera explícita la variabilidad individual, entre otras (Bürkner, 2017; Gelman & Hill, 2006; McElreath, 2018; Vuorre & Bolger, 2018).

En nuestro caso, como se mostró en la Figura 5, el *modelo jerárquico* nos permitió captar de manera muy precisa cómo afectaron las condiciones experimentales las decisiones de cada sujeto y las diferencias entre ellos. Por ejemplo, hay sujetos evidentemente *instrumentales* como el número 8 o 12 que son muy sensibles a la identidad de sus inversores y modifican su cooperación en virtud de la cercanía social. Al mismo tiempo existen sujetos notablemente *intrínsecos* como el 9 y el 10 que todo el tiempo cooperan independientemente de quiénes sean sus compañeros. El *modelo jerárquico* incluye naturalmente esta información para la estimación de los efectos poblacionales, que de no considerarse daría lugar a imprecisiones en la inferencia.

Por otra parte, los *modelos acumulativos* nos permitieron modelar la fuerza del compromiso expresado a través de las promesas. Un problema muy frecuente que se ha señalado recientemente, es que analizar datos ordinales con métodos que asumen que las observaciones son *métricas* puede llevar a graves errores de inferencia (Liddell & Kruschke, 2018). Los *modelos acumulativos* asumen que las respuestas observadas provienen de la categorización de una variable latente continua (Bürkner & Vuorre, 2019), por lo que en nuestro estudio pudimos identificar que los grupos con distintas tasas de pago, que en otros artículos se

han llamado de *deshonestos* y *honestos*, difieren en el compromiso que expresan a través de sus promesas. Considerando que las herramientas para realizar los modelos acumulativos son relativamente recientes, sería valioso realizar la replicación de estudios publicados (por ejemplo, Baumgartner et al., 2009, 2013) considerando las nuevas alternativas de inferencia.

Aunque los *modelos jerárquicos* y *acumulativos* no son herramientas exclusivamente de la inferencia Bayesiana, su aplicación desde dicha aproximación representa varias ventajas en comparación con la aproximación frecuentista. Problemas clásicos de la anterior como son: las múltiples comparaciones, la decisión de cuándo detenerse en la recolección de sujetos (stopping rule), o el uso de comparaciones planeadas versus *post hoc*, no son factores que afecten el apoyo que brindan los datos a las hipótesis desde la aproximación Bayesiana (Dienes, 2011). Cómo las inferencias y el contraste de hipótesis desde la aproximación que utilizamos, hacen uso de la probabilidad para expresar la plausibilidad de las teorías, en nuestro caso pudimos obtener razones de evidencia respecto a la medida en que nuestros datos apoyan nuestras hipótesis, independientemente de si se éstas fueron propuestas *a priori* o *post hoc*.

Una limitación del presente trabajo es el uso de recompensas hipotéticas en comparación con recompensas reales. Un estudio relativamente reciente sobre toma de decisiones, reportó que existe más aversión a la pérdida cuando los sujetos tienen recompensas monetarias reales en comparación con recompensas hipotéticas en una tarea de riesgo (Xu et al., 2016). Sin embargo, existen otros trabajos que no reportan diferencias entre el uso de ambos tipos de recompensas en tareas de autocontrol, descuento temporal y descuento social (Johnson & Bickel, 2006; Locey, Jones, & Rachlin, 2011; Rachlin & Jones, 2006). Asimismo, se puede argumentar que los resultados tienen congruencia teórica (Fehr & Schurtenberger, 2018) y están en la misma dirección que otros trabajos que utilizan dinero real (Croson, 1996; Glaeser et al., 2000), por lo que no hay muchas razones para esperar que otro tipo de recompensas hubieran modificado nuestros resultados.

Otra posible limitación es el uso de ensayos múltiples con cada compañero de juego. Aunque la decisión de utilizar medidas repetidas durante nuestro diseño atiende más al propósito de reducir fuentes de error extrañas e incrementar la validez interna (Maxwell, Delaney, & Ken, 2018), para tener más claridad respecto a la diferencia en las motivaciones *intrínseca* e *instrumental*, los estudios podrían beneficiarse de sólo incluir un ensayo para cada compañero. Jugar un *one-shot* con cada compañero evita la posibilidad de que los múltiples ensayos pudieran generar la creencia de que la computadora también puede variar su comportamiento de acuerdo las decisiones del *fideicomisario*.

5 Conclusiones

Los sujetos cumplen sus promesas por una combinación de motivaciones entre la *intrínseca* y la *instrumental*. Son predominantemente *nice guys* porque mantienen su palabra inclusive ante inversores con nula cercanía social. Sin embargo, su integridad dista de ser perfecta, ya que existe una proporción pequeña, pero confiablemente mayor que cero, de traiciones cometidas hacia este socio. Tal parece ser, que si los sujetos no pueden asegurar que su inversor será un *potential cooperator*, la deshonestidad puede emerger en la forma de promesas rotas. Las predicciones teóricas de la norma social *cooperación condicional* enmarcan el

hallazgo de que la cercanía social disminuya la probabilidad de romper la promesa y aumente la cooperación. De la misma forma que sugiere la correspondencia entre el compromiso expresado a través de las promesas y la conducta posterior.

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>
- Croson, R. T. (1996). Partners and strangers revisited. *Economics Letters*, 53(1), 25–32. [https://doi.org/10.1016/S0165-1765\(97\)82136-2](https://doi.org/10.1016/S0165-1765(97)82136-2)
- 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>
- Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, 6(3), 274–290. <https://doi.org/10.1177/1745691611406920>

- Fareri, D. S., & Delgado, M. R. (2014). Differential reward responses during competition against in- and out-of-network others. *Social Cognitive and Affective Neuroscience*, 9(4), 412–420. <https://doi.org/10.1093/scan/nst006>
- 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>
- Glaeser, E., Laibson, D., Scainkman, J., & Soutter, C. (2000). Measuring Trust. *The Quarterly Journal of Economics*, 115(3), 811–846.
- 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>
- Johnson, M. W., & Bickel, W. K. (2006). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior*, 77(2), 129–146. <https://doi.org/10.1901/jeab.2002.77-129>
- 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>
- Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, 79(November 2017), 328–348. <https://doi.org/10.1016/j.jesp.2018.08.009>
- Locey, M. L., Jones, B. A., & Rachlin, H. (2011). Real and hypothetical rewards in self-control and social discounting. *Judgment and Decision Making*, 6(6), 552–564.
- Maxwell, S. E., Delaney, H. D., & Ken, K. (2018). *Designing experiments and analyzing data: a model comparison perspective* (3rd edition, pp. 611–612). New York: Routledge.

- 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>
- McElreath, R. (2018). *Statistical rethinking: A bayesian course with examples in R and stan*. <https://doi.org/10.1201/9781315372495>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., . . . Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2. <https://doi.org/10.3389/neuro.11.010.2008>
- Rachlin, H., & Jones, B. (2006). Social Discounting. *Psychological Science*, 17(4), 283–286.
- 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
- Sip, K. E., Smith, D. V., Porcelli, A. J., Kar, K., & Delgado, M. R. (2015). Social closeness and feedback modulate susceptibility to the framing effect. *Social Neuroscience*, 10(1), 35–45. <https://doi.org/10.1080/17470919.2014.944316>
- 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>
- Vuorre, M., & Bolger, N. (2018). Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience. *Behavior Research Methods*, 50(5), 2125–2143. <https://doi.org/10.3758/s13428-017-0980-9>
- Xu, S., Pan, Y., Wang, Y., Spaeth, A. M., Qu, Z., & Rao, H. (2016). Real and hypothetical monetary rewards modulate risk taking in the brain. *Scientific Reports*, 6. <https://doi.org/10.1038/srep29520>