UNIVERSIDAD
DE GRANADA

TESIS DOCTORAL

Learning to Individuate versus Categorize People: The Role of Attention in Social Learning

Doctoranda
Maïka Ysaora Telga

Director
Juan Lupiáñez Castillo

Departamento de Psicología Experimental

Programa de Doctorado en Psicología

Octubre 2019

Editor: Universidad de Granada. Tesis Doctorales
Autor: Maïka Ysaora Telga
ISBN: 978-84-1306-414-7
URI: http://hdl.handle.net/10481/58821

À mes parents.

ACKNOWLEDGMENTS

Los últimos cuatro años han pasado con una velocidad increíble. Parece que fue ayer cuando firmé el contrato, ilusionada por empezar el doctorado, y (casi) sin darme cuenta, hoy me toca con mucha ilusión y sin duda, emoción, agradecer a todas las personas que me acompañaron en esta aventura.

Gracias Juan por ofrecerme la mejor supervisión con la que podía soñar. Por supuesto, por la inquietud y la curiosidad científica, por poner tu mente a nuestra disposición cual siervo, y por contagiarnos con tu entusiasmo. Pero también por estas cualidades que hacen que sea un privilegio trabajar contigo, no solo como científico, pero sencillamente, como persona. Por la humildad, la honradez, el apoyo técnico, moral y emocional. Por las reflexiones sobre jamón, maíz y lenguaje inclusivo. Porque ni el índice *h* alto ni el gran número de citas te hacen justicia.

Gracias Elena por marcar el camino y por haber estado tan disponible en los últimos cuatro años. A Sole por confiar en mí incluso antes de la tesis, por literalmente, iniciarme a la investigación y por inspirarme con una trayectoria investigadora ejemplar. A Rosa por transmitir tanto como docente y por ser el ejemplo por antonomasia de alta *communion* y *competence*.

Thank you Kerry for your warm welcome, for offering me a different perspective on my work, for allowing me to get involved in your incredible research, and for treating me as if I had always been part of your lab. Thank you Elysia, Francine, Xia, Channel and Madeline for making me feel part of the group from the moment I arrived. Thank you Michal and Amanda for your kindness and for making Friday afternoons easier.

Gracias a las más de 1200 personas que participaron en mis estudios, incluidos o no en este trabajo, porque desde luego, sin ellas, hoy no habría tesis. En particular, gracias a mis adorables participantes mayores por creer que voy a cambiar el mundo y hacer que la recogida de datos sea un verdadero placer. En especial, gracias a Pepe por hacer posible toda una serie experimental, y ayudar incluso fuera del laboratorio.

Gracias a las personas que contribuyeron de forma más indirecta a que esta tesis saliera adelante, en especial a las alumnas de TFGs y TFMs que me acompañaron durante un trocito de camino, a Guillermo, Rosa y Eulalia por vuestra paciencia.

Many thanks to Teresa Garcia-Marques, Ana Guinote, María Ruz, Miguel Vadillo and Guillermo Willis for accepting to evaluate this dissertation.

Cependant, l'aventure du doctorat a commencé bien avant ces quatre dernières années. Dans mon cas précis, en première année de licence, lors d'un des cours magistraux de Jean-Claude Croizet, que je ne peux que remercier pour m'avoir, sans le savoir, encouragée à m'embarquer dans cette aventure qui m'a permis de connaître des personnes formidables.

Merci à Jordy et Cécile d'être encore présents dix ans après le « MP Crew ».
Merci à mes amies de licence pour avoir fait de ces années une époque inoubliable, particulièrement à Nadia, Julie, Marlène et Marie. Et bien entendu, à Adé pour les matchs de hand, les glaces Ben & Jerry, et ton soutien pendant toutes ces années. A Louis pour les grandes et les « petites pensées », le thatanisme, et l'espoir d'un Skype.
Gracias a mis compañeras y compañeros de máster por el apoyo mutuo tan necesario durante estos años, particularmente a Lucía, Alexandra y Ana.

Gracias a los grupos de Neurociencia Cognitiva y de Psicología Social por proporcionar un ambiente de trabajo maravilloso, por inspirar, colaborar y compartir. A las y los compañeras/os del CIMCYC por estar siempre presentes, en lo académico y lo social. En especial, a Nuria y Jorge por alegrarme más de un día. A Ángela por los cafés de desahogo y apoyo mutuo. A Ana Paqui y Paloma por compartir la frustración ante el mundo de los trámites administrativos, penas y alegrías, en especial en los últimos momentos pre-depósito. A Ana Paqui por los abrazos tanto en los momentos buenos como en los momentos menos buenos. A Paloma por cantar conmigo de Alex Ubago a Ed Sheeran. A Fernando por mantener la amistad en la distancia, por (creer que se puede) arreglarlo todo con un matecito. A Enzo, Vicente, Conchi, Rafa y Greta por soportarme en la intimidad del despacho, por estar pendientes y por el apoyo pre-tesis.

Gracias Paqui por enseñarme a enseñar, y más allá de eso, por estar siempre disponible para escucharme y compartir conmigo tu experiencia y consejos. A Eli por ser un modelo tan inspirador de investigadora, mujer, y persona.

Gracias a las personas que han marcado estos últimos años de una forma tan significativa, que con toda la sinceridad, no sé qué hubiera sido de mí sin ellas.

Gracias Mafer y Cristina por las constantes muestras de cariño, por las tardes de juego de mesa y los viajes lejos del laboratorio. A Lara y Alba por verme mucho mejor que de lo que soy, por animarme, apoyarme y aconsejarme.

Gracias a las y los Alcalá, Barruz, Martín y Monzón por ser la mejor familia (política) con la que podía soñar. Por vuestro apoyo y cariño en estos últimos años, de corazón, gracias.

Gracias Carlos por el apoyo en todo momento, por estar ahí, incluso en la lejanía. A Javi, por haberme aguantado con una paciencia sin límite, por subirme los ánimos, por confrontar mis ideas con tu *manspreading* y *mansplaining*. A Itsaso por la complicidad dentro y fuera del CIMCYC, por ayudarme a deshacer mis sesgos, por conocerme tan bien y quererme igualmente. A María por apoyarme en las etapas complicadas, por escucharme y entenderme, por compartir el gusto por las cosas *bien* hechas (y la frustración cuando no lo están). A Juan Eloy por compartir tantísimo conmigo, por las historias normales y paranormales, por las tartas pasadas y las que quedan, y por ofrecerme siempre tu punto de vista optimista, pase lo que pase. A Carmen por esta relación a caballo entre amistad y amor platónico, por reírte de mí, por dejarme reírme de ti, y sobre todo, por reírnos juntas ante los pequeños y grandes dramas de la vida.

También he de agradecer a estas personas cuyo apoyo transciende el doctorado, habiendo contribuido a convertirme en la persona que soy a día de hoy.

Merci à Fanny d'être si proche tout en étant si loin, d'être encore et toujours l'une des rares personnes avec qui je peux *tout* partager, avec ou sans (virgin) mojito, repas chinois ou Domino's Pizza. Merci d'être la générosité et la bonté incarnées, d'être tout simplement l'une des plus belles personnes que cette terre a portée. À Vincent pour ces vingt années d'amitié, tout juste cette année, qui nous ont permis

de nous voir grandir, de changer ensemble ou séparément. Pour la guitare, les (terribles) marches sportives, Grenade. Pour tout ce qui a fait de ces vingt années ce qu'elles sont. Et pour la certitude de pouvoir encore compter sur toi durant les vingt prochaines années.

Merci à mes parents, à Yorick et à Alison pour l'amour inconditionnel et le soutien permanent. Merci d'être toujours disponibles même à six heures de différence et plus de 6000 km. Merci de croire en moi plus que quiconque et d'avoir tout mis en œuvre pour que je sois où j'en suis aujourd'hui. Merci pour les conseils et le soutien lors des décisions difficiles. Merci de m'avoir, en dépit de tout, laissée partir.

Gracias José Andrés por seguir uniendo contrarios. Por poner caos en mi orden y dejarme poner orden en tu caos. Por aportar locura a mi cordura, y cordura a mi locura. Por ser mi razón y mi perdición. Por estar cerca en la lejanía. Por entenderme cuando yo no me entiendo. Gracias por tanto y más aún, pero sobre todo, gracias por dejarme entrar en tu mundo y ocupar un lugar tan especial en el mío.

El camino no ha sido fácil, entre los artículos rechazados, los congresos a los que no he ido, las estancias que no he hecho, o el ser "demasiado social" o "demasiado cognitiva". Pero a pesar de ello y gracias a vuestro apoyo y cariño, me alegra poder decir que ha sido un camino que he disfrutado.

CONTENTS

TABLES AND FIGURES

ABSTRACTS

Abstract

Faces are essential stimuli in social perception, as they comprise a wealth of information regarding a person's individual identity, social category, emotional state, gaze, and so on (Hugenberg & Wilson, 2013). This information is readily encoded at early stages of face processing, and further integrated to make sense of others. The information extracted from faces substantially influences impressions formation processes, especially at zero acquaintance, when nothing is known about the target. However, all faces are not equally attended and the strategies used to form impressions of others lie on a continuum stretching from social categorization to individuation (Fiske & Neuberg, 1990).

Social categorization consists of using information diagnostic of a person's social category to make inferences about him or her. For instance, skin color, lips shape or hair texture are informative of a person's race. Once these features have been encoded, the target person is classified in a social category and the knowledge associated with this category, based for instance on ethnic stereotypes, is used to make inferences about this particular individual. This strategy allows perceivers to integrate and organize efficiently the diversity of information that may be extracted from social stimuli. Because of its remarkable resources-saving function, social categorization is the default strategy to make sense of others (Macrae & Bodenhausen, 2000). Alternatively, perceivers may further their level of analysis beyond social categories by using the idiosyncratic attributes of a person to predict his or her behaviors. However, as compared to social categorization processes, individuation requires increased attention to integrate the variety of individual attributes contributing to the person's uniqueness (Gilbert & Hixon, 1991). Therefore, perceivers do not engage in individuation processes unless they are both *cognitively able* and *motivated* to acquire an individuated judgment about a target person (Pendry, 1998). Such motivation may arise from different situational and personal variables. For instance, if a perceiver is outcome-dependent on understanding a target (e.g., the perceiver is economically rewarded according to the accuracy of his or her predictions), he or she will likely engage in individuation processes (Neuberg & Fiske, 1987). In that sense, perceivers are often perceived as "motivated tacticians" who possess different cognitive strategies from which they chose the one that best fits their interests (Fiske & Taylor, 1991). Cognitive efficiency is therefore achieved by means of a trade-off between cognitive cost and goals attainment.

Beyond the cognitive economy, social categorization is associated with a series of attentional and behavioral biases with important consequences on social judgments. For instance, categorizing a target as an ingroup member (i.e., a person who belongs to the same social category as oneself) may promote a more positive attitude towards this person than toward an outgroup member, even at zero acquaintance. This effect, known as ingroup favoritism, may be related to trivial misattributions of traits, but also to more serious intergroup discriminations (Brewer, 1979, 2001). Another consequence of social categorization is reflected in the outgroup homogeneity effect, that is, the perception of higher intragroup similarity among outgroup than among ingroup members. Such biased perception may be related to impaired learning about outgroup members' traits (Park & Rothbart, 1982). Broadly speaking, both effects reflect how overgeneralizations based on social group membership may negatively impact social relationships. Understanding the specific circumstances promoting individuation or social categorization strategies in social perception is crucial to reduce the negative implications of category-based judgments, including stereotypes and prejudice.

Considering this framework, the aim of the present dissertation was to evaluate the impact of perceivers' motivational, emotional and cognitive states on the differential use of social categorization and individuation strategies, and their consequences on social judgments regarding the principal dimensions of impressions formation (i.e., gender, ethnicity, age and emotional expression). To achieve this general goal, we developed an adaptation of the trust game paradigm (Berg, Dickhaut, & McCabe, 1995) allowing us to achieve the following specific goals:

1. To explore impressions formation at zero acquaintance in relation to motivational, cognitive and emotional factors.
2. To analyze how individuation and categorization strategies contribute to update first impressions across repeated interactions in a learning process.
3. To examine the impact of learned associations between specific social categories and group behaviors on the perception of new individuals from these social categories.

With this adaptation of the trust game, we conducted seven experimental series distributed in five chapters in which different cognitive, emotional and motivational factors were manipulated. In Experimental Series 1 and 2, we explored impressions formation processes according to the main dimensions of social categorization, namely gender, ethnicity and age. In Experimental Series 3, we examined the potential effect of social power on social perception. In Experimental Series 4 and 5, we directly

manipulated motivation and cognitive resources to understand their impact on social judgments. Finally, in Experimental Series 6 and 7, we analyzed the relationship between emotions and social perception.

In our adaptation of the trust game, participants had to predict the cooperative behaviors of unfamiliar partners to behave accordingly, cooperating with equitable partners, and not cooperating with non-equitable partners. Moreover, accurate predictions in the task were rewarded with real monetary outcomes. The task was divided in three phases with the following specific manipulations:

First, in the baseline, participants played with partners from two social categories, all partners being equitable in half of the trials and non-equitable in the other half. This phase allowed us to verify whether participants were spontaneously biased to cooperate more with one of the two social groups.

Second, in the learning phase, we manipulated the group behavior by associating the two social categories with opposite cooperative behaviors, one being equitable and the other one being non-equitable. Crucially, within each social category, a small proportion of individuals were inconsistent regarding their group, displaying a pattern of cooperative behavior opposite to the group behavior. This phase allowed us to explore learning across repeated interactions and more importantly, to verify whether participants were impacted by their partners' category or individual behavior. In this regard, cooperation decisions with inconsistent partners were informative of the strategy used. A categorization strategy would be reflected in a similar pattern of cooperation with consistent and inconsistent individual within the same social group. Conversely, an individuation strategy would be reflected in opposite patterns of cooperation with consistent and inconsistent individuals within the same social group.

Finally, in the transfer phase, participants were presented with new individuals from the two social categories manipulated, but with whom they lacked individual experience, all of them being equitable in half of the trials. This final phase allowed us to verify whether the associations established in the learning phase between group behavior and social categories would be applied to new individuals in a categorical way.

Across seven experimental series, we never observed a complete categorization pattern with human partners, as participants did not apply the category-related knowledge to make predictions about inconsistent individuals. This finding suggests that the

outcome-dependency intrinsic to the game settings may have motivated participants to predict their partners' behaviors as accurately and as individually as possible. Moreover, and importantly, the few hints of categorical thinking found in the present dissertation are better accounted by motivational than by cognitive factors. Simply stated, participants rely on categorical information only when their motivation to individuate was reduced, but never because of the cognitive cost of individuation processes, as attempts to individuate people were observed even in situations where individuation was impossible. Altogether, these results challenge the trade-off approach of impressions formation theories. Moreover, and importantly, they suggest that with the adequate motivation, perceivers may fully explore their cognitive capacity to individuate, offering an encouraging perspective on motivation-oriented interventions to reduce over-categorization and its negative consequences.

Resumen

Las caras son estímulos claves para la percepción social, ya que nos proporcionan una gran cantidad de información relativa a la identidad, categoría social, mirada o estado emocional de una persona, entre otros muchos aspectos (Hugenberg & Wilson, 2013). Esta información se codifica en etapas tempranas del procesamiento facial, y a continuación es integrada para conformar nuestra percepción de los demás. La información extraída de las caras es determinante en el proceso de formación de impresiones, especialmente en las primeras interacciones con una persona desconocida, cuando aún no se sabe nada acerca de ella. Sin embargo, numerosos factores influyen en el procesamiento de la información social. Las estrategias que usamos para formar una impresión acerca de una persona desconocida se distribuyen en un continuo que va desde la categorización social hasta la individualización (Fiske & Neuberg, 1990).

La categorización social consiste en utilizar la información de la categoría social a la que pertenece una persona para hacer inferencias sobre ella. Por ejemplo, el color de piel, la forma de los labios o la textura del pelo son características informativas acerca de la raza de una persona. Una vez que estas características han sido codificadas, la persona objetivo es clasificada dentro de su correspondiente grupo social y el conocimiento asociado a dicha categoría (por ejemplo, a partir de los estereotipos étnicos) puede usarse para hacer inferencias sobre ella en particular. Esta estrategia nos permite integrar y organizar de forma efectiva la gran cantidad de información que se deriva de los estímulos sociales. Cumpliendo una función de economía cognitiva, la categorización social es la estrategia predominante para hacer inferencia sobre personas desconocidas (Macrae & Bodenhausen, 2000). Sin embargo, de forma alternativa, podemos optar por un nivel de análisis más profundo e independiente de las categorías sociales, usando los atributos idiosincráticos de una persona para determinar su comportamiento. Ahora bien, en comparación con el proceso de categorización social, la individualización requiere más recursos atencionales por la necesidad de analizar e integrar la variedad de atributos que hacen a las personas únicas (Gilbert & Hixon, 1991). Por lo tanto, no empleamos una estrategia de individualización a no ser que dispongamos de los *recursos cognitivos* y la debida *motivación* para adquirir un juicio individual sobre una persona objetivo (Pendry, 1998). Dicha motivación surge de distintas variables situacionales y personales. Por ejemplo, si obtener una recompensa o un resultado valorado positivamente depende de nuestra relación con una determinada persona (por ejemplo, si recibimos una recompensa

económica en función de la precisión de nuestras predicciones sobre la persona objetivo), entonces seremos más propensas/os a usar una estrategia de individualización (Neuberg & Fiske, 1987). En este sentido, somos "estrategas motivadas/os" disponiendo de varias habilidades cognitivas, y eligiendo, según el contexto, la estrategia que mejor encaje con nuestros intereses y objetivos (Fiske & Taylor, 1991). De acuerdo a esta aproximación basada en la eficiencia cognitiva, un uso óptimo y eficiente de nuestros recursos cognitivos se consigue encontrando un equilibrio entre coste cognitivo y objetivos.

Más allá de la economía cognitiva, la categorización social está asociada a una serie de sesgos atencionales y comportamentales con importantes consecuencias en los juicios sociales. Por ejemplo, categorizar a una persona como miembro del endogrupo (es decir, perteneciente a la misma categoría social que una/o misma/o) puede promover una actitud más positiva hacía dicha persona en comparación con una persona del exogrupo. Este efecto, conocido como favoritismo endogrupal, se relaciona con atribuciones erróneas de rasgo sin mayor consecuencia, pero también con comportamientos discriminatorios de cierta gravedad (Brewer, 1979, 2001). Otra consecuencia de la categorización social se ve reflejada en el efecto de homogeneidad exogrupal, esto es, la percepción de que los miembros del exogrupo son más similares entre sí que los miembros del endogrupo. Dicho sesgo tiene como consecuencia, por ejemplo, un peor aprendizaje acerca de los rasgos individuales de los miembros del exogrupo (Park & Rothbart, 1982). En términos generales, ambos efectos reflejan cómo atribuciones basadas en la pertenencia a un determinado grupo social pueden afectar negativamente a las relaciones sociales. Comprender las circunstancias específicas que promueven el uso de una estrategia de individualización o de categorización social es crucial para reducir los efectos negativos de los juicios superficiales basados en pertenencia grupal, incluyendo estereotipos y prejuicios.

Considerando este marco teórico, el objetivo de la presente tesis doctoral fue evaluar el impacto de factores motivacionales, emocionales y cognitivos sobre el uso de las estrategias de categorización e individualización, y sus consecuencias en los juicios sociales basados en las principales dimensiones de formación de impresiones (es decir, género, etnia, edad y expresión emocional). A partir de este objetivo general, desarrollamos una adaptación del paradigma del juego de confianza, (Berg et al., 1995) con el fin alcanzar los siguientes objetivos específicos:

1. Explorar los procesos de formación de impresiones con personas completamente desconocidas, en relación a factores motivacionales, cognitivos y emocionales.

2. Analizar cómo las estrategias de individualización y categorización contribuyen a la actualización de las primeras impresiones a través de repetidas interacciones durante un proceso de aprendizaje.
3. Examinar el impacto de las asociaciones aprendidas entre categorías sociales y conducta del grupo a la hora de predecir el comportamiento de individuos de dichas categorías sociales.

Con esta adaptación del juego de la confianza, realizamos siete series experimentales distribuidas en cinco capítulos en las cuales manipulamos diferentes factores cognitivos, emocionales y motivacionales. En las Series Experimentales 1 y 2, exploramos el proceso de formación de impresiones a partir de las principales dimensiones de la categorización social, es decir, género, etnia y edad. En la Serie Experimental 3, examinamos el potencial efecto del poder en los juicios sociales. En las Series Experimentales 4 y 5, manipulamos directamente motivación y recursos cognitivos para entender cómo impactan la percepción de las y los demás. Finalmente, en las Series Experimentales 6 y 7, analizamos la relación entre emociones y percepción social.

En nuestra adaptación del juego de la confianza, las/os participantes tenían que predecir el comportamiento cooperativo de compañeras/os de juego desconocidas/os para cooperar con las/os compañeras/os equitativas/os, y no cooperar con las/os compañeras/os no equitativas/os. Además, las/os participantes recibieron dinero real de acuerdo con su desempeño en la tarea. Dicha tarea se dividía en tres fases con las siguientes manipulaciones:

Primero, en la línea base, las/os participantes interactuaban con compañeras/os de juego de dos categorías sociales, todas/os siendo equitativas/os en la mitad de los ensayos y no equitativas/os en la otra mitad. Esta fase nos permitió verificar si las/os participantes tenían una tendencia de cooperación sesgada, cooperando en mayor medida con alguno de los dos grupos sociales.

En segundo lugar, en la fase de aprendizaje, manipulamos el comportamiento de las/os compañeras/os de juego a nivel de grupo, asociando las dos categorías sociales con comportamientos opuestos. Concretamente, un grupo social era equitativo mientras el otro grupo era no equitativo. Crucialmente, dentro de cada categoría social, una pequeña proporción de individuos eran inconsistentes con su grupo, mostrando un patrón de cooperación opuesto al de su grupo. Esta fase nos permitió explorar el aprendizaje a través de repetidas interacciones con las/os mismas/os compañeras/os de juego y, lo que es más importante, examinar si el aprendizaje de las/os participantes estaba afectado por la información categórica o individual. En efecto, las decisiones de cooperación de las/os participantes con las/os compañeras/os de juego inconsistentes reflejaba el tipo de estrategia empleada. Una estrategia de categorización se vería reflejada en un patrón de cooperación similar con compañeras/os consistentes e inconsistentes dentro del mismo grupo social. Al contrario, una estrategia de individualización se vería reflejada en un patrón de cooperación opuesto con las/os compañeras/os consistentes e inconsistentes dentro del mismo grupo social.

Finalmente, en la fase de transferencia, las/os participantes eran expuestas/os a nuevos individuos de las categorías sociales previamente manipuladas, siendo estas/os compañeras/os equitativas/os en la mitad de los ensayos. Esta fase final nos permitió comprobar si las asociaciones establecidas en la fase de aprendizaje entre el comportamiento grupal y las categorías sociales se aplicarían a nuevos individuos en un proceso categórico.

Cabe destacar que, a través de siete series experimentales, no se observó nunca un patrón de categorización completo con compañeras/os humanas/os, dado que las/os participantes no utilizaron la información relativa al grupo para hacer predicciones sobre los individuos inconsistentes. Este hallazgo sugiere que las características de la tarea asociada a recompensas económicas pueden haber motivado que las/os participantes predijeran el comportamiento de sus compañeras/os de juego de la forma más precisa e individual posible. Además, y más importante, los pocos indicios de categorización observados en la presente tesis doctoral están mejor explicados por factores motivacionales que cognitivos. En otras palabras, las/os participantes usaron la información categórica solo cuando su motivación para individualizar era reducida, pero nunca como consecuencia del coste cognitivo de la individualización. En efecto, observamos intentos por parte de las/os participantes para de individualizar a sus compañeras/os de juego incluso en situaciones en las cuales la individualización era imposible. En conjunto, estos resultados cuestionan la aproximación basada en eficiencia cognitiva en las teorías de formación de impresiones. Además, y de forma importante, sugieren que con la adecuada motivación, podemos explotar al máximo nuestras habilidades cognitivas para individualizar, ofreciendo una perspectiva esperanzadora en cuanto a intervenciones basadas en motivación para reducir la sobre-categorización y sus consecuencias negativas.

CHAPTER 1. Introduction

> "O'Brien was a large, burly man with a thick neck and a coarse, humorous, brutal face. In spite of his formidable appearance he had a certain charm of manner. He had a trick of resettling his spectacles on his nose which was curiously disarming—in some indefinable way, curiously civilized. It was a gesture which, if anyone had still thought in such terms, might have recalled an eighteenth-century nobleman offering his snuffbox. Winston had seen O'Brien perhaps a dozen times in almost as many years. He felt deeply drawn to him, and not solely because he was intrigued by the contrast between O'Brien's urbane manner and his prize-fighter's physique. Much more it was because of a secretly held belief—or perhaps not even a belief, merely a hope—that O'Brien's political orthodoxy was not perfect. Something in his face suggested it irresistibly. And again, perhaps it was not even unorthodoxy that was written in his face, but simply intelligence. But at any rate he had the appearance of being a person that you could talk to if somehow you could cheat the telescreen and get him alone. Winston had never made the smallest effort to verify this guess: indeed, there was no way of doing so."
>
> George Orwell – *Ninety eighty-four*, p.14

By the end of George Orwell's *Ninety eighty-four* masterpiece, you may have wondered why on earth did Winston trust O'Brien's apparent non-orthodoxy. What facial cues were so informative of his intelligence or trustworthiness? How is it possible that Winston formed such a detailed impression of O'Brien's state of mind, only after a dozen encounters without exchanging any word. These questions come with the revelation that Winston mistakenly judged O'Brien. However, when reading the first description of O'Brien quoted above, you probably did not think that Winston was out of his mind for making such complex inferences with such little ground. In fact, inferring other's cognitive and affective states from simple facial cues is a daily exercise, hopefully associated with more positive outcomes than Winston's. As perceivers, we constantly use the information extracted from facial cues to predict others' internal states. Importantly, like Winston, we often form complex impressions of others with few information.

Notably, in impression formation processes, bodily and facial cues are readily processed to make sense of others. For instance, if somebody gazes at you while smiling, you may infer that this person is interested in establishing contact, and respond by smiling back to communicate your positive disposition to talk. Alternatively, if someone consistently avoids eye contact, you may react by avoidance-related responses as well. Therefore, perceivers' behaviors are determined by (a) the processing of relevant visual cues and (b) the attribution of internal states to others formed on the basis of this perceptual information. Both processes are modulated by top-down factors, potentially resulting in attentional biases and misattributions of traits, which may partially explain Winston's fate.

With the aim of understanding the factors impacting impressions formation in trustworthiness judgments, the first section of this chapter focuses on face processing, reviewing how visual cues, in relation to top-down factors, are integrated to achieve a meaningful perception of people. The second section reviews how, on the basis of the information extracted from faces, perceivers elaborate inferences and predictions about others, which will, in turn, determine their own attitude, decision-making and behaviors. Finally, the last section will focus on trustworthiness judgments, briefly reviewing how factors related to individual identity, group membership or emotional expressions impact trust decisions.

1. Perception of faces

Faces are unarguably essential stimuli in social perception. From the mere presentation of a face, perceivers can readily extract relevant information related to a person's transient state, as for instance his or her emotional state or gaze direction, to more stable features including social categories and individual identity. All this information crucially impacts perceivers' inferences and expectations from targets, which in turn influence their behaviors and the outcomes of the interaction.

1.1. Individual identity

Humans' capacity to recognize a person's identity from his or her face is remarkable. Although faces usually share the same pattern where the principal facial features (i.e., two eyes, one nose and one mouth) are located according to a unique configuration, perceivers are highly capable of attending to the subtler features diagnostic of a target's individual identity to discriminate one face from another. Such ability has generated a large body of research investigating the mechanisms underlying face perception and identity coding, highlighting two perceptual coding mechanisms: adaptive norm-based coding and holistic processing.

In adaptive norm-based coding (Leopold, Rhodes, Muller, & Jeffery, 2005; Rhodes & Jeffery, 2006; Valentine, 1991), faces are considered in a computational space in which each face is coded as a deviation from an average or "norm" situated in the center of the space. Several dimensions of the space (i.e., facial features) are considered to extract the norm. This average is constantly calibrated and updated by experience as perceivers are exposed to new faces, by for instance, adding new dimensions to discriminate between frequently encountered faces (Valentine & Endo, 1992), which makes the process “adaptive”. In addition, being able to code identity information with respect to a norm allows reducing the cognitive cost of face processing as compared to a strategy in which all facial features would be equally attended. Norm-based coding has been extensively investigated by means of face identity aftereffects. Specifically, the underlying idea is that a prolonged exposure to a face (e.g., anti-Dan) shifts the average toward that face, resulting in an adaptive perceptual bias toward the opposite computational face (e.g., Dan). For instance, if anti-Dan is characterized by having thick lips, an aftereffect would be reflected in a facilitation to code faces featuring the computational opposite characteristic, namely, thin lips. In fact, the degree to which new faces are perceived as more similar to Dan’s face after the exposure to anti-Dan reflects the perceptual switch in the perceiver’s average. In face identity aftereffects paradigms, participants are tested on their recognition of a target face after being exposed to a different face. In match trials, the faces from the exposure and test phases lie from opposite identities (i.e., anti-Dan and Dan), while on mismatch trials, the faces from the exposure and the test phases lie from different but not opposite identities (i.e., anti-Dan and Jim). The size of the aftereffect is computed by subtracting accuracy scores on mismatch trials from accuracy scores on match trials such that higher scores reflect a larger shift in average as a result of exposure.

Norm-based coding has been related to identity identification in different instances. Notably, face identity aftereffects are reduced in patients with impaired face perception. For instance, patients with congenital prosopagnosia, a condition characterized by a selective deficit in face recognition, show reduced aftereffects (Palermo, Rivolta, Wilson, & Jeffery, 2011). A similar pattern is observed with autistic patients who present an atypical pattern of face recognition (Ewing, Leach, Pellicano, Jeffery, & Rhodes, 2013; Ewing, Pellicano, & Rhodes, 2013). Moreover, Rhodes, Jeffery, Taylor, Hayward, and Ewing (2014) showed that individual differences in norm-based coding correlated with individual differences in face recognition in neurotypical subjects, supporting the role of adaptive coding in face expertise. Importantly, such correlation was not found with regard to non-face object indicating that the norm-based coding is specific to face instead of a general visual effect. Finally, it is interesting to note that the impact of norm-based coding is found within 250 ms after the presentation of a face (Burkhardt et al., 2010) coinciding with face discrimination time course (Barragan-Jason, Cauchoix,

& Barbeau, 2015). Altogether, these data suggest that perceivers' capacity to adaptively code the facial features that make a particular face different from the norm is crucial for accurate identity recognition.

On the other hand, holistic processing has also been well documented as a specific marker of individual identity recognition (Maurer, Grand, & Mondloch, 2002; Tanaka, Kiefer, & Bukach, 2004). Specifically, unlike non faces stimuli, facial features are integrated simultaneously in a unified representation. The extent to which faces are considered as a meaningful whole instead of a combination of independent facial features has been investigated with, among other paradigms, the composite face illusion. In this task, participants are consecutively presented with two faces whose top and bottom halves may be either aligned or misaligned. Participants are instructed to indicate whether the two top halves presented are from the same or from different faces. Importantly, the illusion arises in the aligned condition: when the top halves are identical but the bottom halves are different, the holistic coding of faces creates the illusion of a different identity, resulting in longer reaction times and more frequent errors to indicate that the top halves are the same than in the misaligned condition. Similarly to norm-based coding, impairments in holistic processing reflected in reduced composite effects have been found in patients with prosopagnosia (Palermo, Willis, et al., 2011) and autism (Teunisse & de Gelder, 2003). Moreover, holistic face processing has been related to face recognition (Richler, Cheung, & Gauthier, 2011) suggesting that the degree to which perceivers encode faces as a meaningful whole is related to their capacity to discriminate one face from another.

Interestingly, a growing literature (Chua, Richler, & Gauthier, 2014, 2015; Richler & Gauthier, 2014; Richler, Mack, Gauthier, & Palmeri, 2009) has recently suggested a new approach of holistic processing of faces as a "failure of selective attention". Specifically, by manipulating whether information diagnostic of an individual identity is located in the top half, the bottom half or both halves of the face, Chua et al. (2014) showed that holistic coding arises through learned attentional strategies. To the extent that perceivers have learned that information diagnostic of a person's identity may be found in a specific part of his or her face, this part becomes difficult to ignore, although it might be no longer relevant to discriminate between new faces. Indeed, attention allocation at early stages of face processing is crucial to determine identity recognition. In fact, evidence of holistic coding is found at 170 ms after the presentation of the face, together with the first face-specific neural responses, namely, the N170 event-related potential (Jacques & Rossion, 2009).

Although norm-based coding and holistic processing have traditionally been considered separately, there have been recent attempts to understand the contribution of each mechanism to identity recognition (McKone, 2009; Susilo et al., 2010). Notably, Engfors, Jeffery, Gignac, and Palermo (2017) showed that norm-based and holistic coding are correlated but contribute differentially to individual differences in faces recognition. Specifically, the authors suggest that holistic coding allows creating a robust representation of faces (McKone, 2008), and this perceptual representation is subsequently used to evaluate new faces according to norm-based processes (Goffaux & Rossion, 2006). This model is consistent with electrophysiological evidence situating holistic coding prior to norm-based analyses in the time course of face processing, as stated above (Burkhardt et al., 2010; Jacques & Rossion, 2009). Altogether, these findings suggest that accurate identification of people's identity is determined by attention to the facial cues that are diagnostic of their identities in a configural fashion, and the evaluation of these characteristics with respect to an evolving norm sustained in experience. Therefore, depending on whether perceivers pay attention to these facial cues in a configural fashion, a face will be more or less individually processed.

Finally, it is worth mentioning that a significant body of research suggests that among all facial cues, eyes are particularly relevant. Studies using an eye tracker camera showed that when participants are presented with faces, approximately half of the fixation time is spent on the eyes (Henderson, Williams, & Falk, 2005; Janik, Wellens, Goldberg, & Dell'Osso, 1978) and this preference for eyes against other facial features arises early after birth (Farroni, Csibra, Simion, & Johnson, 2002). This asymmetrical attention allocation to the eyes is related to face recognition as removing eye-related facial cues significantly impairs faces recognition (McKelvie, 1976). Moreover, individuals with autism attend to nose and mouth to a greater extent than non-autistic perceivers (Klin, Jones, Schultz, Volkmar, & Cohen, 2002), which has been related to atypical patterns face recognition as mentioned in a previous section. In neurotypical population, individual differences in attention to the eyes account for later memory of these faces (Heisz, Pottruff, & Shore, 2013). The relationship between attention to the eyes and the extraction of individual-related information has been directly tested in a recent study by Kawakami et al. (2014), observing that instructions to individuate targets led to an increased attention to the eyes during the coding phase and higher performance on a subsequent recognition test. Altogether, these findings suggest that although faces are processed in a configural manner, asymmetrical attentional weights to particular features such as the eyes may contribute to identity recognition.

Therefore, in face processing, information related to individual identities is extracted by means of several mechanisms. Featural (norm-based coding) and configural (holistic coding) information are integrated to account for the uniqueness of each face. Among the featural information extracted from facial cues, the eyes seem to provide valuable information diagnostic of a person's identity. Importantly, the individual features extracted from a face impact downstream the attribution of specific traits to this individual (Hassin & Trope, 2000; Hugenberg & Wilson, 2013). Judgments based on facial cues include health condition (Tskhay, Wilson, & Rule, 2016), dominance (Oosterhof & Todorov, 2008), professional success (Rule & Ambady, 2008), leadership (Re & Rule, 2017) or trustworthiness (Todorov, Said, Engell, & Oosterhof, 2008), which will be explored in the last section of this chapter. Altogether, these findings suggest that the extraction of information related to a person's individual facial features and therefore identity, is a crucial first determinant of the impression formed of this person.

1.2. Social categories

Aside from the information distinctive of a person's individual identity, faces also inform of a person's social category membership. Information related to a person's sex, race and age is readily and efficiently extracted from faces (Bruce & Young, 1986; Hugenberg & Wilson, 2013; Stolier & Freeman, 2016). Unlike individual identity coding requiring the integration of featural and configural information, as stated in the previous section, the extraction of category-related information seems more straightforward.

Each social category is associated with specific facial features cueing a target's social group membership. For instance, features related to hair length, eyes area, jaw and chin shapes allow perceivers to distinguish between men and women (Brown & Perrett, 1993; Macrae & Martin, 2007). Skin color, hair texture or lips shape are informative of a person racial group (Blair, Judd, & Fallman, 2004; Blair, Judd, Sadler, & Jenkins, 2002). Finally, wrinkles, facial sagging, age-related spots and eyes and lips size are facial cues informative of a person's age (Porcheron, Mauger, & Russell, 2013). Moreover, there is evidence that attention is drawn to those category-diagnostic facial cues at very early stage of face processing, and that very few category-related information is necessary to trigger categorization processes (Martin & Macrae, 2007; Stolier & Freeman, 2016). For instance, there is consistent evidence suggesting category-cueing facial features readily prompt the categorization of targets into the aforementioned social groups (Blair, Judd, & Chapleau, 2004; Blair, Judd, & Fallman, 2004; Blair et al., 2002).

Focusing on sex-based categories, Macrae and Martin (2007) used a priming paradigm to investigate whether hairstyle alone was sufficient to cue the activation of gender-related knowledge. Concretely, in one experimental condition, participants were asked to indicate whether a series of names (e.g., John, Julie) were characteristically male or female. Before the presentation of the name, participants were primed with either male or female faces. Importantly, to test whether hairstyle alone may trigger categories activation, the authors also manipulated that the primes either comprise all facial features (face condition), or only a "floating" hairstyle without any additional facial cues (hairstyle condition). The pattern of data indicated that face and hairstyle primes similarly activated gender-related knowledge, with reduced reaction time and fewer errors on congruent compared to incongruent trials, broadly suggesting that the mere detection of a category-cueing facial feature trigger categories activation. Additionally, the extraction of categorical information seems to be resistant to manipulations that normally impair identity identification. For instance, Cloutier, Mason, and Macrae (2005) had participants reporting either targets' identity or sex from the presentation of their faces. They observed that suboptimal viewing conditions affecting either holistic coding (e.g., inverted faces) or featural coding (e.g., blurred face) impacted the extraction of identity-related information to a greater extent than category-related information, both in terms of reaction time and proportion of errors.

Therefore, the extraction of category-related information from faces seems to require less effort and is more robust to alterations than the extraction of individual-related information. The predominance of category-cueing feature in face processing is also supported by electrophysiological evidence. In fact, attention to facial features diagnostic of social categories based on sex, race and age occurs at very early stages of face processing, regardless of the relevance of such categories for performing the assigned task. Specifically, sensitivity to categorical facial features has been found as early as 122 ms for race, and between 145 ms to 185 ms for gender and age (Ito & Urland, 2003; Mouchetant-Rostaing & Giard, 2003; Mouchetant-Rostaing, Giard, Bentin, Aguera, & Pernier, 2000), all being earlier components than the 250-ms face discrimination component discussed in the previous section (Burkhardt et al., 2010).

The fast and efficient identification of social categories from the presentation of faces has important consequences on later stages of face processing, influencing the allocation of attention to these faces. For instance, in an event-related potential (ERP) study, Ito and Urland (2003) reported that white perceivers showed an attentional bias for same-race (i.e. ingroup) compared to black (i.e., outgroup) faces, reflected in a deeper processing of white faces at late stages of face encoding. Moreover, and just as observed with individual facial features, the early processing of category-relevant information impacts the subjective evaluation of these faces. Adolphs, Tranel, and Damasio (1998) showed that across repeated presentations of faces, activity in the amygdala is reduced for same-race but not other-race individuals. Interestingly, such activation in the amygdala has more recently been associated with superficial judgments (Freeman, Schiller, Rule, & Ambady, 2010), implicit negative racial bias (Phelps et al., 2000), and implicit and explicit gender stereotypes (Quadflieg et al., 2009). In the same vein, in an ERP study, Tortosa, Lupiáñez, and Ruz (2013) observed a larger amplitude of the N170 potential for black compared to white faces, which has been associated with implicit racial bias (Ibañez et al., 2010). Importantly, social judgments based on structural qualities of faces are modulated by the typicality of this face, that is, the degree to which a person's face is representative of his or her group. The more typical the face is, the more the judgments of this face are impacted by category-related knowledge. For instance, facial cues associated with black individuals include dark skin, coarse hair, large nose and full lips. Within the same racial group, individuals with more typical traits (e.g., darker skin tone, coarser hair, and so on) trigger more category-based stereotypical judgments (Blair, Judd, & Chapleau, 2004; Blair et al., 2002). Similar effects are found regarding age-related facial cues. Specifically, features related to a baby-face such as round cheek, large eyes and high eyebrows trigger age-related traits, in such a way that more typical baby-faces are readily associated with traits such as submission or dependence (Montepare & Zebrowitz, 1998).

Finally, it is worth mentioning that facial features also allow perceivers to categorize people in subtler social categories such as professional category (Goldstein, Chance, & Gilbert, 1984) sexual orientation or religious group (Tskhay & Rule, 2013). However, the extraction of category-related information based on those groups is different from categorization based on sex, race and age for different reasons. First, focusing on sexual orientation and religious groups, Tskhay and Rule (2013) observed categorization accuracy from faces is much lower for subtle (i.e., 64.5%) than race-based (i.e., 99.4%) categories. Additionally, this low accuracy has been argued to be overestimated in the case of sexual orientation because of discrepancies between the in-lab and real-life ratio gay:straight people (Plöderl, 2014). Moreover, face-based categorization in subtler group seems to be based on a combination of sex-related and age-related facial features (e.g., gay men possess more feminine and more babyish facial features) (Rule, 2017), instead of being associated with their own markers. Finally, to our knowledge, accurate detection of the facial cues diagnostic of a person's sexual orientation or religious group has not been consistently related to any neuro- or electrophysiological marker. Therefore, although facial features may accurately cue a person's membership to subtle social categories such as religious groups or sexual orientation, the mechanisms underlying this categorization process seems substantially different from the ones observed with sex, race and age.

Altogether, these findings indicate that information related to sex-, race- and gender-related social categories are encoded at very early stages of face processing, and affects attention to and evaluation of these faces outside of the perceiver's awareness.

1.3. Emotional expressions

As well as stable information related to a target's social category or identity, faces also provide relevant information about a person's transient state. Evolutionary theories of emotion suggest that humans' outstanding capacity to portray and interpret others' affective states has evolved to favor efficient social communication and, to a broader extent, survival (Ekman, 2003). Within this theoretical framework, six basic emotions including happiness, sadness, anger, disgust, fear, and surprise, are thought to be associated with specific facial features and movements identifiable across cultures (Ekman & Cordaro, 2011), involuntarily, and outside from awareness (Tracy & Robins, 2008).

A vast body of research suggests that on the basis of static and dynamic facial cues, perceivers can readily and efficiently determine a target's affective state and behave accordingly (Parkinson, 2005). For instance, negative emotional expressions prompt avoidance responses while positive emotions elicit approach-related behaviors (Adams Jr, Ambady, Macrae, & Kleck, 2006; Marsh, Ambady, & Kleck, 2005). In that sense, emotion processing shares notable similarities with gender and sex encoding from faces. For instance, larger amplitudes of ERP are observed for emotional faces compared to neutral faces in early stages of face processing, ranging from 120 to 180 ms after the presentation of the face (Eimer & Holmes, 2007). Furthermore, analyzing the interaction between race (black vs. white) and emotion (happy vs. angry) in early faces coding, Tortosa, Lupiáñez, et al. (2013) suggested that the processing of emotional and category-related cues are interdependent. Specifically, the authors reported an interaction between emotion and race in the N170 potential. When targets were black, both happy and angry emotional expressions were associated with larger amplitudes of the N170 potential compared to neutral expressions. In contrast, when targets were white, only angry emotional expression was associated with a larger amplitude of the N170 compared to happy and neutral expressions. Because the N170 potential has been linked to increased attention and deeper processing (Olofsson, Nordin, Sequeira, & Polich, 2008; Stahl, Wiese, & Schweinberger, 2008), it seems that white angry faces were more cautiously processed while this effect was extended to both black emotional faces, independently of their valence. The perceived threatening value of negative (for Whites) and emotional (for Blacks) stimuli may account for these data (Ohman & Mineka, 2001), likely related to the lack of familiarity of white participants with outgroup faces (Scott, 2006). Overall, these data suggest that while threatening emotional stimuli received more attention for same-race faces, this increased in attention may be generalized to any emotional expression for other-race faces.

On a different note, emotional expressions allow perceivers to infer others' preferences and intentions by means of mentalizing processes (North, Todorov, & Osherson, 2010). Thus, emotional expressions impact, for instance, perceivers' cooperative behaviors. Notably, targets' positive expressions cue cooperation decisions while negative expressions are associated with non-cooperative behaviors (Alguacil, Madrid, Espín, & Ruz, 2017; Mussel, Göritz, & Hewig, 2013; Tortosa, Strizhko, Capizzi, & Ruz, 2013). Moreover, perceivers fail at ignoring emotional cues even when they are irrelevant to perform the assigned task, and even when this results in impaired performance (Hodsoll, Viding, & Lavie, 2011; Liang, Chen, Yan, Qu, & Fu, 2018; Ruz, Madrid, & Tudela, 2013). Interestingly, even unconscious affective information may bias social judgments. For instance, Sweeny, Grabowecky, Suzuki, and Paller (2009) used an affective priming paradigm in which faces portraying surprise, an emotion that may be interpreted of positive or negative valence, were subliminally primed (i.e., 30 ms) with faces expressing either fear, happiness or a neutral emotional expression. Next, participants were asked to evaluate the valence of the target surprised faces and reported more positive evaluation for faces subliminally primed with happiness and more negative evaluations for faces subliminally primed with fear, indicating that the affective information from the primes impacted the evaluation of subsequently presented faces, even when participants when unaware of this affective information. Moreover, 24h after this coding phase, participants performed a recognition test in which they were presented with new faces among the faces evaluated the previous day. Even after 24h, the memory of the faces was still impacted by the affective information from the primes, as faces primed with happiness were better recalled than faces primed with fear or a neutral expression. Interestingly, these results echo previous research showing better memory of the identities of faces portraying happiness compared to faces portraying a negative emotional expression such as sadness (Ridout, Astell, Reid, Glen, & O'Carroll, 2003). Therefore, this literature indicates that emotional information from faces is processed effortlessly and efficiently at early stages of face processing, and has a crucial impact on predictions and inferences about others' states of mind, and the subsequent associated behavior.

Given the importance of face processing in traits attribution and the consequent attitude and behaviors, it is not surprising that faces have been widely used in social cognition research to investigate how perceivers understand others and regulate their behavior according to these inferences. Notably, the processes underlying social impressions formation have received much attention in past and present theorizing. A large body of research focuses on understanding whether individuals are judged according to their category membership, namely, on the basis of the categorical cues extracted from their faces, or understood as unique individuals, relying instead on their idiosyncratic attributes. Some of the factors impacting social perception are considered in the next sections.

2. Social perception

As it has been argued in the previous section, faces are crucial in impression formation processes, as they provide perceivers with relevant information about targets identity, group membership and emotional states, among others. Therefore, our impressions of others and further judgments and trait attributions are closely related to the way we pay attention to their faces. Decades of research in social cognition has focused on whether perceiver considers the uniqueness of each individual or rather rely on categorical features to inform their judgments. These two approaches are conceptualized as individuation for the former and social categorization for the latter.

2.1. Social categorization and individuation processes

The most prominent models of social perception identify social categorization and individuation as the main strategies of impression formation (Brewer, 1988; Fiske & Neuberg, 1990). On the one hand, social categorization consists of attending targets' attributes diagnostic of their group membership to make inferences about them. In these processes, the facial cues informative of a person's category are readily attended. Once a person has been categorized, the information related to his or her group, in form of previous experience or stereotypes, impact the perception of this particular target. Imagine for instance a character called Jim who enters the waiting room of his son's pediatrician and sits next to one of the strangers in the room. From her facial features, Jim easily identifies this stranger as a woman, which may activate some stereotype about women being high on communal behaviors (Cuddy et al., 2009). Believing that she is surely friendly, he may, after a while, start a conversation about her children, the reasons that brought her to the pediatrician, and other casual topics making the wait more bearable. Crucially, Jim would not have started this conversation if the stranger next to him was a fit large and tall man. In this case, he would have rather focused his attention on the Real Madrid t-shirt worn by the little girl accompanying the stranger, and would have started a football conversation. In this situation, categorical visual features may have cued category-related judgments which, in turn, has guided his behavior.

On the other hand, individuation consists of relying on the idiosyncratic attributes and qualities of a target to inform one's judgments. In contrast to social categorization, an individual assessment of a particular target may not be achieved by merely attending category-relevant facial cues. Instead, individuated judgments require the piecemeal integration of diverse individual characteristics to consider the uniqueness of each individual. In consequence, as stated above, increased attention is needed to individuate (compared to categorize) individuals. Following the previous example, imagine now that the people in the waiting room are other candidates for the job Jim is applying for, as the clinic is searching for a new pediatrician. In this context, the fact that the other people are competitors makes them relevant to him. He might not rely as much on gender categories to evaluate the woman seated next to him and choose a topic of conversation. Instead, he may be motivated to understand her in a more individuated manner, and try to know more about her personal attributes and qualities, and consequently, his chances to get the job. Importantly, in this context, Jim may be less likely to pick a children-related topic to engage a conversation.

Although there is less consensus regarding whether categorization and individuation processes may occur in parallel (Brewer, 1988), or are the two extremes of a continuum (Fiske, Lin, & Neuberg, 1999; Fiske & Neuberg, 1990), theorists agree on the fact that categorization occurs immediately upon encounters, and a set of motivational and cognitive factors determine whether a person is understood in terms of his or her unique attributes and qualities, or as a member of a determined social group. For instance, Fiske and Neuberg (1990) proposed that in first encounters, targets are readily categorized on the basis of category-relevant facial cues. In most occasions, impressions formation processes stop at this initial stage, and the judgments obtained from such processes are rather superficial. However, as shown in Figure I, when the target is relevant enough (e.g., this woman is a competitor), perceivers attend to his or her individual attributes, either confirming the former categorization or initiating a recategorization when the individual-based information does not match the initially activated category. When the target cannot be classified into any available category, perceivers engage in a piecemeal integration of the target's individual attributes until reaching a fully individual assessment evaluation of the target. For these processes to be successful, perceivers need, in addition to being motivated to attend to individual attributes, to be cognitively able to assess the target's individual features. Put simply, whether a person will be categorized or further evaluated according to idiosyncratic attributes broadly depends on the perceiver's *capacity* and *motivation* to individuate.

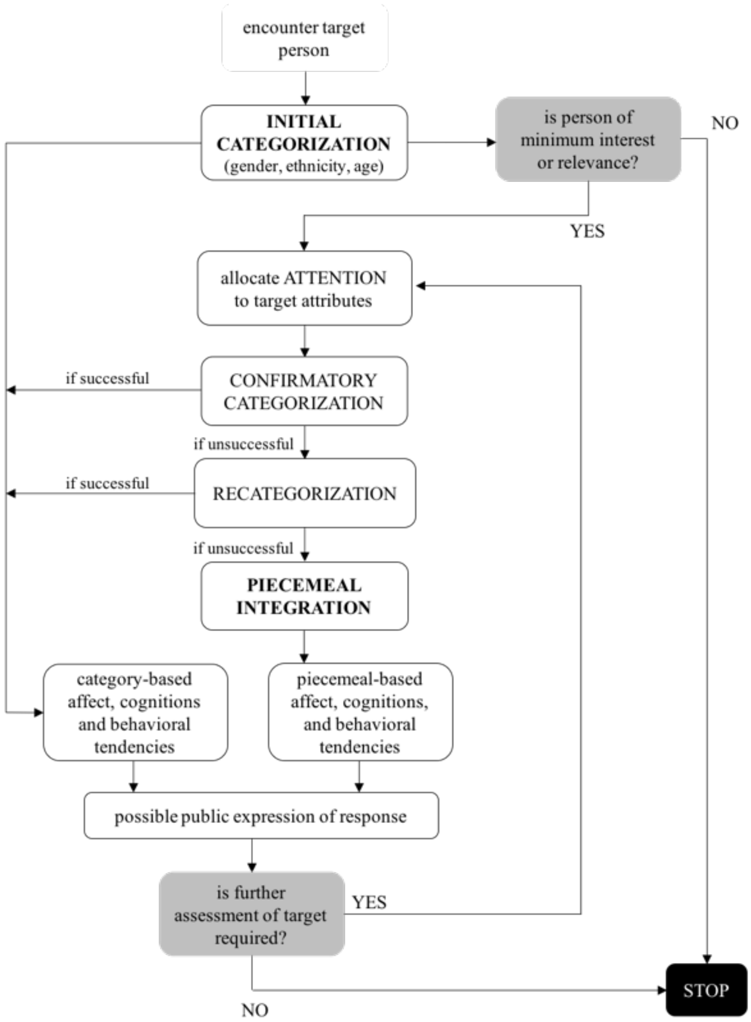

Figure I. Schematic representation of the continuum model of impressions formation adapted from Fiske and Neuberg (1990).

According to Fiske and Neuberg's (1990) model, shown in Figure I, social categorization allows perceivers to make predictions about others effortlessly, on the basis of information readily and easily identifiable from their faces. In this sense, it fulfills a cognitive economy function, saving perceivers "the trouble of thinking" (Gilbert & Hixon, 1991). Consistently, perceivers may aim at individuating, but be forced to rely on categorical thinking when their cognitive resources are being depleted by contextual demands. Going back to the clinic and the pediatrician job offer, Jim may want to evaluate his competitors individually but be unable to do it as he has to fill a job application form, review his CV, prepare for potential questions during the interview and maintain updated his partner constantly asking "How is it going?". Conversely, perceivers may be able to individuate but lack motivation to involve in this costly process, and rather rely on a less demanding categorization strategy. This would be the case of Jim having nothing else to do but wait for the pediatrician to call him and his son, but not being motivated to know the individual attributes that make the woman seated next to him, unique. Accordingly, the impact of cognitive and motivational factors on social judgments are reviewed in the following sections.

2.2 Cognitive factors impacting social perception

Human's mental capacity is limited, and this cognitive constraint determines how social stimuli are processed and integrated to make sense of people (Allport, 1954; McGarty, Yzerbyt, & Spears, 2002; Tajfel, 1969).

A large body of research suggests a greater impact of categorical (versus individual) information in social judgments when perceivers' attentional resources are depleted (Fiske, 1980; Gilbert & Hixon, 1991; Leyens, Yzerbyt, & Schadron, 1994; Macrae, Hewstone, & Griffiths, 1993; Macrae, Milne, & Bodenhausen, 1994). For instance, Pendry and Macrae (1994) investigated the impact of a concurrent resources-depleting task (i.e., memorizing a digits string) while trying to form an impression of an unfamiliar target by reading her description. Moreover, motivation was manipulated by making half of the participants outcome-dependent on the accuracy of their predictions to motivate them to individuate the target, while the other group was not outcome-dependent and therefore not necessarily motivated to individuate. The results indicated that non-motivated participants made stereotypic judgments of the target, on the basis of category-related knowledge. However, participants in the motivated group made individuated judgments, but only when their attentional resources were fully dedicated to the evaluation of the target. Altogether, these results suggest that even in contexts in which motivation to individuate is high, reduced capacity may promote categorical thinking. These data were further replicated in more ecological task settings, in which participants were not instructed to perform any particular task but were rather distracted by "incidentally" hearing a potentially relevant, but unrelated to the task, conversation (Pendry, 1998).

Importantly, category-related knowledge is thought to be activated upon the identification of target's social categories (but see Gilbert & Hixon, 1991). The impact of reduced attentional resources lies in whether, once activated, this categorical information will be applied or not to evaluate targets (Gilbert & Hixon, 1991; Govorun & Payne, 2006; Sherman, Macrae, & Bodenhausen, 2000).

The limitation of perceivers' attentional resources is a concrete constraint for efficient individuation processes. Research in social cognition has convincingly documented that in spite of its potential negative consequences, social categorization is a remarkable resources-saving tool allowing us to make sense of the social world effortlessly. However, limited attentional resources are not sufficient to explain why people rely on categorical thinking in such a large range of situations. As observed in Pendry & Macrae's (1994) research described above, the use of categorical knowledge is the default strategy to make sense of others, even in contexts of low cognitive demands. In fact, when participants were not outcome-dependent on making accurate decisions, they did not engage in individuating processes even if the cognitive demands of the task were low. In this regard, some theorists have argued that perceivers essentially rely on the cognitive economy principle, dedicating the least cognitive resources to impression formation (as well as other cognitive) processes (Fiske & Taylor, 1991; Sherman et al., 2000). Although not flattering, this theory has received a considerable empirical support, showing that unless being highly motivated, perceivers rarely perform at the best of their capacity.

2.3. Motivational factors impacting social perception

After the initial categorization of others into their corresponding social group, perceivers' specific motives may promote that the category-related knowledge about this specific group is ignored in favor of individuated information. Among the factors promoting individuation, Fiske and Neuberg (1990) highlighted three main motivating agents: outcome-dependency on the target, third-party and personal values. Within these three general agents, different specific factors contribute to enhance the motivation to individuate.

Situations of interdependence between perceivers and target impact the impressions formation in that being outcome-dependent on someone promotes individuation (Darley, Fleming, Hilton, & Swann, 1988; Fiske & Neuberg, 1990). For instance, Neuberg and Fiske (1987) made participants outcome-dependent on an alleged schizophrenic patient, and evaluated their impressions of this target. Specifically, in the outcome-dependency group, participants were told that they would have to realize a creative task with the patient, and that both of them would be rewarded with $20 if their work were selected as the most creative one. Alternatively, participants who were not outcome-dependent on the target were told that their contribution to the creative work would be evaluated separately, including the possibility that only one of the two receives the $20 prize, despite working together on the task. The authors reported that participants in the outcome-dependency condition paid more attention to the individual attributes of the target and provided more individuated judgments about him.

A different instance in which people may be outcome-dependent on others is when the relationship between a perceiver and a target is qualified by power differences. Critically, power differences conceptualized in social terms (Fiske, 2010) confer to one individual (i.e., powerholder) control over other's resources (i.e. powerless). Because of outcome-dependency, powerless people are likely motivated to understand and please powerholders (they need to understand the powerful, as they depend on them), while the opposite is not necessarily true (they do not need to individuate the powerless, as they do not depend on them). A common paradigm to investigate the effect of power consists of assigning participants to a group task, and crucially manipulating whether within the group, they are powerful (i.e., they are able to decide how alleged rewards will be distributed within the group) or powerless (i.e., a different member of the group will decide the distribution of the rewards within the group) members. Prior to any contact with the other members of the group, participants are presented with a series of traits supposedly describing their counterparts, either category-consistent or category-inconsistent. The time spent reading the different traits informs of the allocation of attentional resources to understand the targets. In this setting, powerholders have been consistently found to pay more attention to category-consistent information (Fiske, 1993; Fiske & Dépret, 1996; Goodwin, Gubin, Fiske, & Yzerbyt, 2000; Rodríguez-Bailón, Moya, & Yzerbyt, 2000). These data have been replicated in ecological contexts in which power differences existed between participants prior to the participation in the experiment (i.e., managers vs. subordinates) (Guinote & Phillips, 2010). However, and importantly, power is also related to specific personal values which may, instead, promote individuation.

In fact, perceivers may be their own motivating agents according to the internalized personal values that are relevant for their self-esteem (Fiske & Neuberg, 1990). Regarding powerholders, one important intrinsic motivation regards the need to preserve their powerful position (Rodríguez-Bailón et al., 2000), which may impact their strategies of social perception. For instance, Rodriguez-Bailon, Moya, and Yzerbyt (2006) observed that powerholders may be motivated to make individuated judgments about others when they illegitimately acquired their powerful position. In this experiment, participants were assigned to a powerful position after receiving either positive (i.e., legitimate condition) or negative (i.e., illegitimate condition) feedback about their supervisory skills. Further, participants were asked to choose one among two candidates to go to a university conference. Importantly, the profiles of the candidates were manipulated such that one of them appeared to be high on competence and sociability while the other one was rather low on the two dimensions. Although one of the candidates seemed clearly better skilled than the other, participants in the illegitimate power condition chose the less skilled partner more frequently than participants in the legitimate power condition. Moreover, and importantly, this difference was found despite illegitimate powerful participants asked more information about the positively described candidate than participants in the legitimate group. Therefore, their decisions were not due to a lack of information about the candidates, but were rather made in spite of having individuated them.

Altogether, these results suggest that when powerholders are motivated to preserve their position because of its illegitimacy, they may engage in more demanding individuation processes to ensure that their position is not threatened. In a similar vein, Ruscher and Fiske (1990) showed that competition settings promote individuation. To the extent that making accurate predictions about a competitor serves one's own benefits, he or she may be willing to engage in individuating processing to understand the strengths and the weaknesses of his or her competitor. Finally, powerholders goal-driven attitudes also impact basic cognition as powerful individuals may show higher attentional flexibility and inhibition (Guinote, 2007). Altogether, these data indicate that according to personal goals, power may either promote individuation or categorization strategies.

Perceivers' personal values may also be impacted by transient emotional states. For instance, anger is elicited in negative circumstances in which perceivers feel a lack of control over the situation. Therefore angry individuals are motivated to restore control which prompts the use of heuristic-based strategies allowing fast decision-making (Scott, 1980). In contrast, sadness is related to the acceptation of a loss, such that sad individuals are motivated to avoid the thoughts related to the situation that elicited sadness, which makes them prone to use details-oriented processing of information (Wenzlaff, Wegner, & Roper, 1988). Accordingly, Bodenhausen, Sheppard, and Kramer (1994) observed that participants induced with anger were more prone to evaluate others according to their group membership, while participants induced with sadness were more inclined to individuate targets.

Finally, personal values also broadly includes perceivers' intrinsic motivation to be accurate in their predictions (Neuberg, 1989), low prejudice level (Blair, 2002), or group motives related to self-categorization processes, which will more extensively described in a later section of this chapter.

The third motivating agent impacting impressions formation according to Fiske and Neuberg (1990) regards the presence of a third-party aside from perceivers and targets. Specifically, perceivers are more prone to engage in individuation processes when they are accountable to a third-party. For instance, Tetlock (1983) had participants reporting their thoughts on controversial social issues in four experimental conditions: with no expectation to justify their attitudes (a), with expectation to justify their attitudes to a conservative third-party (b), to a liberal third-party (c) or to a third-party with unknown views (d). The pattern of data observed indicated that participants engaged in more complex information processing when they were expected to justify their attitudes to others, especially when they knew nothing about the third-party's beliefs about the topic. Therefore, accountability may promote people engage the cognitive resources necessary to make accurate judgments. This effect was replicated when participants were directly asked to judge unfamiliar people instead of broad social issues (Lerner, Goldberg, & Tetlock, 1998; Pendry & Macrae, 1996)

Altogether, these findings consistently indicate that several motivational factors may promote the differential use of categorization and individuation strategies. Although the aforementioned variables (i.e., outcome-dependency, power, emotional state, prejudice level, motivation to be accurate, and accountability) are generally considered in relation with the self, perceivers may, in some situations, activate the goals and values related to their group (Ellemers, 2012; Ellemers & Haslam, 2012). This focus on their social identity elicits specific attentional, behavioral and attitudinal strategies that also impact impression formation.

2.4. Self-categorization and social perception

Among the motivational variables associated with perceivers' personal values, people may be differentially motivated to individuate members from specific social categories, depending on self-categorization processes. In fact, upon activating social categories, one can easily classify targets as ingroup or outgroup members (Ellemers, 2012; Haslam, Oakes, & Turner, 1996). This process of self-categorization is associated with a series of attentional and behavioral biases (Spears, Oakes, Ellemers, & Haslam, 1997).

Specifically, while ingroup faces are processed along with the features diagnostic of their individual identities, outgroup faces are rather processed on the basis of category-relevant features. This difference in encoding stages influences later memory of the faces, as ingroup members are better recognized than outgroup members, an effect first conceptualized as the other-race effect (ORE, Hugenberg, Young, Bernstein, & Sacco, 2010; Levin, 1996; Meissner & Brigham, 2001), and also known as own-race (Slone, Brigham, & Meissner, 2000) or cross-race (Sporer, 2001) bias. Initial theorizing about the ORE considered it as a consequence of a reduced expertise in processing outgroup compared to ingroup faces, likely because of lesser contacts with people from a different race (Lewis, 2004; McKone, Aitkin, & Edwards, 2005). This idea is framed within the norm-based coding theory described in the first section of the current chapter (Valentine, 1991). In fact, the norm-based model posits that experience affects perception such that encountering new faces allows perceivers to calibrate the norm and use a variety of dimensions to discriminate among them. Considering that ingroup faces are more often encountered than outgroup faces, ingroup faces are perceived according to a more diverse range of dimensions than outgroup faces. For instance, white compared to black individuals are characterized by having thinner lips. For a white perceiver exposed to numerous white faces across his or her life, coding lips thinness alone is not sufficient to discriminate between white targets. In consequence, white individuals will consider new dimensions to process Whites' faces and will achieve a more detailed analysis of the differences between faces featuring thin lips. However, the same white perceiver encountering a black individual experiences a facilitation to identify the thickness of his or her lips, but lacks the experience to further attend the subtler differences between individuals with full lips. The direct consequence is that ethnic outgroup faces are poorly coded compare to ingroup faces.

Although this argument seems quite convincing, more recent theorizing indicated that the ORE is not fully accounted by experience-based models, but instead, a motivated bias seems to underlie the effect (Hills, Pake, Dempsey, & Lewis, 2018; Hugenberg & Wilson, 2013; Kawakami, Amodio, & Hugenberg, 2017). In fact, ORE is not restricted to social categories with whom perceivers lack experience and visual expertise, as other-sex (Wright & Sladden, 2003) and other-age (Wright & Stroud, 2002) effects are also well-documented. Similarly, faces categorized as sharing one's same sexual orientation are better remembered than faces of people of a different sexual orientation (Rule, Ambady, Adams, & Macrae, 2007). Simply stated, classifying faces as ingroups or outgroups is sufficient to elicit attentional and memory biases. For instance, Bernstein, Young, and Hugenberg (2007) presented participants with faces of unknown people on either a green or a red background. In one experimental condition, participants were told that the background color was irrelevant for the task whereas in the other experimental group, it allegedly informed of the target's university affiliation. In the former condition, no difference was observed in the recognition of targets. However, in the second experimental group, recognition was better for university ingroup than outgroup targets. Similar results were found manipulating that targets were believed to be ingroup or outgroup members with regard to religious categories (Rule, Garrett, & Ambady, 2010). Therefore, both ascribed and acquired group membership bias face processing. To the extent that perceivers may identify targets as ingroups, processing, and consequently memory of these targets are considerably impacted. Importantly, motivating people to individuate outgroup members by the means of monetary (Kawakami et al., 2014) and non-monetary (DeLozier & Rhodes, 2015) rewards may reduce the ORE.

Echoing the ORE effect, greater reliance on categorical information is not restricted to face perception, but is also found in trait attribution processes. In fact, categorical knowledge is more readily used to make inferences about outgroups. Conversely, judgments about ingroup members are more often based on the personal characteristics of this individual. The fact that ingroup members are more individuated than outgroups results in a higher perceived intragroup similarity among outgroup members compared to ingroup members. Put simply, outgroup members are often perceived "all alike", an effect known as outgroup homogeneity (Haslam et al., 1996; Judd & Park, 1988; Linville, Salovey, & Fischer, 1986; Park & Rothbart, 1982; Quattrone & Jones, 1980). Outgroup homogeneity effect is thought to be the result of using different levels of social categorization to code the behaviors of ingroup compared to outgroup members (Park & Rothbart, 1982). Specifically, outgroup behaviors are appraised at the superordinate level (i.e., "a black person shows remarkable athletic qualities") while ingroup behaviors are understood according to subordinate dimensions (i.e., "a fit person shows remarkable athletic qualities", in which race is not as relevant as for outgroup members). Because perceivers are more motivated to understand ingroups than outgroups, they engage in differential learning strategies to infer their attributes, in which outgroups' behaviors seem more homogeneous than ingroups'.

Finally, one's group membership may impact the evaluation of others in a more straightforward way. Because self-concept is highly determined by the social groups one belongs to, the presence of an outgroup may increase the salience of a person's social identity, that is, his or her self-concept as a member of a specific social group. Considering oneself in terms of social membership may promote attitude and behavior in line with the group (versus individual)'s goal and values (Ellemers, 2012). Moreover, to the extent that people are willing to enhance their group identity, they may engage in behaviors favoring ingroup over outgroup members, an effect known as ingroup favoritism (Brewer, 2001; Brewer, 2007; Turner, Brown, & Tajfel, 1979). In fact, people generally show a more positive attitude toward ingroups compared to outgroups in different domains of responses such as empathy (Xu, Zuo, Wang, and Han (2009) or cooperation (Balliet, Wu, & De Dreu, 2014; Wilson & Kayatani, 1968).

Importantly, the mere distinction we/they is sufficient to elicit differential responses to unfamiliar ingroup and outgroup members (Tajfel, 1970; Tajfel & Turner, 1979). Accordingly, ingroup favoritism has also been observed in minimal paradigms where artificially created (instead of ascribed) groups are used to distinguish between ingroup and outgroup members (Hartstone & Augoustinos, 1995; Otten & Moskowitz, 2000; Tajfel, Billig, Bundy, & Flament, 1971). For instance, Ahmed (2007) randomly assigned participants to "heads" and "tails" groups by tossing a coin. Participants were later involved in the Prisoner Dilemma (Poundstone, 1992), in which they had to make a decision to trust or not a partner at zero acquaintance. Critically, Ahmed (2007) manipulated that this partner was either an ingroup or an outgroup members according to the heads/tails dimension mentioned above, and observed that participants spontaneously trusted more artificial ingroup than outgroup members. In this sense, and just as the ORE and the outgroup homogeneity, classifying someone as an ingroup members may be sufficient to trigger a differential approach of this person.

A large range of motivational and cognitive factors impact social impressions by biasing whether perceivers use category-related or individual information about others.

The analysis of the interaction between cognitive and motivational factors has contributed to define perceivers as “motivated tacticians” (Fiske & Taylor, 1991) searching for the ideal trade-off between cognitive economy and accurate predictions. The specific circumstances favoring the most accurate predictions with limited cognitive resources are still being explored. The present dissertation aims at contributing to the understanding of the factors impacting the differential use of categorization and individuation strategy and their consequences on social judgments at zero acquaintance and across repeated interactions, focusing on cooperation decisions related to trust.

3. Trustworthiness judgments

Determining whether or not to trust a stranger is not a trivial decision. In fact, trust implies having a positive expectation about somebody’s behavior, while accepting vulnerability (Dunning & Fetchenhauer, 2010; Dunning, Fetchenhauer, & Schlösser, 2012; Lewicki, McAllister, & Bies, 1998). The risk associated with trust decisions has led theorists to argue that trust at zero acquaintance is nothing less than an “anomaly” (Schlösser, Fetchenhauer, & Dunning, 2016, p. 216), and yet people do trust strangers who in turn, reciprocate (Berg, Dickhaut, & McCabe, 1995). One may intuitively think that trustworthiness judgments are thoughtful processes implying a detailed analysis of pros and cons in a determined situation. However, and as stated in the first section of this chapter, people readily evaluate others’ trustworthiness on the basis of facial features (Willis & Todorov, 2006) just as information related to social categories.

In fact, trustworthiness judgments seem based on the overgeneralization of facial cues resembling emotional expressions. Specifically, neutral expression faces whose structural variations resemble a happy expression are considered more trustworthy (Said, Sebe, & Todorov, 2009; Todorov et al., 2008) and predict people’s decisions to trust targets (van ’t Wout & Sanfey, 2008). Conscious trustworthiness judgments are made as soon as 100 ms after the presentation of the face (Willis & Todorov, 2006). Moreover, reliable trustworthiness judgments have been found even under the perceptual consciousness threshold. For instance, Todorov, Pakrashi, and Oosterhof (2009) reported that people’s trustworthiness judgments for faces presented during 33 ms were in agreement with judgments made without any time constraint. Additionally, and similarly to the Sweeny et al.’s (2009) study reported above, the trustworthiness of prime faces presented during 20 ms influences the subsequent evaluation of a target neutral face, suggesting that the trustworthiness of a face may be evaluated even when perceivers are not aware of the presentation of the face. Therefore, a person's trustworthiness is appraised at very early stage of face processing and may impact perceivers' responses outside of their awareness.

Moreover, and consistently with impression formations theories, social group membership also impacts trustworthiness judgments. For instance, female partners are trusted to a greater extent than their male counterparts (Buchan, Croson, & Solnick, 2008; Carragher, Thomas, & Nicholls, 2018; Orbell, Dawes, & Schwartz-Shea, 1994). Because of gender stereotypes associating women with communal behaviors and men to competence (Cuddy et al., 2009; Cuddy, Fiske, & Glick, 2008; Fiske, Cuddy, Glick, & Xu, 2002), people readily associate women with social-related skills and behaviors. Moreover, and according to the facial cues of trustworthiness described above, feminine facial cues are related to more trustworthiness than masculine facial cues (Oosterhof & Todorov, 2008). Similarly to gender, age also impacts social judgments. Baby-faced features are associated with higher trustworthiness as compared to mature traits (Oosterhof & Todorov, 2008). However, moving to later stages of lifespan, older adults are generally perceived as more trustworthy than younger adults (Bailey et al., 2015). In fact, just as for women, age-relate stereotypes associate older adults with communion and warmth, including trustworthiness (Cuddy, Norton, & Fiske, 2005; Schniter & Shields, 2014). Completing the big three, ethnicity also impacts trust decisions. For instance, black individuals are less trusted because of ethnic stereotypes associated them with crime, violence, or more broadly, threat (Eberhardt, Goff, Purdie, & Davies, 2004; Quillian & Pager, 2001). Consistently, perceivers' endorsement of implicit racial bias predicts the impact of targets' ethnicity on trustworthiness judgments, in that perceivers with pro-whites attitudes find white faces more trustworthy than black faces, and vice versa (Stanley, Sokol-Hessner, Banaji, & Phelps, 2011). Regarding the relationship between social group membership and trust, it is worth mentioning that the aforementioned ingroup favoritism also apply to trustworthiness judgments in that ingroups are more trusted than outgroups according to gender, ethnicity and age (Ahmed, 2007; Bailey et al., 2015; Balliet et al., 2014; Wilson & Kayatani, 1968).

Finally, targets' transient states such as emotions influence trust decisions. In fact, people's emotional expressions are informative of their internal states and intentions (Fridlund, 1994). People may readily react to affective information by approach versus avoidance behaviors (Adams Jr et al., 2006; Marsh et al., 2005; Phaf, Mohr, Rotteveel, &

Wicherts, 2014). Specifically, positive emotions such as happiness cue approach responses while negative emotions such as anger cue negative behaviors. Consequently, and consistently with the overgeneralization theory described above (Oosterhof & Todorov, 2008), happiness promotes trust responses to a greater extent than negative expressions as anger (Tortosa, Lupiáñez, et al., 2013; Tortosa, Strizhko, et al., 2013). Interestingly, emotional expressions may also bias trust decisions in a subtler way. When individual identities and emotional expressions predict incompatible responses (e.g., this particular individual is trustworthy but portrays an angry expression) participants take longer to make a decision, even when they are explicitly told to ignore the irrelevant emotional information (Alguacil, Tudela, & Ruz, 2015).

Among the different paradigms used to investigate trust decisions, the trust game (Berg et al., 1995) received much attention and has been widely modified to explore the factors impacting trust. In its classical version, participants receive a certain amount of money (e.g., $10) and have to decide how much of this initial amount to send to an anonymous partner. In a second stage, the partner receives this amount tripled (e.g., $30) and in turn decides how much of the received money, if any, he or she sends back to the participant. Trust decisions are analyzed according to the amount sent by the participant in the first stage of the game. Moreover, the amount sent back by the partner informs of the extent to which this partner is equitable (i.e., sending back a fair amount of money) or non-equitable (i.e., keeping the money for him or herself). In a later version of this paradigm, the multi-round trust game (King-Casas et al., 2005), the same partners are presented several times such that participants have several opportunities to learn whether their partners tend to be equitable or non-equitable. In this case, trust decisions are not only impacted by first impression, but also by learning across repeated interactions.

Social decisions including trust are biased by a series of factors including social categories, emotional states or previous information about a particular target (Díaz-Gutiérrez, Alguacil, & Ruz, 2017). These factors differentially contribute to the impressions formation processes and the outcomes of the relationship. Moreover, decisions made at zero-acquaintance may evolve across repeated interactions as a result of learning. In the present research, we aim at understanding the impact of the aforementioned factors on social judgments both at zero acquaintance and across repeated interactions.

CHAPTER 2. Motivation and Overview of the Research

In social perception, human faces are rich stimuli comprising unique attributes specific to a particular individual, but also, among other features, information about his or her emotional state or social group membership (Bruce & Young, 1986; Hugenberg & Wilson, 2013). All this information needs to be processed and integrated efficiently for perceivers to behave accordingly. As described in Chapter 1, in the impression formation process, perceivers may pay attention to categorical attributes, attending the information diagnostic of the target's group membership. Alternatively, they may further this level of analysis to individuation processes, allowing them to understand targets as unique individuals, independently of the social category they belong to (Brewer, 1988; Hugenberg, Young, Bernstein, & Sacco, 2010). Whether a target is categorized, individuated or rather falls somewhere on the categorization-individuation continuum broadly depends on the perceiver's cognitive capacity and inclination to individuate (Fiske & Neuberg, 1990). These two main determinants of impression formation processes result themselves from the complex interaction of several factors including the characteristics of the stimuli (Mullen, Brown, & Smith, 1992), the contextual demands (Sherman, Macrae, & Bodenhausen, 2000), the interdependence between perceivers and targets (Neuberg & Fiske, 1987), or the perceiver's power (Guinote & Phillips, 2010), or affective state (Bodenhausen, Sheppard, & Kramer, 1994), to mention only a few examples.

Although each social encounter is unique, there is only one "first impression", which makes interactions at zero acquaintance special. In these situations, in which perceivers lack previous experience with a particular target, social biases may likely drive their judgments (Bodenhausen & Macrae, 1998; Macrae & Bodenhausen, 2000), which may be either confirmed or corrected across repeated interactions. The differential impact of categorical, individual or emotional cues in the different phases of the process of social learning is an extensive issue that is still a hot research topic in social perception (Kawakami, Amodio, & Hugenberg, 2017).

The aim of the present research was to evaluate the impact of perceivers' motivational, emotional and cognitive state on the differential use of categorization versus individuation strategies, and their consequences on social judgments regarding the principal dimensions of impression formation (i.e., gender, ethnicity, age, and emotional expression).

To achieve this goal, we developed an adaptation of the trust game paradigm (Berg, Dickhaut, & McCabe, 1995), allowing us to evaluate the impact of the aforementioned factors on impression formation at zero acquaintance, during learning across repeated interactions, and after learning. Specifically, participants had to predict the cooperative behaviors of unknown game partners to earn economic rewards across three phases, as shown in Figure II.

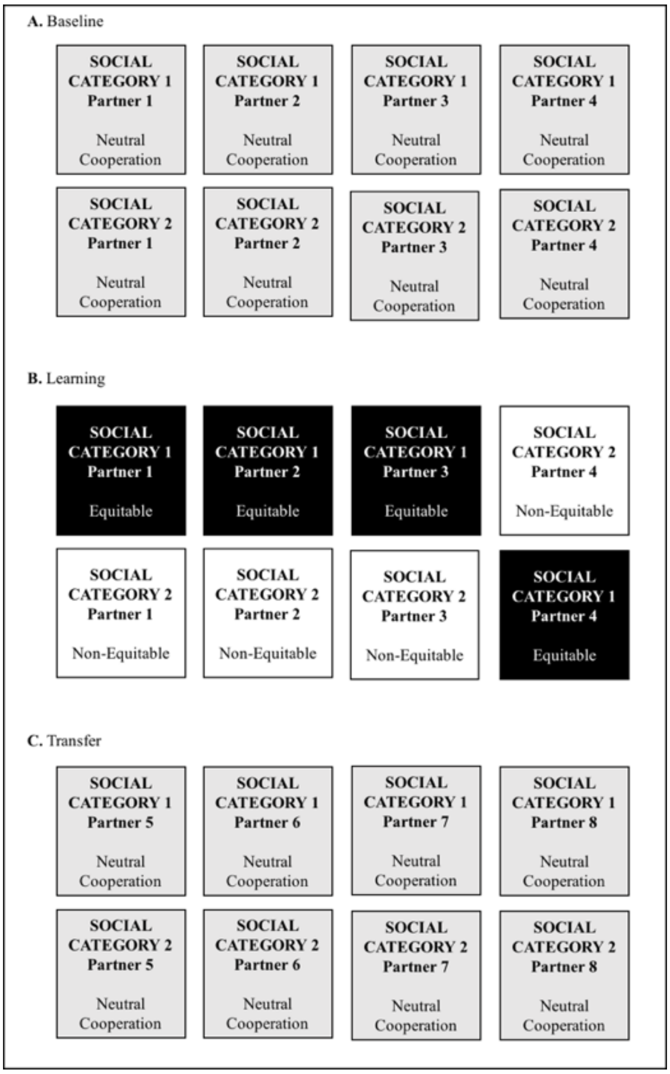

Figure II. Example of the general procedure and manipulations established in the baseline (A), learning (B) and transfer (C) phases. In the current example, participants play with 8 partners (4 from each social category). In the baseline and the transfer phases, all individuals show a neutral cooperation pattern (i.e., they cooperate in half of the trials and do not cooperate in the other half). In the learning phase, 25% of individuals (i.e., one partner in each social group) are inconsistent, and therefore display a pattern of cooperation opposite to the group behavior. In this and all the figures of the dissertation, partners are represented in black when they are equitable (i.e., they cooperate on 75% of the trials), in white when they are non-equitable (i.e., they cooperate on 25% of the trials), and in gray when their cooperation tendency is neutral (i.e., they cooperate on 50% of the trials).

First, in a baseline phase, participants were presented with partners from two different social categories, all of them cooperating in half of the trials and not cooperating in the other half. This phase allows us to verify whether participants were spontaneously biased to cooperate more with one of the social groups manipulated, either because of specific stereotypes affecting beliefs about these groups or ingroup favoritism promoting cooperation with partners belonging to the same social group as the participant.

Next, in a learning phase, we manipulated the group behavior by associating the two social categories with opposite cooperative behaviors, one being equitable and the other one being non-equitable. Moreover, partners' behaviors were also manipulated at the individual level, by introducing a small proportion of inconsistent individuals within each social category, displaying a pattern of cooperation opposite to the group behavior. Taking into account that attention determines what is learned (Jiménez & Méndez, 1999), with this procedure, we sought to understand whether participants pay attention to individual or categorical information, and consequently learn about the individual or the group he or she belongs to. In particular, participants' cooperation with inconsistent partners is informative of whether they use a categorization or an individuation strategy, as shown in Figure III. In fact, if participants use a categorization strategy, they should apply their knowledge about the group to all particular individuals within this group, independently of their consistency. This should be reflected in a similar pattern of cooperation with consistent and inconsistent individuals from the same social group, as shown in Figure IIIa. Conversely, if participants use an individuation strategy, they should notice that inconsistent individuals do not behave as the group, and thus, participants should revert their strategy of cooperation with inconsistent individuals as compared to consistent individuals, as shown in Figure IIIb. Whether participants categorize or individuate their partners may also be analyzed by computing a learning index, subtracting participants' cooperation with partners individually non-equitable from their cooperation rate with individually equitable. With the learning index as a dependent variable, the crucial analysis consists of comparing learning about consistent vs. inconsistent partners. Opposite patterns of learning for consistent and inconsistent individuals (and therefore negative values of learning for inconsistent individuals) would reflect a categorization strategy, as shown in Figure IIIa, while similar patterns of learning (and therefore positive values of learning for inconsistent individuals) would reflect an individuation strategy, as shown in Figure IIIb.

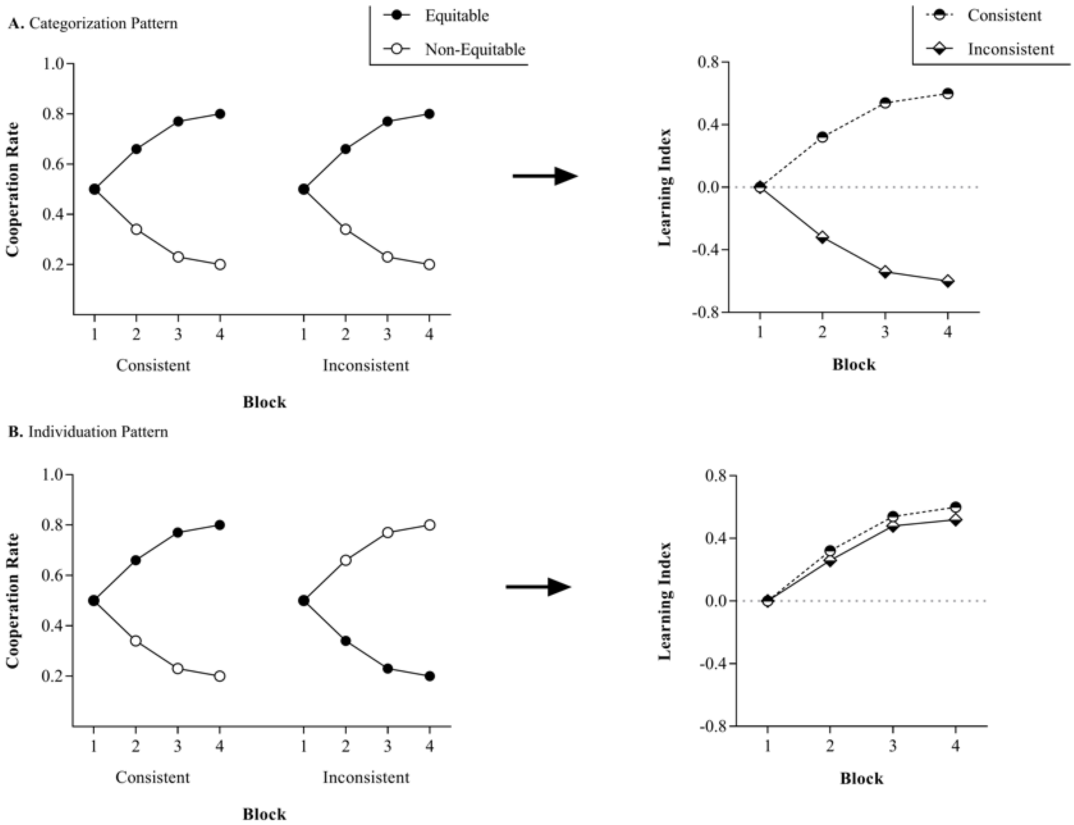

Figure III. Expected patterns of cooperation for a categorization strategy (A) and an individuation strategy (B) in the learning phase. On the left side, the dependent variable is cooperation rate for partners belonging to the equitable vs. non-equitable behavior, as a function of their individual consistency and the block of trials. A categorization strategy should be reflected in Group Behavior x Block interaction, not modulated by the consistency variable. Alternatively, an individuation strategy should be reflected in a Group Behavior x Consistency x Block interaction, with opposite patterns of cooperation for consistent vs. inconsistent partners within the same social group. On the right side, the dependent variable is a learning index computed by subtracting cooperation rate with individually non-equitable partners from cooperation rate with individually equitable partners. Therefore, a categorization strategy should be reflected in a Consistency x Block interaction, while an individuation strategy should be reflected in a main effect of the block variable, not modulated by the individual consistency.

Finally, in a transfer phase, participants were presented with new partners belonging to the two social groups manipulated, but with whom they had no prior experience, all of them cooperating in half of the trials and not cooperating in the other half. This phase allows us to verify whether participants used the information learned in the learning phase to categorize new partners accordingly, cooperating with them in accordance with the previously learned group behavior. Specifically, if participants categorize the new targets, they should cooperate more with the partners belonging to the group that was equitable in the learning phase, and less with partners belonging to the group that was non-equitable in the learning phase, as shown in Figure IVa. Conversely, if participants individuate their partners, they should cooperate with the new targets independently of their group membership, as shown in Figure IVb. The strategy employed may be analyzed by looking at participants' cooperation with new targets, as a function of these partners' group membership and the information learned in the trust game. Alternatively, we may also directly compare participants' cooperation with target from the group that was equitable vs. non-equitable, independently of which one of these groups was learned to be equitable or non-equitable in the transfer phase, as shown in Figure IV.

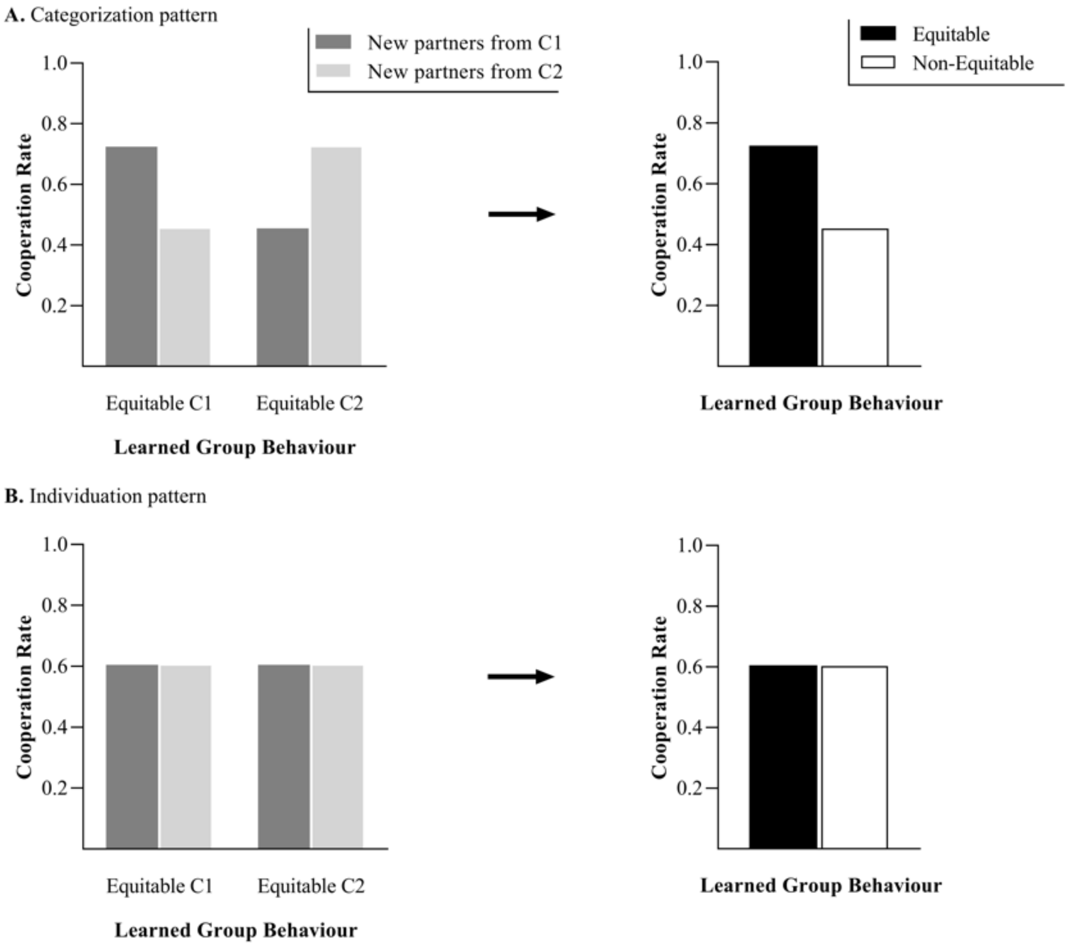

Figure IV. Expected patterns of cooperation for a categorization strategy (A) or an individuation strategy (B) in the transfer phase. On the left side, cooperation rates are displayed as a function of the partner's category membership (C1 vs. C2) and the learning about the group behavior (C1 was equitable vs. C2 was equitable). Therefore, a categorization strategy should be reflected in a Partner Category x Equitable Group interaction while an individuation strategy should be reflected in the absence of any effect. On the right side, the counterbalanced variable associating a specific category to equitable variable is ignored, and cooperation rates are displayed as a function of the partner's group behavior in the learning phase (equitable vs. non-equitable). Therefore, a categorization strategy should be reflected in a main effect of group behavior, and an individuation strategy should be reflected in the absence of any effect. In this figure, as in all the figures of the dissertation, cooperation with the equitable group are represented in black while cooperation with the non-equitable group are represented in white. Different tones of gray are used to represent cooperation with specific social categories (i.e., men, women, etc.).

With this paradigm, we explored the general aim described above according to the following specific goals:

1. To explore how motivational, cognitive-related and emotional factors interact to impact cooperation decisions at zero acquaintance, when participants know nothing about their partners. This goal was explored in the baseline phase of the studies of the present dissertation. We broadly expected that categorical cues would impact cooperation decisions to a greater extent than individual cues.

2. To analyze how the use of different categorical cues is impacted by learning across repeated interactions with the partners. During these interactions, we generally expected participants to progressively draw attention from categorical to individual information, as long as they are cognitively able and sufficiently motivated to individuate. However, because salience and previous knowledge differ between gender, ethnicity, age and emotional state, specific predictions were made according to the dimensions manipulated.

3. To examine how learning about groups' behaviors impacts decisions with new individuals from these social groups with whom participants have no prior experience. In this context, we expected participants to use the information previously learned to categorize their new partners, especially in task settings with high cognitive demands.

These goals were achieved across seven experimental series distributed in five chapters (Chapters 3, 4, 5, 6, and 7), with specific hypotheses pre-registered on Open Science Framework. Considering that one of the experimental series has already been published and several others are currently under review for their publication, each experimental series is structured as a manuscript, with its corresponding introduction, method, results and discussion. The chapters are organized as follows:

1. The Big Three (gender, ethnicity and age)

Research in social psychology has extensively demonstrated that gender, ethnicity and age are the main three axes of social categorization, granting them the label of "Big Three" (Hugenberg et al., 2010; Macrae & Bodenhausen, 2000; Stolier & Freeman, 2016). Information related to race, sex and age is accurately extracted from unfamiliar faces from early stages of face processing, allowing a fast classification of targets into their corresponding groups (Bruce & Young, 1986). Importantly, once a target has been categorized, the set of knowledge and beliefs held about his or her group shapes downstream perceivers' evaluation and assumptions about this person (Freeman & Ambady, 2011; Kawakami et al., 2017). Different processes guide category-based social judgments. On the one hand, stereotypical associations pair certain categories with specific behaviors or traits. For instance, black individuals are readily associated with aggressive behaviors, resulting in a large range of negative attitudes towards Blacks from trivial erroneous assumptions to fatal shooting decisions (Correll, Park, Judd, & Wittenbrink, 2002). On the other hand, attitude toward a specific social group is highly determined by whether or not perceivers fall into the target group, resulting in more positive (i.e., ingroup favoritism) and more individual-based (i.e., outgroup homogeneity effect) judgments for ingroups as compared to outgroups.

In **Chapter 3**, we examined ingroup favoritism (Turner, Brown, & Tajfel, 1979) and outgroup homogeneity effects (Tajfel & Wilkes, 1963) on gender, ethnicity and age dimensions. Notably, white participants played the trust game with black and white partners, female participants played with male and female partners, and both younger and older adults performed the task with older and younger partners, allowing to explore cooperation decisions about ingroup and outgroup partners in all three dimensions of categorization. We broadly expected all participants to show an ingroup favoritism at zero acquaintance, while learning would be impacted by categorical processes only for outgroup partners, through an outgroup homogeneity effect. The two experimental series of this chapter offer a global overview of social perception strategies with gender, age and ethnic ingroup and outgroup members.

2. Social power

In **Chapter 4**, we extended the results from the previous chapter by investigating male participants' strategies in social learning. The study of male participants was conducted in light of power differences between men and women, potentially affecting social perception. In fact, power affects mechanisms related to the self on the one hand, such as increased self-regulation and higher focus on goal pursuits (Guinote, 2007), and to social perception on the other hand, with a greater use of category-consistent information (Guinote & Phillips, 2010). Both mechanisms likely impact social learning, by either promoting the use of category-related information and hindering learning, or enhancing attention to relevant individual-based information with a positive impact on performance.

Because of patent social inequalities granting men a higher status than women in a large range of domains and the overlap between status and power (Fiske, 2010), we expected male and female participants to use different strategies to make inferences about their partners in the trust game. Specifically, male participants were expected to make a greater use of categorical information during learning than female participants. This chapter contributes to a better understanding of the impact of power on social perception.

3. Cognitive resources

Despite the large body of research investigating the negative consequences of social categorization (Allport, 1954; Dovidio, Glick, & Rudman, 2005; Kawakami et al., 2017), researchers unanimously acknowledge its essential resources-saving function (Macrae & Bodenhausen, 2000; Macrae & Bodenhausen, 2001; Macrae, Milne, & Bodenhausen, 1994; Quinn & Macrae, 2005). More than a strategy, social categorization has been argued to be *the* cognitive tool allowing efficient processing of social information. Certainly, the flexible allocation of attention to category-related or individual characteristics seems to be the best strategy to achieve a balanced trade-off between cognitive cost and accurate predictions.

This argument is explored in **Chapter 5** by manipulating across four experimental groups the cognitive cost of individuation, increasing the number of partners with whom participants played the trust game from 8 up to 64. The reasoning was that when individuation is too costly to be successful, participants should make use of the resources-saving tool of social perception, and hence adopt a social categorization strategy. This chapter allows us to discuss the actual role of cognitive economy in social perception.

4. Motivational factors

In line with Fiske & Neuberg's (1990) continuum model of impression formation, we further aimed at directly exploring the influence of motivation on learning. To achieve this goal, in **Chapter 6**, we adapted the trust game to manipulate whether participants made predictions about the outcomes associated with either social (i.e., humans), social-like (i.e., artificial races) or non-social (i.e., paintings) targets. Participants motivation to individuate was expected to gradually decrease from a high motivation to individuate humans to a low motivation to individuate paintings. In fact, several motives justify driven attention to individual features mostly for human targets, including higher perception of shared characteristics between human targets and the self (Ellemers, 2012), attributions of cognitive and affective states to humans but not to non-social categories (Frith & Frith, 1999), higher expectations of future interactions with humans, and epistemic motives related to the desire to understand the social world (Bodenhausen, Todd, & Becker, 2006). Therefore, we predicted a pattern of individual learning for human targets, progressively switching to categorical learning for social-like and non-social targets. With this chapter, we consider the interconnection between motivation and perceptual expertise in social decision-making and learning.

5. Emotion and social perception

Finally, in the last chapter of the present dissertation, we investigated the relationship between affective state and social categorization across two experimental series. In the first experimental series, we manipulated partners' emotional states as predictors of their cooperative behaviors in a categorical way. Notably, partners were presented portraying either a happy or an angry emotional expression in different trials. Moreover, we established counter-intuitive associations between emotional expression and cooperative behaviors such that participants had to learn that angry faces were associated with equitable behaviors and happy faces were associated with non-equitable behaviors. We also manipulated that participants played with either 8 or 32 partners, to examine the impact of differential cognitive demands on learning strategies. Importantly, with this procedure, individual identities were made completely irrelevant to perform the task. Although emotional states may be treated as categorical cues of a person's behavior, the results from this experimental series indicated that attention is withdrawn from individual identities to categorical information only under very specific circumstances.

In a second experimental series, participants learned specific group- and individual-based associations between ethnic categories and cooperative behaviors and were further induced with sadness, anger or a neutral affective state. Next, we tested whether participants expressed the categorical or the individual pre-induction learning when making decisions about new individuals from the two ethnic categories manipulated in a transfer phase. Anger was expected to prompt heuristic-based reasoning and therefore trigger a categorization strategy while sadness was expected to promote detail-oriented analyses, and therefore, a greater reliance on individual attributes (Bodenhausen, 1993; Bodenhausen et al., 1994). The data from **Chapter 7** offer promising results regarding the impact the perceiver's affective state on social perception and decision making.

CHAPTER 3. The Big Three. Gender, Ethnicity and Age

EXPERIMENTAL SERIES 1

The content of this experimental series has been published as:

Telga, M., de Lemus, S., Cañadas, E., Rodríguez-Bailón, R., & Lupiáñez, J. (2018). Category-based learning about deviant outgroup members hinders performance in trust decision making. *Front Psychol, 9*(1008). doi:10.3389/fpsyg.2018.01008

but has been slightly adapted for the present dissertation to maintain coherence regarding the name of the variables.

CATEGORY-BASED LEARNING ABOUT DEVIANT OUTGROUP MEMBERS HINDERS PERFORMANCE IN TRUST DECISION MAKING

Abstract

The present research examines whether individuation and categorization processes influence trust decisions about strangers at first and across repeated interactions. In a partial replication of the Experiment reported by Cañadas, Rodríguez-Bailón, and Lupiáñez (2015), participants played an adaptation of the multi-round trust game paradigm and had to decide whether or not to cooperate with unknown partners. Gender (Experiment 1a) and ethnicity (Experiments 1b, 2 and 3) served to create distinct social categories among the game partners, whose behaviors were manipulated at group and individual levels. At the group level, two social groups (i.e., ingroup vs. outgroup) were associated with opposite cooperative behaviors (i.e., equitable vs. non-equitable). At the individual level, consistency was manipulated by altering the cooperative behavior of one out of four members in each social group. Notably, one inconsistent individual in each group showed a pattern of cooperation opposite to the group behavior. Our data, contrary to Cañadas and colleagues' findings, suggested that ingroup partners were individuated as participants made their decisions to cooperate with them according to their individual cooperative behavior and independently of the group behavior. In contrast, decisions about outgroup partners (i.e., men in Experiment 1a and Blacks in Experiment 1b, 2 and 3) were affected by category-based thinking. At the same time, in comparison with ingroup, greater cooperation was observed with ethnic outgroup but not with gender outgroup. The consistency of our results with the previous literature on social categorization and across the three experiments suggests that they are reliable, supporting the hypothesis that categorization and individuation processes guide trust decision-making, promoting individuation mainly for ingroup and categorization among outgroup members.

Keywords: categorization, individuation, motivation, trust, outgroup homogeneity

In our daily life, plenty of situations require us to make decisions about people we do not know, from helping a beggar to hiring someone's services. When we get involved in these interactions, we surely have a confident positive expectation regarding the behavior of these people, that is, we trust them (Lewicki, McAllister, & Bies, 1998). But, once we decide to interact with them, we have to deal with uncertainty since we have no further control over the outcomes. This is why trust has often been considered as irrational or inconsistent with self-interested decisions (Berg, Dickhaut, & McCabe, 1995). Indeed, trusting someone unknown is risky given one exposes him or herself to deception or exploitation. In fact, trust has also been defined as an "intention to accept vulnerability" (Dunning, Anderson, Schlosser, Ehlebracht, & Fetchenhauer, 2014, p. 123). But it is also "an important lubricant of social system" (Arrow, 1974, p. 23) since trust promotes cooperation between individuals (Barnard, 1968; Deutsch, 1973), which in turn leads to reciprocity in addition to being rewarding on its own (Tomasello, 2009). Most theories consider that people engage in trust behaviors when their tolerance of risk has not been trespassed (Dunning et al., 2014). However, people often trust strangers with whom they have no prior experience, thus bearing a high risk of deception. (Johnson & Mislin, 2011; Wilson & Eckel, 2011). This paradox has been investigated in psychology (e.g., Balliet & Van Lange, 2013), sociology (e.g., Paxton, 2001) or political science (e.g., Wilson & Eckel, 2011). But the topic particularly caught attention among economists who have provided theories and procedures to examine how we engage in interactions involving trust (Johnson & Mislin, 2011).

The trust game (Berg et al., 1995) is a useful paradigm to investigate under which circumstances people place their trust in someone else's hands. In its classical version, participants are endowed with $10 and have to decide how much of this initial amount they will send to an anonymous partner. In a second stage, the amount sent is tripled and participants' partner can decide how much of the received money, if any, they would send back to the participant. Thus, participants are "trustors", whereas the partner is the "partner" who has the power to make a decision that affects both the trustors and themselves. From the participants' perspective, the most rational decision is to send nothing since they have no guarantee to receive something back. However, research has shown that participants do trust strangers who, in turn, reciprocate (Balliet & Van Lange, 2013; Johnson & Mislin, 2011; Wilson & Eckel, 2011). Therefore, rational decision-making based on risk attitudes is not sufficient to explain how we decide to place our trust in someone. Several social factors such as socioeconomic status (e.g., Blue, Hu, & Zhou, 2018; Bogliacino, Jiménez Lozano, & Reyes, 2018), emotion (e.g., Alguacil, Madrid, Espín, & Ruz, 2017; Tortosa, Strizhko, Capizzi, & Ruz, 2013), or face appearance (e.g., Li, Liu, Pan, & Zhou, 2017; van 't Wout & Sanfey, 2008) have been shown to affect trust decision-making at zero acquaintance. All the social variables (e.g., facial expression, gaze direction, gender, ethnicity, attractiveness) that might influence the impression formation process (Stolier & Freeman, 2016; Uleman & Kressel, 2013) can, in turn, affect the decisions being made. Understanding the processes underlying trust decisions requires understanding what factors influence social perception and impression formation.

Social stimuli are complex and contain considerable information. Body language (e.g., de Lemus, Spears, & Moya, 2012; Tiedens & Fragale, 2003), facial expression (e.g., Cañadas, Lupiáñez, Kawakami, Niedenthal, & Rodríguez-Bailón, 2016), gaze direction (e.g., Macrae, Hood, Milne, Rowe, & Mason, 2002), skin color (e.g., Sommers, 2006), gender and attractiveness (Solnick & Schweitzer, 1999) are some of the numerous cues that influence our perception and expectations about others. Processing social information is cognitively demanding so we need to deal with this information efficiently. Social categorization allows us to make sense of our social world effortlessly (e.g., Fiske & Neuberg, 1990) by using noticeable information to classify others on the basis of the diagnostic characteristics of the social groups to which they belong. Categorization is a prominent strategy when we perceive social stimuli (Brewer, 1988; Cuddy, Fiske, & Glick, 2004; Fiske, Lin, & Neuberg, 1999; Fiske & Neuberg, 1990; Macrae & Bodenhausen, 2000), but this basic tendency to attend to social information can be overcome by activating instead the motivation to focus on individuating characteristics. Indeed, several factors such as prejudice level (Lepore & Brown, 1997), personal relevance (Fiske & Neuberg, 1990), instructions (Cañadas, Rodríguez-Bailón, Milliken, & Lupiáñez, 2013), power (Goodwin, Gubin, Fiske, & Yzerbyt, 2000; Rodríguez-Bailón, Moya, & Yzerbyt, 2000), interdependence (Bukowski, Moya, de Lemus, & Szmajke, 2009) or some contextual variables (Blair, 2002) can selectively direct attention towards individual-based attributes. Given the potential negative consequences of misattribution of traits, being able to flexibly adopt an individuation or a categorization strategy is crucial for understanding the social world.

Once social categories are established, one necessarily realizes that he or she falls into some social groups (i.e., ingroups), and remains excluded from others (i.e., outgroups) (Ellemers & Haslam, 2012). These processes of self-categorization are crucial for self-perception. Depending on the context, people can categorize themselves according to different social identities (e.g., gender, ethnicity, etc.) which are associated with different emotional significance (Ellemers & Haslam, 2012; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). The salience of the social identity and the relevance of the ingroup for the self will determine how much are people willing to use certain strategies to enhance the group identity (Tajfel, 1978), broadly resulting in a more positive attitude toward ingroup than outgroup members (Tajfel & Turner, 1979). Such ingroup bias can be observed in a large range of responses, from resources distribution (Tajfel, Billig, Bundy, & Flament, 1971) to empathy (Xu, Zuo, Wang, & Han, 2009), including trust (Romano, Balliet, Yamagishi, & Liu, 2017; Tanis & Postmes, 2005; Wilson & Kayatani, 1968). Thus, the motivation to enhance or maintain a positive social identity should lead people to cooperate more with ingroup than with outgroup members (Brewer, 2008).

Beyond group identity, cooperation can lead to more global positive outcomes such as humans' survival. From an evolutionary approach, group organization allows exchange network necessary for survival (Henrich & Henrich, 2007). According to the Bounded Generalized Reciprocity theory, cooperative individuals within a group help to achieve this goal and gain the reputation of being reliable cooperators, which enhances their probability to remain part of the group (Yamagishi, Jin, & Kiyonari, 1999). Importantly, when it comes to trust decision-making in intergroup contexts, both interests in achieving a positive social identity or maximizing the groups' outcomes converge in promoting ingroup favoritism and intergroup discrimination.

As well as ingroup bias, a different consequence of social categorization is reflected in the outgroup homogeneity effect (Tajfel & Wilkes, 1963), that is, a category-based perception of the outgroup resulting in a greater perceived similarity among outgroup members than among ingroup members. This effect has been observed for both physical features directly observable (see Hugenberg, Young, Bernstein, & Sacco, 2010; Meissner & Brigham, 2001 for a review) and more complex personality traits (Freeman, Schiller, Rule, & Ambady, 2010; Linville, Fischer, & Salovey, 1989; Linville, Salovey, & Fischer, 1986). For instance, same-race faces are better recognized (Hugenberg et al., 2010) and differentially attended (Kawakami et al., 2014) than other-race faces. Therefore, the use of social categories to extract information about unknown people has important consequences for our judgments, our expectations from others, and in general the way we interact with them (Allport, 1954; Bodenhausen, Kang, & Peery, 2012; Kawakami, Amodio, & Hugenberg, 2017; Macrae & Bodenhausen, 2000). Altogether, these effects suggest that despite their cognitive efficiency, social categorization processes might also lead us to biased perception and flawed decision-making. For instance, the outgroup homogeneity effect can lead to overgeneralization, failure to distinguish among the members of the same category and stereotyping (Allport, 1954; Blair, Judd, & Chapleau, 2004; Levin, 1996, 2000; Stroessner, 1996). In cooperation settings, individuation should be a more efficient strategy, leading to more accurate predictions of people's cooperative tendency. In this research, we aim to deepen our understanding of how categorization and individuation processes are used in social interactions, and how they modulate the way we learn who is trustworthy.

In an attempt to clarify whether categorization and individuation processes affect the way we learn whether to trust unknown game partners depending on their ethnicity, Cañadas et al. (2015) conducted an adaptation of the trust game paradigm. They used the multi-round version of the trust game (King-Casas et al., 2005) in which participants interact several times with the partners. Because of these repeated interactions, the best strategy to maximize benefits is to monitor individuals' behavior and to learn as fast as possible their individual cooperation tendency. In Cañadas and colleagues' adaptation, all participants were white and played with white and black partners. The two ethnic groups were associated with opposite cooperative trends. For instance, black partners were equitable and cooperated on 75% of the trials whereas white partners were non-equitable and cooperated only on 25% of the trials. Furthermore, in each group, one individual was inconsistent with respect to the other members, that is, this person was associated with the cooperative behavior of the other ethnic group. Following the same example, one black partner was non-equitable whereas one white partner was equitable. With this procedure, participants' cooperation with the inconsistent partners is critical. If participants individuate their partners, they should cooperate with the inconsistent individual according to his or her individual cooperative behavior and independently of the group behavior. Alternatively, if participants categorize their partners, they should apply the group knowledge to the inconsistent individuals and show a similar pattern of cooperation for consistent and inconsistent partners within the same ethnic group.

Participants were expected to individuate their partners using the trial-by-trial feedback to guide their decisions as this strategy maximizes profits. Moreover, and according to previous research, this pattern was expected mainly for ingroup members, who are generally individuated to a greater extent than outgroup members (Tajfel & Wilkes, 1963). Contrary to this hypothesis, participants showed a pattern of categorization for the ingroup (i.e., white partners) and a pattern of individuation for the outgroup (i.e., black partners). Cañadas et al. (2015) argued that despite the manipulations related to the group behavior were made according to partners' ethnicity, participants may have relied on gender over ethnicity to identify and categorize their game partners, so that ethnicity did not have the relevance expected in their experiment. In the current research, we go beyond the studies reported by Cañadas and colleagues by experimentally distinguishing the effects of race and gender on trust decision-making. We adapted and replicated Cañadas and colleagues' procedure across three experiments in which we also investigated a possible effect of experimenter's ethnicity on participants' decisions.

One of the main aims of the present research was to clarify the results reported by Cañadas et al. (2015) by disentangling the effects of gender and ethnicity in a multi-round trust game task. To achieve this goal, we manipulated either partners' gender or ethnicity in different experiments, ensuring that the not manipulated dimension remained constant across all partners. Specifically, we presented participants with men and women, all belonging to the participants' ethnic group (i.e., white) in Experiment 1a, or with blacks and whites, all belonging to the participants' same gender category (i.e., women) in Experiment 1b. In Experiment 2, we focused on the ethnic categories and repeated the same experimental procedure as in Experiment 1b introducing a between-participants manipulation of the experimenter's ethnicity. This allowed us to explore a possible effect of social desirability boosted by the presence of an outgroup experimenter, which may have influenced participants' responses. Finally, in Experiment 3, we directly replicated the experiment reported by Cañadas et al. (2015) including men and women in each ethnic group, while we maintained the between-participants manipulation of experimenter's ethnicity.

We broadly expected participants to individuate their partners since they were provided with both the motivation (i.e., economic outcomes) and the means (i.e., feedback after each trial) to do so. Moreover, since the salience of gender may have confounded the identification of ethnicity as the relevant social dimension in Cañadas and colleagues' experiment, we had no strong theoretical motives to expect a replication of the data they reported. Thus, in Experiments 1 and 2, we predicted on both gender and ethnicity dimensions a stronger pattern of individuation for ingroup members than for outgroup members (Tajfel & Wilkes, 1963). Finally, in Experiment 3, we directly replicated the experimental design of Cañadas et al. (2015) while testing for a possible experimenter effect.

Experiment 1

The goal of Experiment 1 was to distinguish between the effects of gender and ethnicity in cooperation decisions. Participants (all white female) were randomly assigned to one of two experiments: in Experiment 1a gender was manipulated while ethnicity remained constant across all partners (all whites) whereas in Experiment 1b ethnicity was manipulated while gender was identical across all partners (all women).

Method

Participants. Experiments 1a and 1b were conducted concurrently and participants were randomly assigned to one of the two conditions[1]. Forty[2] Caucasian female students (mean age: 20.48 years, range: 18-26 years) participated in Experiment 1a and forty-one Caucasian female students (mean age: 20.11 years, range: 18-27 years) participated in Experiment 1b. All the participants were volunteers from the local university and took part in exchange for financial compensation proportional to their performance in the task (€5.82 on average). In these experiments and the following ones, all participants had normal or corrected to normal vision and were naïve as to the purpose of the study. Written informed consent was obtained from all participants and the experiments were conducted according to the guidelines set forth by the local university on the use of human participants in research.

Apparatus and stimuli. PCs with E-Prime 2.0 software package (Schneider, Eschman, & Zuccolotto, 2002) were used for stimuli presentation and data acquisition. Stimuli were presented on a 17-in. computer screen and consisted of full color photographs of an emotionally neutral face with a direct gaze on a gray background. The photographs were taken from the NimStim Set of Spatial Expressions (Tottenham et al., 2009) as Cañadas et al. (2015). However, given that we needed to introduce more faces of white (Experiment 1) or black partners (Experiment 1b), some of the stimuli were taken

[1] After data collection, we realized there was a mistake in the gender manipulation condition. One photograph of a male partner was repeated, such that participants played with three instead of four partners. This condition was replicated with forty new participants. Here, we present the data of the corrected new experiment. As a consequence, the original between-participants manipulation is therefore presented as two different experiments: Experiment 1a and Experiment 1b.

[2] At the time this experiment was carried out, it was not usual in our lab to perform power analyses to estimate sample size. However, a sensitivity power analysis assuming an alpha criterion of .05 and a power criterion of .80 revealed that with our sample of 40 and 41 participants respectively in Experiments 1a and 1b, the smallest effect size that could have been detected for the critical Partner Group x Group Behavior x Consistency interaction was $f = .23$.

from a different database[3]. Overall, twenty-four different photographs were used to represent the partners, corresponding to eight white women (Experiments 1a and 1b), eight white men (Experiment 1a) and eight black women (Experiment 1b).

Procedure. As a cover story, participants were told that they would take part in a study about economic decision-making and were motivated to be as accurate as possible as they would be economically rewarded proportionally to their performance in the task.

Participants played a multi-round trust game adapted from King-Casas et al. (2005). Each trial consisted of a round with a virtual partner represented by one of the twenty-four faces that served as stimuli. At the beginning of the trial, participants were presented for 190 ms with a euro symbol (i.e., "€") indicating the endowment of €1. Then, a fixation cross was presented during 500 ms followed by the photograph of the partner (4.30º x 3.34º) for 1500 ms. Participants were asked to indicate whether or not they cooperated with the partner by pressing the '1' key to cooperate and the '0' key not to cooperate. If the participants cooperated, the partner of this round would receive the initial €1 multiplied by 5 (i.e., €5) and decide to either keep the whole money for him or herself or to cooperate by sending back half of the amount received (i.e., €2.50) to the participant. If the participant did not cooperate, he or she would keep the initial €1 and the partner of this round would receive nothing. Once participants responded (or after 1500 ms), a second fixation cross was presented for 500 ms followed by visual feedback displayed for 1000 ms. Three symbols (i.e., "o", "*" and "#") in three different colors (i.e., blue, brown and green) were used to display the feedback, corresponding the following meanings: "You have kept the money", "You have cooperated and your partner also cooperated" and "You have cooperated but your partner has not cooperated". The association between symbols, colors and their meanings was counterbalanced across participants. The message "You did not answer" was displayed when participants did not answer after 1500 ms.

[3] The photographs were ceded by the Social Cognition lab of York University,

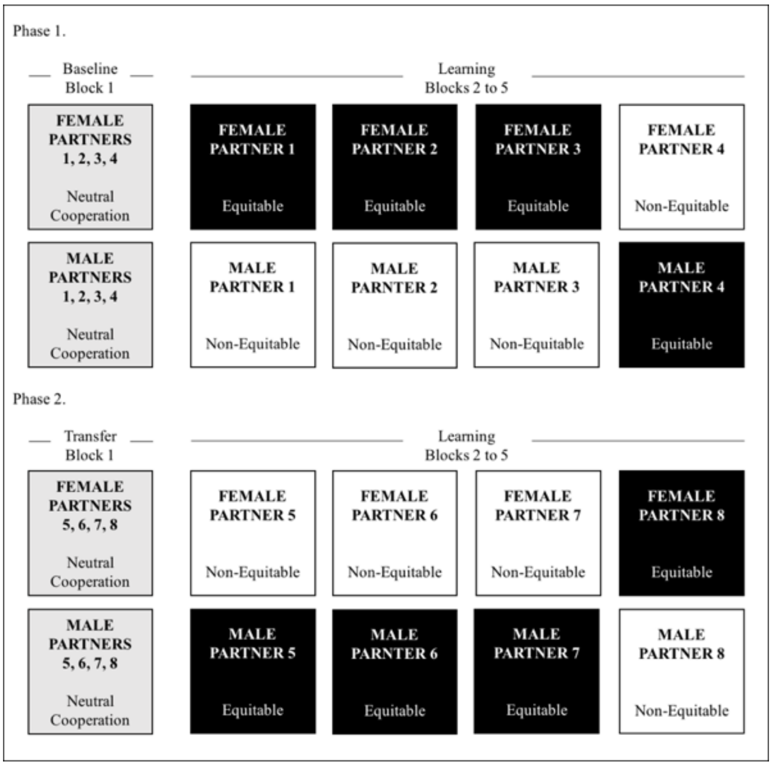

Figure 1.1. Example of the manipulation of partners' cooperative behaviors in Experiment 1a, adapted from Cañadas et al. (2013).

Participants played forty times with each of the sixteen partners, resulting in 640 trials. The task was divided into two phases of five blocks, as shown in Figure 1.1. The first block of the first phase was the baseline, in which all partners cooperated on 50% of the trials. This block was introduced to verify whether participants were biased toward one of the social groups manipulated, which would be reflected in a higher cooperation rate with one of the two groups.

Next, Blocks 2 to 5 corresponded to a learning phase, in which we introduced specific manipulations of partners' behaviors. At the group level, ingroup and outgroup (Experiment 1a: women vs. men; Experiment 1b: blacks vs. whites) displayed opposite patterns of cooperation. For instance, when one group (e.g., the ingroup) was equitable (i.e., cooperating on 75% of the trials), the other one (e.g., outgroup) was non-equitable (i.e., cooperating in 25% of the trials). Moreover, partners' behavior was manipulated at the individual level such that consistent individuals (3 out of 4 members) displayed the group behavior, whereas inconsistent individuals (1 out of 4 members) displayed a pattern of cooperation opposite to the group behavior. The faces associated with the inconsistent condition were counterbalanced across participants.

The first block of the second phase was the transfer block, in which participants were presented with 8 new partners (4 from each social category) with whom they had no prior experience. Again, these partners started cooperating on 50% of the trials, which allowed us to analyze whether the learning from the first phase would generalize to new individuals in the second phase. This would be reflected in a biased cooperation rate toward the group that was equitable in the learning phase.

Finally, Blocks 2 to 5 corresponded to a second learning in which we reversed the association established between the social groups manipulated and the group behavior. For instance, if the ingroup was equitable in the first phase, the new partners from the ingroup were non-equitable in the second phase. The order in which ingroup or outgroup started being equitable was counterbalanced across participants. This procedure was used to have a full within-participants design regarding the variables of interest.

At the end of the experiment, participants were endowed according to their percentage of accuracy[4] in the task.

Design. In this and the following Experiments, we manipulated four within-participants independent variables (IV) corresponding to partners' social group membership (i.e., partner group: ingroup vs. outgroup), partners' group behavior (i.e., group behavior: equitable vs. non-equitable), partners' consistency with respect to the group behavior (i.e., consistency: consistent vs. inconsistent), and the block of trials numbered from the first presentation of the partners (i.e., blocks: 1-5 in both phases). The dependent variable (DV) was participants' cooperation rate.

Results

To ensure that participants performed the task without paying attention to the faces, three participants (all from Experiment 1b) were excluded from the analyses for having a mean RT shorter than 200 ms in more than 50% of the trials (which was considered as a signal that participants did not thoroughly respond to faces). Furthermore, we applied the same criterion as Tortosa, Lupiáñez, and Ruz (2013) and excluded trials with RTs shorter than 200 ms from the analyses (4% in Experiment 1a and 5 % in Experiment 1b). First, the cooperation rates for Blocks 1 of each phase (in which participants were exposed to unknown faces) will be presented, followed by the analyses of cooperation rates across Blocks 2 to 5 of both phases, from which it can be deduced whether participants learned about partners as individuals or as members of their group.

[4] Accurate trials were those in which participants' decision matched partners' cooperation or deception.

Baseline and Transfer: Block 1 of each phase. To verify whether participants' cooperation decision was spontaneously biased toward one or the other social group, cooperation rates in **Block 1** of the first phase (i.e., baseline) were subjected to a repeated-measures Analysis of Variance (ANOVA) with partner group as a within-participants factor. In Experiment 1a, when gender was manipulated, we observed a significant effect of partner group, $F(1, 39) = 4.63$, $p = .04$, $\eta_p^2 = .11$, showing that participants cooperated more with women (the ingroup, $M = .68$, $SD = .16$, CI: .63-.73) than with men (the outgroup, $M = .61$, $SD = .19$, CI: .54-.66), thus showing an ingroup favoritism. In Experiment 1b, in which ethnicity was manipulated, the main effect of partner group was also significant, $F(1, 37) = 7.58$, $p = .01$, $\eta_p^2 = .17$. However, in this case, participants cooperated more with black (the outgroup, $M = .70$, $SD = .18$, CI: .64-.76) than with white (the ingroup, $M = .60$, $SD = .22$, CI: .53-.67) partners, thus showing an outgroup favoritism.

Further, to verify whether the experience with partners in the first phase of the experiment (Blocks 2 to 5) impacted participants' decision with new individuals from the same social categories, cooperation rates in **Block 1** of the second phase was subjected to a mixed-design ANOVA with partner group (ingroup vs. outgroup) as a within-participants variable and equitable group (ingroup vs. outgroup) as a between-participants variable. We observed no significant effects, all Fs < 1.4, ps > .24.

Learning: Blocks 2-5 of both phases.

Experiment 1a. (White) women vs. men partners.

To verify how gender categorization or individuation strategies impacted learning, cooperation rates in Blocks 2 to 5 were subjected to a repeated-measures ANOVA with partner group (ingroup vs. outgroup), group behavior (equitable vs. non-equitable), consistency (consistent vs. inconsistent) and block (2-5) as within-participants variables. A pattern of categorization should be reflected in a main effect of group behavior (and absent Group Behavior x Consistency interaction), indicating participants display a similar pattern of cooperation with consistent and inconsistent partners within the same social category. Alternatively, a pattern of individuation should be reflected in a Group Behavior x Consistency interaction, indicating that participants show opposite patterns of cooperation with consistent and inconsistent partners from the same group.

We observed a significant Group Behavior x Consistency interaction, $F(1, 39) = 42.31$, $p < .001$, $\eta_p^2 = .52$ indicating that participants' decision to cooperate was not solely led by partners' group but also by their individual attributes. Moreover, this interaction was moderated by the block variable, $F(3, 117) = 18.67$, $p < .001$, $\eta_p^2 = .32$, suggesting that this pattern of individuation increased across the blocks as a result of learning, as shown in Figure 1.2. We also found a significant Partner Group x Group Behavior x Consistency interaction, $F(1, 39) = 4.43$, $p = .04$, $\eta_p^2 = .10$, indicating that despite participants learned to individuate both ingroup, $F(1, 39) = 54.59$, $p < .001$, $\eta_p^2 = .58$, and outgroup $F(1, 39) = 21.54$, $p < .001$, $\eta_p^2 = .37$, the pattern of individuation was stronger for the former, as shown in Figure 1.2.

Experiment 1b. (Female) black vs. white partners.

The same analysis was conducted in Experiment 1b, in which ethnicity was manipulated. Again, we found a significant Group Behavior x Consistency x Blocks interaction, $F(3, 111) = 19.27$, $p < .001$, $\eta_p^2 = .34$, showing that participants individuated their game partners and this tendency increased across the blocks of trials. The Partner Group x Group Behavior x Consistency interaction was marginal, $F(1, 37) = 2.97$, $p = .09$, $\eta_p^2 = .07$, suggesting that although the pattern of individuation was significant for both ingroup, $F(1, 37) = 43.35$, $p < .001$, $\eta_p^2 = .54$, and outgroup, $F(1, 37) = 34.78$, $p < .001$, $\eta_p^2 = .49$, it seemed to be stronger for the ingroup, as shown in Figure 1.2.

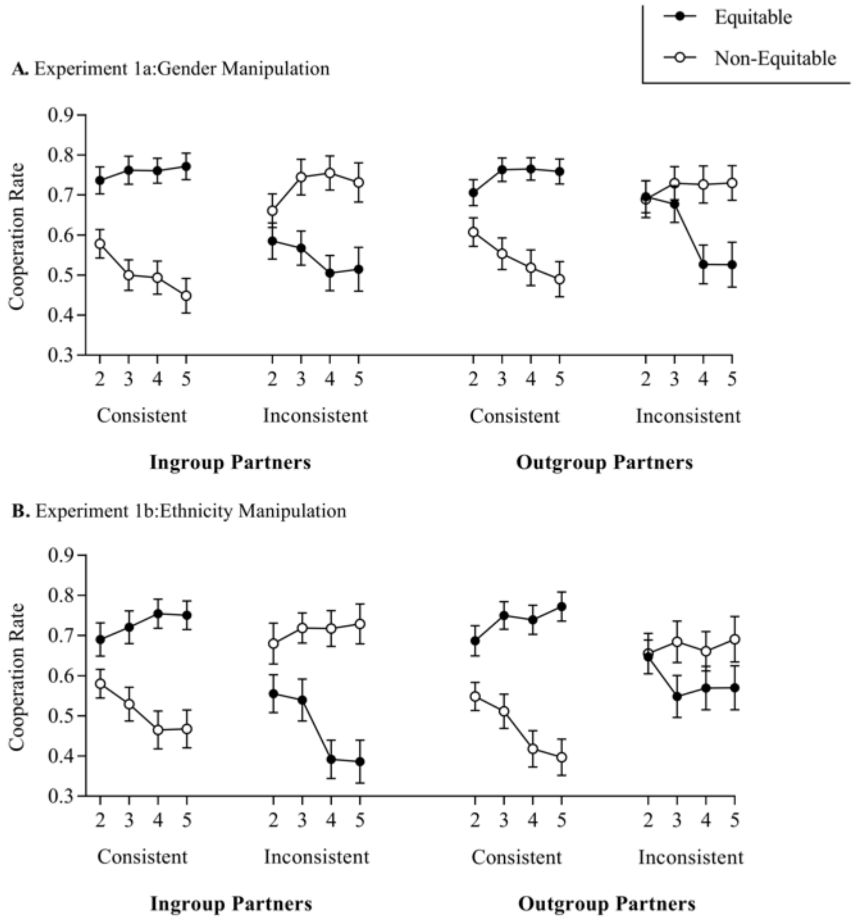

Figure 1.2. Participants' cooperation rates in Experiment 1 as a function of partners' group, consistency and blocks of trials. In this figure and the following ones, Blocks 2 to 5 referred to the learning blocks in both phases of the experiment and error bars represent the standard error of the mean.

Discussion

Experiment 1 aimed at exploring the effects of gender and ethnicity on participants' decisions to trust unknown game partners in a multi-round trust game. We expected participants to spontaneously identify and use social categories to guide their decisions in the baseline, such that they would trust more ingroup (i.e., women in Experiment 1a and Whites in Experiment 1b) than outgroup members (i.e., men in Experiment 1a and Blacks in Experiment 1b). The data were consistent with this hypothesis in Experiment 1a since female participants cooperated more with women than with men. Nonetheless, we found the opposite pattern in Experiment 1b, when ethnicity was manipulated, as white participants cooperated more with black than with white partners.

While inquiring ourselves about this pattern of results, we noticed that the experimenter in this study was a black woman, something very unusual in the context in which the experiment took place. Therefore, it is possible that this pattern might be a result of the *experimenter effect* (Sattler, 1970). In fact, Lowery, Hardin, and Sinclair (2001) argued that the presence of a black experimenter might be a tacit form of social influence that increases participants' social regulation. According to them, social regulation is determined by presumptions about the attitude of others and relationship-specific motives. In Experiment 1a, the experimenter was an ingroup member according to the salient social dimension manipulated in the task (i.e., gender). In contrast, in Experiment 1b, the experimenter was an outgroup member considering the relevant dimension for performing the task (i.e., ethnicity). Therefore, participants might have been more likely concerned about possible discrepancies between their own attitudes toward black people and the experimenter's ones in Experiment 1b, than in Experiment 1a. This may have led them to be particularly careful not to be perceived as holding prejudices against black people, such that participants' cooperation rates with black partners in the baseline (Block 1 of the first phase) may have been artificially enhanced because of the presence of a black experimenter. This issue is addressed in Experiments 2 and 3.

Moreover, given that participants played several times with the same partners, the intrinsic motivation reinforced by economic outcomes, and the feedback after each trial, we expected ingroup members (e.g., white women in both experiments) to be individuated. Conversely, given that outgroup members are perceived more categorically (i.e., outgroup homogeneity effect, Tajfel & Wilkes, 1963) we expected this pattern of individuation to be weaker for outgroup members. The data supported our hypotheses in both Experiments 1a and 1b. Although participants made their cooperation decision according to their partners' individual behavior, cooperation with male (Experiment 1a) and black (Experiment 1b) partners were somehow impacted by the group behavior. Specifically, despite the 40 interactions with each partner, participants were less efficient at making decisions about outgroup partners when their individual behavior differed from the rest of their group (i.e., inconsistent individuals). This suggests that participants' cooperation decisions were impacted by the group behavior to a greater extent for gender outgroup than ingroup partners. This effect was only marginally significant in Experiment 1b, in which partners' ethnicity was manipulated. It is possible that the experimenter effect described above affected not only participants' spontaneous cooperation attitudes with black partners (Block 1 of the first phase), but also their learning about black partners' cooperative behaviors (Blocks 2 to 5 of both phases). As a consequence of an increased social regulation, participants may have paid more attention to black partners during learning, and therefore showed a better performance at the task. Interestingly, this tendency is observed mainly at the beginning of the task, right after interacting with the black experimenter. In fact, the same analysis conducted on the last two blocks of trials confirmed that the tendency to categorize black partners was stronger at the end of the task, $F(1, 37) = 6.53$, $p = .02$, $\eta_p^2 = .15$, when participants were no longer cooperating in a trial/error dynamic but rather according to what they had learned in the previous blocks of trials. The possibility that black partners received greater attention is verified in Experiments 2 and 3.

Overall, we could not provide empirical support to the results reported by Cañadas et al. (2015), who found the opposite pattern of data, (i.e., categorization of the ingroup and individuation of the outgroup). In both Experiments 1a and 1b, the ingroup-outgroup distinction was clearer than in Cañadas and colleagues' as the not manipulated dimension remained constant across all partners. Therefore, the experimental design of the present research may have been more adequate to Experiment learning strategies in intergroup contexts.

Finally, in the transfer block (i.e., Block 1 of the second phase), in which new partners were cooperated on 50% of the trials, we observed that participants did not differ in their cooperation with new ingroup and outgroup partners. The fact that we found no effect of the equitable group variable (i.e., whether ingroup or outgroup was equitable in the first phase of the experiment) confirms that participants did not learn to categorize, and therefore did not apply the categorical knowledge manipulated in the first phase of the experiment to new individuals. However, they did learn something as the ingroup (Experiment 1a) or outgroup (Experiment 1b) favoritism found in the baseline disappeared after the learning phase. Participants' individual-based learning in Blocks 2 to 5 of the first phase seemed to be sufficient for social categories to become less significant as a criterion for decision-making. As a consequence, participants started individuating new partners from the first interaction with them (i.e., first block of the second phase), in contrast with the baseline.

In summary, Experiment 1 did not replicate the results reported by Cañadas et al. (2015) and rather indicated that when making decision about strangers, ingroup members are individuated whereas outgroup members are somehow categorized. Nonetheless, a possible experimenter effect causing the unexpected outgroup favoritism found in

Experiment 1b still needs to be clarified. Experiment 2 aimed to test the experimenter effect and its possible consequences on participants' perception of black partners.

Experiment 2

Experiment 2 investigated a possible effect of experimenter's ethnicity which may have impacted the data observed in Experiment 1b. Thus, we replicated the procedure of Experiment 1b manipulating between groups the experimenter's ethnicity (black vs. white). We expected participants to show in the baseline (i.e., Block 1 of the first phase) an ingroup favoritism with a white experimenter, and an outgroup favoritism with a black experimenter, as in Experiment 1b. Moreover, in line with Cañadas et al. (2015), we aimed at investigating whether the repeated interactions with the partners in the trust game could affect the way they were perceived by participants in terms of trustworthiness and basic dimensions of social perception. Therefore, we included new measures of impression about partners. Particularly, we expected partners who were individually equitable to be perceived as more trustworthy, and generally more positively than partners who were not equitable. We also expected these data to echo the pattern from the trust game task in that the discrimination between equitable and non-equitable partners should be better for ingroup than for outgroup members. Finally, in line with the experimenter effect hypothesis, we expected black partners to be particularly attended and individuated, with a black experimenter, but not with a white experimenter. The hypotheses, methods and analyses of this experiment were registered before data collection on Open Science Framework (osf.io/6dqs7)

Method

Participants. Undergraduates from the local university were invited to participate in the experiment. Outside the classrooms, a paper-pencil list was handed over to several professors for students to sign up. This way, we ensured that participants had no contact with the experimenters before coming to the lab. For ethical reasons, men and foreign students were allowed to participate in the experiment but were excluded from the analyses so we could test our hypotheses and exclude the possibility of cultural or gender biases.

On the basis of the analyses of Experiment 1b, we calculated an estimation of the adequate sample size with G*Power program v. 3.1.9.2. (Faul, Erdfelder, Lang, & Buchner, 2007) to replicate the critical interaction corresponding to a better learning for ingroup vs. outgroup members, with an alpha value of .05 and an estimated power of .90. The estimated effect size *f(V)* = .42 corresponded to the effect size found when the interaction reached significance, that is, when we included only the two last blocks of trials in the analysis (see discussion of Experiment 1). We found that a sample of 39 participants per experimental group (*N* = 78) would be sufficient to replicate the critical interaction. In the end, 107 undergraduates (29 men/foreigners were excluded in line with our criterion of exclusion) (mean age: 19.23 years, range: 18-34 years) participated in exchange for financial compensation according to their accuracy in the task (€5.81 on average).

Apparatus and stimuli. Apparatus and stimuli were identical as in Experiment 1b.

Procedure. The general procedure was identical to Experiment 1b, except that we introduced a post-interaction evaluation of participants' impressions about their partners. After performing the trust game, participants were asked to evaluate each of the 16 partners on different social dimensions, as described in the next section.

Participants were distributed in the two experimental conditions (Experimenter: black vs. white) in the following way: Experimenters A (white woman) and B (black woman) ran the experiment on Tuesdays and Thursdays, and Mondays and Wednesdays respectively, until reaching the sample size estimated as adequate. Experimenters provided participants with identical instructions, which they received in the same lab.

Measures. After performing the trust game, participants were asked to evaluate each one of the 16 partners they had played with on a scale ranging from 1 "not at all" to 7 "totally" on the following dimensions: attractiveness, trustworthiness, competence, threat, and warmth, in line with Cañadas et al. (2015). To get a complementary measure of subjective categorization, participants were also asked about their subjective perception of similarity of each member with the rest of their group on a scale from -3 "very distinctive" to +3 "very indistinctive", and the frequency with which they have been presented with the partners (1 "less", 2 "the same", 3 "more"). Moreover, participants were asked to indicate the perceived frequency of presentation and cooperation rate of each group in percentage.

Finally, apart from their impression about the partners, participants were asked about their own general perception of intragroup similarity on a four-items scale ranging from 1 "not at all" to 7 "totally" with two items of low similarity ("It is easy to differentiate between black/white people", "Most of black/white people are different"), and two items of high similarity ("It is hard to differentiate between black/white people", "Most of black/white people are alike"). Two questionnaires of explicit prejudice toward blacks and women were included at the end of the experiment, but did not provide valuable information since participants showed extreme rates, likely for social desirability effects. Therefore, these data are not analyzed and not included in further experiments.

Results

In line with Experiment 1, one participant was excluded from the analyses for having mean RT shorter than 200 ms in more than 50% of the trials, leaving in 77 participants for the analysis. Trials with RTs shorter than 200 ms (7%) were also excluded. Following the same strategy for the analyses of Experiments 1a and 1b, we first analyzed the cooperation rates in Blocks 1 of each phase (i.e., learning and transfer) in which participants were exposed to unknown faces, and then the cooperation rates resulting from the repeated interactions with them across Blocks 2 to 5 of both phases.

Baseline and Transfer: Block 1 of each phase. Cooperation rates in Block 1 of the first phase were introduced in a mixed-design ANOVA with partner group (ingroup vs. outgroup) as a within-participants factor and experimenter (black vs. white) as a between-participants variable. As in Experiment 1, we observed a main effect of partner group indicating that participants cooperated more with black (M = .71, SD = .16, CI: .67-.75) than with white partners (M = .62, SD = .18, CI: .57-.66), $F(1, 75) = 14.14$, $p < .001$, $\eta_p^2 = .16$. This effect was not moderated by the ethnicity of the experimenter, $F < 1$, $p >$ .50.

Furthermore, we aimed at verifying whether participants used their prior experience with partners to make decision about new partners from the same category in the transfer. Thus, we conducted a 2 (partner group: ingroup vs. outgroup) x 2 (equitable group: ingroup vs. outgroup) x 2 (experimenter: black vs. white) mixed-design ANOVA. As in Experiment 1, we found no significant effect, all Fs < 1.20, ps > .28.

Learning: Blocks 2 to 5 of both phases. A mixed-design ANOVA on cooperation rate was conducted with partner group (ingroup vs. outgroup), group behavior (high vs. low), consistency (consistent vs. inconsistent) and blocks (2-5) as within-participants factors, and experimenter (black vs. white) as a between-participants variable. We found a significant Group Behavior x Consistency x Block interaction, $F(3, 225) = 27.38$, $p < .001$, $\eta_p^2 = .27$, indicating that, as in Experiment 1, participants used an individuation strategy relying across blocks not only on partners' group behavior but also on their consistency.

We also observed a Partner Group x Group Behavior x Consistency interaction, $F(1, 75) = 3.70$, $p = .058$, $\eta_p^2 = .05$. Although the Group Behavior x Consistency interaction was significant for both ingroup, $F(1, 75) = 108.52$, $p < .001$, $\eta_p^2 = .59$, and outgroup, $F(1, 75) = 85.15$, $p <.001$, $\eta_p^2=.53$, the strategy of individuation was clearer for the former, as shown in Figure 1.3, and observed in Experiment 1.

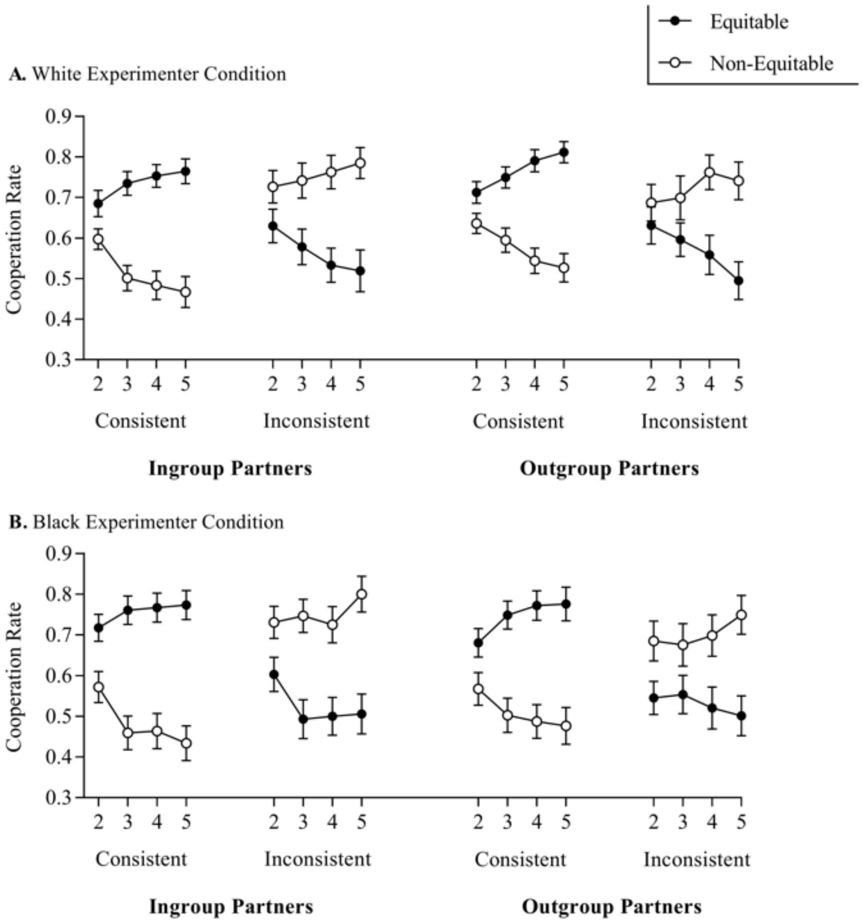

Figure 1.3. Participants' cooperation rates in Experiment 2 as a function of partners' group, Consistency and blocks of trials. (Upper panel (A) is White Experimenter Condition, and lower panel (B) is Black Experimenter Condition).

Impression about partners. To analyze participants' perception of their partners at the group level, scores on the perceived intragroup similarity, perceived frequency of presentation and perceived cooperation rates of each group (i.e., ingroup vs. outgroup) were introduced in different 2 (ethnicity: blacks vs. whites) x 2 (experimenter: black vs. white) mixed-design ANOVAs. We observed that participants perceived black partners as more similar to each other than white individuals. Interestingly, they also perceived that black partners were presented more often and were more cooperative than white partners, as shown in Table 1.1.

Then, we examined participants' perceptions of partners at the individual level. To verify whether the 5 dimensions measured (trustworthiness, attractiveness, competence, warmth and perception of threat) could be classified into components, we conducted several Principal Components Analyses (PCA) with an oblimin rotation. Each PCA was conducted in one of the condition resulting of the combination of our three within-subjects variables (partner group, group cooperation and consistency), thus resulting in eight PCA. All Bartlett's test of sphericity were significant ($p < .001$, smaller $\chi^2(10) = 49.11$), and Kaiser-Meyer-Olkin rates were high (smaller KMO = .55), thus it was acceptable to proceed with the analyses. In most of the PCAs (5 out of 8), the first component was formed by the items trustworthiness, attractiveness, competence and warmth while the fifth item measuring perception of threat was left in a second component. A reliability analysis revealed that Cronbach's alphas were acceptable for the four items of the first component across the eight conditions (range: .59 to .81). Therefore, we averaged participants' ratings of trustworthiness, attractiveness, competence and warmth to form the first component named impression, and separately analyzed their ratings at the scale measuring perception of threat.

Different within-subjects ANOVAs were conducted to examine participants' impression about partners (Component 1) as well as the perception of threat (Component 2). A significant Group Behavior x Consistency interaction was found on impression, $F(1, 75) = 19.75$, $p < .001$, $\eta_p^2 = .21$, indicating that participants had a more positive impression about partners who were individually equitable ($M = 4.03$, $SD = 1.01$), than partners who were individually not equitable ($M = 3.62$, $SD = 1.13$). Moreover, the Partner Group x Group Behavior x Consistency interaction was marginal, $F(1, 75) = 3.16$, $p = .08$, $\eta_p^2 = .04$, echoing the results found in the trust game: despite participants learned to individuate both outgroup, $F(1, 75) = 9.59$, $p = .003$, $\eta_p^2 = .11$, and ingroup, $F(1, 75) = 22.06$, $p < .001$, $\eta_p^2 = .23$, the effect was larger for the latter. Finally, a main effect of Experimenter $F(1, 75) = 5.56$, $p = .021$, $\eta_p^2 = 0.07$, showed that overall participants had a more positive impression about partners when the experimenter was black ($M = 4.03$, $SD = 1.03$) than when she was white ($M = 3.62$, $SD = 1.10$).

On the second component, perception of threat, a significant Group Behavior x Consistency interaction was found, $F(1, 75) = 15.99$, $p < .001$, $\eta^2_p = .18$, revealing that in line with previous data, participants perceived partners individually equitable as less threatening ($M = 2.33$, $SD = 1.30$) than partners individually non-equitable ($M = 2.76$, $SD = 1.57$).

Table 1.1 Impression about partners in Experiments 2 and 3. Means and (standard deviations) of participants' rates in the post-interaction evaluation of their impression about the social groups with whom they had interacted. *F*s and *p*s values correspond to the main effect of partners' social group (ingroup vs. outgroup) on participants' rates in each scale.

Experiment 2				
	Ingroup	Outgroup	*F*	*p*
Easy to discriminate	5.43 (1.41)	5.01 (1.48)	8.25	.01**
Hard to discriminate	1.89 (.80)	2.34 (1.18)	14.79	<.001***
% Cooperation	42.69 (13.72)	56.70 (15.37)	24.36	<.001***
% Presentation	46.70 (11.61)	52.19 (12.43)	5.32	.02*
Experiment 3				
	Ingroup	Outgroup	*F*	*p*
Easy to discriminate	5.42 (1.39)	4.92 (1.70)	5.3	.025*
Hard to discriminate	2.08 (1.14)	2.23 (1.34)	.8	0.38
% Cooperation	47.54 (14.91)	57.98 (14.66)	9.51	.003**
% Presentation	47.67 (11.23)	56.93 (10.76)	12.12	.001**

Discussion

Experiment 2 replicated Experiment 1b while examining the effect of the experimenter's ethnicity on participants' decision whether to trust black and white unknown game partners.

We found that in the baseline, participants spontaneously cooperated more with black than with white partners, independently of the experimenter's ethnicity. Therefore, the presence of a black experimenter is not sufficient to explain the unexpected outgroup favoritism. Despite it has been repeatedly argued that ingroup members are perceived more positively than outgroup members (e.g., Brewer, 1981; Brewer, 2007), it is not the first time that an outgroup favoritism is reported in trust decisions. Tortosa, Lupiáñez, et al. (2013) conducted a series of two trust game experiments in which white participants played with black and white partners. In the first one, participants did not differ in their cooperation with black and white partners. However, in the second experiment, the authors found a tendency to cooperate more with black than with white partners. Interestingly, they also observed an implicit negative bias towards black people in a Blacks/Whites Implicit Association Task (Greenwald, McGhee, & Schwartz, 1998). The authors' interpretation was that in the trust game, participants attempted not to show their prejudice-related biases in an explicit task, by appearing more cooperative with black individuals. A similar conclusion may be drawn from our data. Despite the apparent positive attitude towards black partners in the first block of the experiment, participants learned better the cooperative trends of white as compared to black partners. Therefore, it is possible that the higher cooperation rates toward black partners reflected an active effort to be perceived as egalitarian (Dunton & Fazio, 1997; Kawakami, Dunn, Karmali, & Dovidio, 2009; Maddux, Barden, Brewer, & Petty, 2005; Nosek, Hawkins, & Frazier, 2011).

We also replicated the outgroup homogeneity effect found in Experiment 1. Notably, we found that participants' decision (not) to cooperate was made according to partners' individual behavior when they were white. However, participants were less accurate to predict black partners' individual behaviors when they were inconsistent with the group behavior. Interestingly, these results were not moderated by the experimenter's ethnicity. Therefore, the discrepancy between our pattern of results and the data reported by Cañadas et al. (2015) is unlikely due to the experimenter effect.

In fact, our results are rather consistent with the prior literature stating that outgroup categorized to a greater extent than ingroups (Billig & Tajfel, 1973; Haslam, Oakes, & Turner, 1996). This interpretation is supported by the results from the post-interaction impression about partners, since participants indicated they perceived black individuals as more similar to each other than white people. Moreover, the general perception of black partners within the game was somehow biased in that participants also perceived that black partners were presented more often and were more cooperative than white people. This was not the case since all partners were presented with the same frequency and overall, the groups were evenly matched in cooperative behaviors. According to Fiske (1980), unusual or extreme stimuli are more salient. Given the scarcity of black people in our social context, they may have received more attention in the task such that participants overestimated the frequency of presentation or the cooperative behaviors of black partners. Still, this general striking effect did not encourage more individuating processes during learning.

Finally, the individuation strategy observed in the trust game was confirmed in the post-interaction evaluation of partners. Participants' perception of partners was affected by their individual behavior trend in the two components evaluated (positive impression and perception of threat), although this individuation strategy tended to be greater for ingroup partners, as reflected in the marginal three-way interaction, in line with the data observed in the trust game.

In summary, Experiment 2 allowed to replicate the data from Experiment 1b and to rule out the presence of an experimenter effect. However, it remains unclear whether the discrepancy between our results and the data reported by Cañadas et al. (2015) is explained by the overlap between gender and ethnicity in their Experiment. We address this issue in Experiment 3.

Experiment 3

Across Experiments 1 and 2, we found that despite participants mostly individuated their partners, this pattern was poorer when it came to outgroup members. These results are consistent with our hypothesis and the literature on social cognition (Tajfel & Wilkes, 1963) but inconsistent with the data reported by Cañadas et al. (2015). Thus, we decided to verify whether the confusion between gender and ethnicity among partners accounted for the pattern of data reported by Cañadas and colleagues. We conducted an exact replication of their experiment in which men and women were presented in each ethnic group, while controlling for the ethnicity of the experimenter as done in Experiment 2. If having a white experimenter is crucial for replicating the pattern of data observed by Cañadas and colleagues, we should replicate it with a white experimenter whereas we should obtain the pattern of data observed in Experiments 1 and 2 with a black experimenter. The hypotheses and method of this experiment were registered before data collection on Open Science Framework (osf.io/tz7c8).

Method

Participants. The sample selection was identical as in Experiment 2, except that data from men were included in the analyses and only data from foreign students were excluded. This change in the sample selection was introduced since in this Experiment, only ethnicity determined partner's group membership with respect to participants. As gender was no longer relevant for discriminating between ingroup and outgroup members among partners, neither was it among participants.

We calculated an estimation of the sample size necessary in order to replicate the data reported by Cañadas et al. (2015) using the same parameters as described by the authors (alpha value of .003, estimated power of .90 and estimated effect size *f(V)* = .65) with G*Power program (Faul et al., 2007). We found that a minimum of 29 participants per experimental group ($N = 58$) would be sufficient to replicate the significant three-way interaction corresponding to a better learning for outgroup than for ingroup members. Finally, 67 undergraduates (62 excluding foreigners) (11 males, mean age: 20.21 years, range: 18-27 years) voluntarily participated in exchange for financial compensation according to their performance in the task (€5.61 on average).

Apparatus and stimuli. The same photographs of eight white and eight black partners (half men in each ethnic group) used by Cañadas et al. (2015) were used in this experiment. As for Experiment 2, a white and a black experimenter were provided with the same instructions to the participants, in the same lab, and alternated weekly until reaching the right number of participants in each group. In order to reduce the duration of the task and given that the dimensions measured in Experiment 2 provided similar information, the dimensions evaluated in this Experiment were reduced to trustworthiness and attractiveness, as in Cañadas et al. (2015).

Results

Four participants were excluded from the analyses for having a mean RT faster than 200 ms in more than half of the trials, leaving in 58 participants for the analyses. Trials with RTs shorter than 200 ms (11%) were also excluded from the analyses. Analyses of Block 1 of each phase (i.e., baseline and transfer) are presented, followed by the analyses of cooperation rates during learning (i.e., Blocks 2-5 of both phases). All the analyses were identical to Experiment 2.

Baseline and Transfer: Block 1 of each phase. A mixed design ANOVA with partner group as a within-subjects variable and experimenter as a between-participants factor revealed a main effect of partner group indicating that participants cooperated more with outgroup ($M = .66$, $SD = .23$, CI: .60-.72) than with ingroup members ($M = .60$, $SD = .22$, CI: .54-.65), $F(1, 56) = 4.61$, $p = .04$, $\eta_p^2 = .08$, independently of the experimenter, $F(1,56) = .08$, $p = .78$, in Block 1 of the first phase. A 2 (partner group: ingroup vs. outgroup) x 2 (equitable group: blacks vs. whites) x 2 (experimenter: black vs. white) mixed-design ANOVA on cooperation rates in Block 1 of the second phase showed no significant effect, $Fs < 1.33$, ps > .25.

Learning: Block 2-5 of both phases. Cooperation rates in Blocks 2 to 5 of both phases were introduced in a mixed-design ANOVA with partner group (ingroup vs. outgroup), group behavior (high vs. low), consistency (consistent vs. inconsistent) and block (2-5) as within-participants factors, and experimenter (black vs. white) as a between-participants variable. Once again, we found a Group Behavior x Consistency x Block interaction, $F(3, 165) = 15.04, p < .001, \eta_p^2 = .22$, showing a pattern of individuation increasing across blocks. Moreover, the Partner Group x Group Behavior x Consistency interaction was again replicated in this experiment, $F(1, 55) = 4.95$, $p = .03$, $\eta_p^2 = .08$, indicating that despite participants individuated both ingroup, $F(1, 55) = 43.27, p <.001$, $\eta_p^2 = .44$, and outgroup, $F(1, 55) = 37.78, p < .001, \eta_p^2 = .41$, the pattern of individuation was clearer for ingroup members, as shown in Figure 1.4. This effect was not modulated by the experimenter variable, $F(1, 55) < .01, p = .98, \eta_p^2 < 01$.

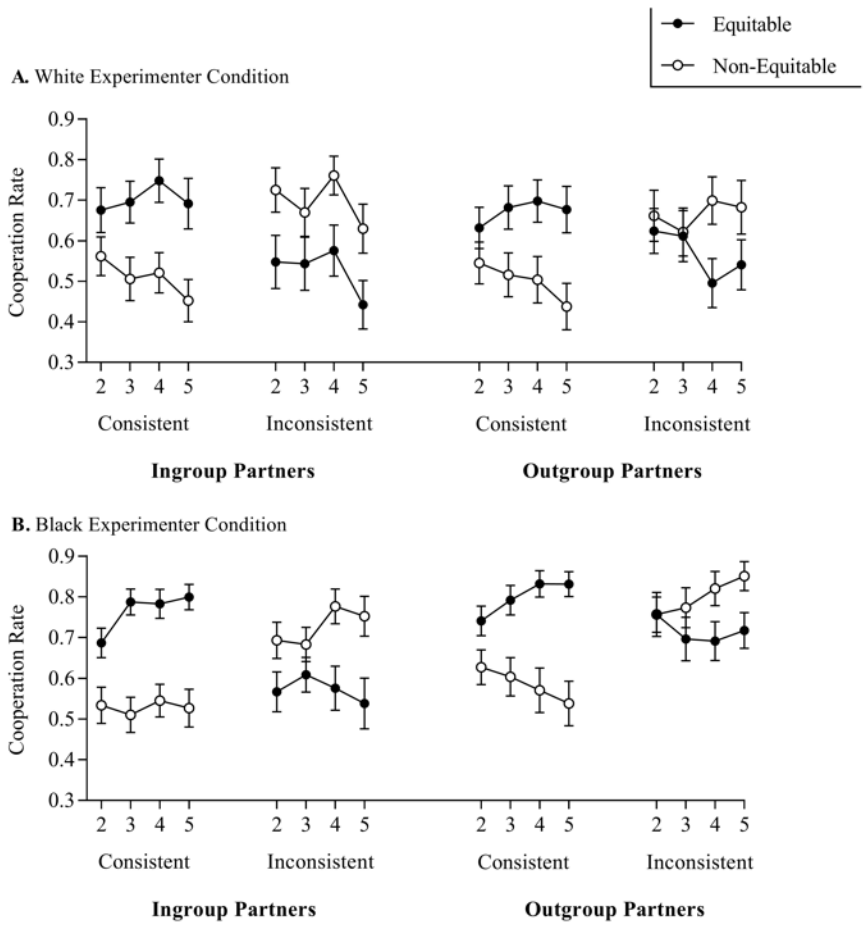

Figure 1.4. Participants' cooperation rates in Experiment 3 as a function of partners' group, Consistency and blocks of trials. (Upper panel (A) is White Experimenter Condition, and lower panel (B) is Black Experimenter Condition).

Impression about partners. The same analyses as in Experiment 2 revealed that black partners were perceived as more similar to each other than white partners. Participants also perceived that black partners were presented more frequently and were more cooperative than white partners, as shown in Table 1.

As in Experiment 2, participants' scores in the scales evaluating trustworthiness and attractiveness were averaged in a unique component evaluating positive impression. A mixed-design ANOVA revealed that participants evaluated the partners according to their individual behavior, as reflected in the Group Behavior x Consistency interaction $F(1, 56) = 28.54$, $p < .001$, $\eta_p^2 = .34$. Partners who were individually equitable were perceived more positively ($M = 3.69$, $SD = 1.43$) than partners who were individually non-equitable ($M = 3.03$, $SD = 1.06$). Moreover, this interaction was modulated by the partner group variable, $F(1, 56) = 7.83$, $p = .007$, $\eta_p^2 = .12$. Despite participants evaluated more positively partners who were individually equitable for both ingroup, $F(1, 56) = 36.71$, $p < .001$, $\eta_p^2 = .40$, and outgroup, $F(1, 56) = 11.58$, $p = .001$, $\eta_p{}^2 = .17$, this pattern of individuation was clearer for the former. We also observed a Experimenter x Consistency interaction, $F(1, 56) = 7.61$, $p = .008$, $\eta_p^2 = .12$, showing that the general more positive impression observed with a black experimenter than with a white experimenter observed in Experiment 2 was replicated here only for inconsistent, $F(1, 56) = 3.87$, $p = .054$, $\eta_p^2 = .07$, but not for consistent partners, $F(1, 56) = .07$, $p = .80$, $\eta_p^2 < .01$.

Discussion

In Experiment 3, we conducted an exact replication of the study reported by Cañadas et al. (2015), while manipulating experimenters' ethnicity as in Experiment 2. In line with the previous studies, we found that participants cooperated more with black than with white partners before learning their particular cooperative behavior. Because of the social norms against racial prejudice (Kawakami et al., 2009; Nosek et al., 2011), it is possible that in the first block, participants were particularly motivated to be egalitarian and actively showed a high readiness to cooperate with black partners. Interestingly, they later showed a tendency to categorize black individuals as they were less efficient at learning about the cooperative behaviors of inconsistent black partners. This pattern is consistent with the evidence suggesting discrepancies between explicit intention to be egalitarian and the implicit negative bias against black people. In the studies of the present research, explicit intention to be egalitarian may be observed in high cooperation rates with black partners while negative implicit biases against black people may be reflected in a subsequent category-based learning about them. The fact that we used an infrequent outgroup in our social context may have increased this tendency. In further research, it would be convenient to control for the frequency of contacts with outgroup members. Otherwise, the inclusion of an implicit measure of prejudice toward black people may help to disambiguate whether there is a negative implicit attitude towards black people subjacent to the initial positive impression of black people.

Further, we replicated the pattern of learning observed in the previous studies of the current research. While participants clearly individuated ingroup members, they rather seemed to categorize outgroup members to some extent. These results are consistent with previous literature arguing that outgroup members are perceived in a more categorical way than ingroup members (e.g., Haslam et al., 1996; Park & Rothbart, 1982; Tajfel & Wilkes, 1963). However, we could not provide an empirical replication of the results reported by Cañadas et al. (2015). Despite using the same experimental procedure and materials as Cañadas and colleagues, and controlling for other variables such as the experimenter's ethnicity or the instructions across the experiments, we did not replicate the pattern of categorization for ingroup and individuation for outgroup members described in their report. The fact that the pattern of data presented in the present research is consistent with the literature and replicated across the three experiments leads us to believe that it is reliable, while the reasons why Cañadas and colleagues found a different pattern of results remain unclear.

To shed light on these discrepancies, we meta-analyzed Experiments 1b, 2 and 3 and the original study reported by Cañadas et al. (2015) using random-effects model and the mean effect size of difference in cooperation rates with equitable vs. non-equitable partners, separately for ingroup and outgroup inconsistent members. The pattern of categorization for ingroup members observed by Cañadas and colleagues was not supported. Indeed, across the four studies, the strategy of individuation was significant for both ingroup, $Z = 1.95$, $p = .050$, and outgroup, $Z = 6.54$, $p < .001$. However, in the 'ingroup' condition, we observed a very high heterogeneity, $I^2 = 88.89\%$, $\tau^2 = .26$, with a significant Cochran's $Q = 23.58$, $p < .001$, which was drastically reduced when removing Cañadas et al.'s (2015) study, $I^2 = 29.07\%$, $\tau^2 < .01$, $Q = 2.68$, $p = .26$. Influential case diagnostic confirmed that Cañadas and colleagues' study had a strong influence on the results reflected in large DFBETAS (DFBETAS = -2.48). New meta-analyses removing Cañadas and colleagues' study confirmed the consistency of the results across the three experiments of the current research with a significant individual learning for both ingroup and outgroup members, larger for ingroup members, as shown in Figure 1.5.

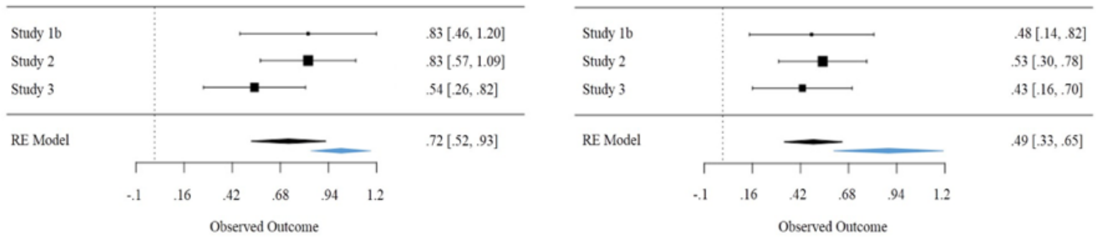

Figure 1.5. Confidence interval of the effect sizes in each study and confidence intervals of the meta-analytic average. Negative values would correspond to category-based learning while positive values correspond to individual-based learning. Blue areas represent the effect size observed with consistent individuals, .97 and .89 respectively for ingroup and outgroup partners. (Left panel (A) is Ingroup Inconsistent Condition, and right panel (B) is Outgroup Inconsistent Condition).

A possible relevant difference between Cañadas and colleagues' study and our Experiment 3 is the sample. Despite we could control that all participants were born and grown in the Spanish context, it is possible that a group of participants (i.e., the sample used by Cañadas and colleagues) was by chance more motivated to individuate outgroup members. Indeed, participants indubitably bring to the lab their real life goals, which may be reflected in their performance in an experimental task if it is related enough. With a sample of undergraduates, several circumstances may have increased participants' will to perceive outgroup members in a more individuated manner such as the area of study that may be more or less related to social concerns, or a particular class about social categorization, prejudice or discrimination. Given the impossibility of predicting the particular motivational context of each one of the participants, future research should include measures of attitudes toward the outgroup, as well as motivation for controlling prejudices, to assess individual differences. In addition, systematic replication is a useful strategy to control for broader contextual variables.

In addition, an interesting result is that in both Experiments 2 and 3, participants perceived that black partners were presented more often, were more cooperative and were more similar to each other. The scarcity of contacts between our participants and black people may account for these results. Finally, in line with our hypothesis and the data from the trust game, we found that partners who were individually equitable were perceived more positively than partners who were individually non-equitable, reflecting an individuated impression about partners. This effect was clearer for white than for black partners, thus showing that the individuation processes observed in the trust game seem to transfer to the subsequent explicit evaluation of partners in the exact same way.

In summary, Experiment 3 replicated the data found in Experiments 1b and 2 indicating that in a trust decision-making, ingroup members are highly individuated whereas the decision to trust outgroup members is also impacted by categorical information. These results appeared to be independent of the experimenter's ethnicity or of the presence of male and female partners within each group.

General Discussion

Across three experiments, we investigated whether participants used individuation or categorization processes to interact with strangers in a trust game paradigm. Our results generally support the hypothesis that when deciding whether or not to trust individuals with whom they had no prior experience, participants quickly identified and used social categories to make their decisions.

The first important result of the current set of experiments comes from the data from the baseline (i.e., first block of the first phase), in which participants made decisions at zero acquaintance. Gender and ethnicity manipulations resulted in different patterns of responses. When gender was manipulated, an ingroup favoritism was found, as participants cooperated more with ingroup than outgroup members (Experiment 1a). However, when ethnicity was manipulated, participants showed an unexpected outgroup favoritism, trusting more outgroup than ingroup members (Experiment 1b, 2 and 3).

The results from the manipulation of gender support the hypothesis that ingroup is generally perceived more positively than outgroup (e.g., Brewer, 2007). However, the outgroup favoritism for black partners was rather unexpected. In this sense, it is interesting that after performing the trust game, participants inaccurately perceived black partners as being presented more often and being more cooperative than white partners. Therefore, it seems that participants' perception of black partners was generally biased, possibly because the population of black people is very scarce in the Spanish context and thus more salient (Fiske, 1980). For this reason, participants may have been particularly motivated to be perceived as equalitarian people (Plant & Devine, 1998), and in consequence cooperated to a great extent with black partners.

A second key result is that across repeated interactions, participants learned about the cooperative behaviors of their partners and their decisions were impacted by both individuation and categorization strategies. Individuation was the general trend with all partners. However, learning about outgroup (men in Experiment 1a and Blacks in Experiments 1b, 2 and 3) was also affected by categorical thinking. Outgroup members were somehow categorized, as participants could not learn whether or not to trust outgroup partners deviating from the group behavior, as accurately as they learned about ingroup members. This result is consistent with the predictions of the outgroup homogeneity effect stating that we have a more categorical perception of outgroup than ingroup members (Tajfel & Wilkes, 1963). The categorization of black individuals was confirmed in participants' post-interaction impression about partners in Experiments 2 and 3. Indeed, black individuals were perceived as more similar to each other than white individuals. Moreover, while white partners who individually equitable were easily identified and consistently evaluated as more attractive than white partners who did not, the evaluation of black partners' attractiveness was inconsistent across the studies and did not seem to depend reliably on their individual cooperative behavior. Together, these data suggested an important effect of intergroup context in trust decision-making: white participants failed to fully individuate black partners (i.e., men and women) when they were presented together with white partners, in the same way that women failed to individuate male partners, when they were presented together with female partner.

There are a number of implications of these results for real life relationships. When categorical judgments affect trust decisions, the inferences made about strangers are necessarily less accurate and the perception of the outgroup is mistaken. In intergroup contexts, such inaccuracies may become particularly important since they prevent people from establishing reliable trust relationships. The inefficiency to learn about outgroup members may result in a more negative perception of the group, opening the path to intergroup conflicts and prejudice.

Limitations and further directions

Despite shedding light on the processes underlying social learning in intergroup contexts, the current research also leaves some questions open. For instance, the reasons why we observed an unexpected outgroup favoritism across Experiments 1b, 2 and 3 are still to clarify. Further research may address this issue by using a more frequent outgroup, measuring participants' prejudice toward black people (see the experimental procedure used by Tortosa, Lupiáñez, et al., 2013), their motivation for controlling prejudice, or manipulating partners' and/or participants' ethnicity. A complementary option is to use a general population sample who is less likely familiarized with social perception concerns compared to the undergraduate samples used in the present studies.

The need for maintaining a clear ingroup-outgroup distinction between partners and participants led us to use a sample of white females in Experiment 1, which limits the generalizability of the results to a broader population. When manipulating ethnicity, this limitation was addressed in Experiment 3 in which men and women participated in the study and showed similar results as in Experiments 1b and 2. However, regarding our manipulation of gender, this issue should be addressed in further research by conducting the same experiment with a male sample of participants.

The present research did not replicate the data reported by Cañadas et al. (2015), although it is consistent with the previous literature on outgroup categorization. It is possible that Cañadas et al. (2015) pattern of results is spurious, or that our samples of participants differed from the sample used by Cañadas et al. (2015) on one or several dimensions such as the level of prejudice toward black people or the motivation for controlling prejudice. Controlling the aforementioned variables related to the sample would shed light on these potential explanations. Importantly, the discrepancies between different experiments in the process of replication actually help to deepen our knowledge of a particular topic. What is often considered as a *failure* to replicate may rather be a step ahead in the understanding of the specific variables and circumstances that affect the results, an issue of the most importance which can hardly be addressed in a single study (see Open Science Collaboration, 2015, for a similar purpose).

Finally, the procedure developed by Cañadas et al. (2013) and extended in this paper is a remarkable contribution in the study of individuation and categorization processes. The inclusion of an inconsistent member within a social group is a key manipulation to understand to what extent the group knowledge may be generalized to all individuals in a stereotyped manner, independently of their particular behavior. It appeared to be a useful tool to investigate these processes in different social contexts (emotion: Cañadas et al., 2016; ethnicity: Cañadas et al., 2015; gender: Cañadas et al., 2013) and related to different dependent variables such as RT or cooperation rates, thus granting its external validity. Further research could reliably use this paradigm to examine the circumstances that lead individuals to prefer one or another strategy.

Conclusions

In conclusion, our results broadly indicated that, in impression formation processes, social categories play a key role in our interactions at zero acquaintance, but also across repeated interactions when we learn about targets. Social categorization may occur even when participants are highly motivated for individuating, and provided with the means to do so. The current results suggest that in cooperation settings, the information that categorizes people with whom we interact might sometimes bias our decisions and hinder our performance.

Acknowledgments and funding

We are grateful to all the students who took part in the studies , to Kerry Kawakami and the Social Cognition lab of York University for ceding part of the photographs that have been used as stimuli, and to Lucía E. Reina and Laura Barea Domínguez for their help in data collection of Experiments 2 and 3. This research is part of MT' thesis dissertation under the supervision of JL.

This work was supported by the Spanish Ministry of Education, Culture and Sports, with pre-doctoral FPU fellowship FPU14/07106 to Maïka Telga, and the Spanish Ministry of Economy and Competitiveness, with research projects PSI2014-52764-P to Juan Lupiáñez and PSI2013-45678-P to Rosa Rodríguez-Bailón and Soledad de Lemus.

Correspondence concerning this article should be addressed to Maïka Telga, Departamento de Psicología Experimental, Universidad de Granada, Campus de Cartuja s/n, 18 071 Granada, Spain. Email: mtelga@ugr.es.

EXPERIMENTAL SERIES 2

The content of this experimental series is currently under review as:

Telga, M., & Lupiáñez, J. (2019). *Use of heuristics in social learning by younger and older adults.* [Under review]

USE OF HEURISTICS IN SOCIAL LEARNING BY YOUNGER AND OLDER ADULTS

Abstract

We explored older and younger adults' strategies to make prediction about unfamiliar individuals at zero acquaintance and across repeated interactions. In an adaptation of the trust game, participants had to predict the cooperative tendencies of their partners to earn economic rewards. We expected older participants to use group membership information to a greater extent than younger participants. While a large body of research suggests that aging is associated with a greater use of heuristics in different domains of cognition, our data indicate that older adults do not rely on heuristics in interpersonal settings, performing as well as their younger counterparts.

Keywords: aging, social categorization, heuristic strategies, learning.

Aging is associated with cognitive losses related to memory, speed processing, reasoning or executive functions (Reuter-Lorenz & Park, 2014; Salthouse, 2010). These functions are involved in many daily activities, including social interactions (von Hippel, 2007). In fact, cognitive impairments affect several social abilities such as behavioral restraint, emotion recognition or perspective-taking (Kalokerinos, von Hippel, & Henry, 2017; Moran, Jolly, & Mitchell, 2012). The impact of aging on social responses is not as simple as a lineal decrease in performance. Instead, a large body of research suggests a differentiated impact of age-related deficits on controlled versus automatic processes. While automatic processes are rather well preserved across life-span, controlled and inhibitory processes are more deficient in older than younger adults (von Hippel & Henry, 2012). Notably, research has demonstrated that older adults' impairments in inhibitory processes lead them to rely more on the preserved automatic processes, resulting in a larger use of heuristic strategies in contexts related to learning (Price & Murray, 2012), language (Kim, Goldstein, Hasher, & Zacks, 2005), financial decision-making (Chen & Sun, 2003) or memory (Fine, Shing, & Naveh-Benjamin, 2018).

In intergroup settings, heuristic strategies may also guide our perception of others. Impressions formation models (e.g., Fiske & Neuberg, 1990) argued that in first encounters, information related to social categories such as gender, ethnicity or age, are automatically and effortlessly processed and integrated to make sense of people. Such information allows to make inferences about them in a categorical way, according to our knowledge about their social groups, but leaves out individual attributes, opening the path to over-categorization, stereotyping and prejudices (Allport, 1954). The default tendency to categorize people can be overcome by inhibiting category-based information in favor of individual characteristics. However, individuating processes are more demanding (Gilbert & Hixon, 1991; Macrae, Milne, & Bodenhausen, 1994) and people only adopt an individuation strategy when they are highly motivated and cognitively able to individuate (Fiske & Neuberg, 1990). In this sense, category-based judgments are social heuristics that people rely on when they lack the motivation or the capacity to engage in individuation processes. In this domain, research has shown that older adults are less able to inhibit the automatic activation of stereotypes than younger adults, resulting in more stereotypic decisions and prejudiced judgments in the former (Gonsalkorale, Sherman, & Klauer, 2009; Radvansky, Copeland, & von Hippel, 2010).

Although stereotyping and prejudices have been extensively investigated, different consequences of category-based social heuristics are also well documented. For instance, the ingroup bias is a more positive attitude towards the members of one's group compared to members of a different social group (Tajfel & Turner, 1979), promoting more pro-social behaviors within the ingroup, even at zero acquaintance. For instance, Telga, de Lemus, Cañadas, Rodríguez-Bailón, and Lupiáñez (2018) conducted an adaptation of the multi-round trust game in which participants had to learn the cooperative tendencies of unfamiliar game partners across repeated interactions. Data from the first interactions showed that female participants spontaneously cooperated more with gender ingroup partners. This tendency observed in the first interactions later disappeared to give way to more individual-based decisions. However, across repeated interactions, heuristic strategies kept impacting learning in a subtler way. While ingroup partners were perfectly individuated, learning about deviant outgroup members was poorer, affected by categorical thinking.

The present research examines the use of heuristic versus more controlled strategies in social learning in different age groups. Specifically, we aimed at testing the hypothesis that older adults use more heuristic processing than younger adults at zero acquaintance, and also across repeated interactions. As the use of categorical heuristic information is related to the cognitive demands of individuating (Gilbert & Hixon, 1991; Macrae et al., 1994), we went beyond previous studies testing trust decisions in a one-shot paradigm (Bailey et al., 2015) and created a context in which participants had to keep in mind the cooperative tendencies of 8 different partners across time. To achieve this goal, we had older and younger participants playing an adaptation of the multi-round trust game (Telga et al., 2018) with both younger and older partners. First, in the baseline, participants were presented with 4 older and 4 younger partners reciprocating in half of the trials. This phase allowed to determine whether participants' decision to cooperate with their partners was based on social heuristics at zero acquaintance.

Second, during a learning phase, we manipulated partners' behavior such that each age group was associated with a specific cooperation tendency. For instance, older partners were equitable, cooperating with participants in most trials, and younger partners were non-equitable, being uncooperative in most trials. This manipulation allowed to examine the use of category-based information in learning. Further, within each social group, one out of four partners were inconsistent with respect to the category tendency. Following the same example, one old partner was non-equitable and one young partner was equitable. By introducing an inconsistent individual within each group, we could explore whether participants' learning is guided by social heuristics, or if in contrast, they used more demanding individuation processes. If participants used categorical information, they should cooperate indistinctively with consistent and inconsistent individuals, according to the group cooperation tendency. Alternatively, if participants individuated their partners, they should show opposite patterns of cooperation with consistent and inconsistent partners cooperating with each of them according to their individual behavior.

Finally, in a transfer phase, we presented participants with 4 older and 4 younger new partners, with whom they had no prior experience. All partners in the transfer phase cooperated in half of the trials, allowing to explore whether the associations between age groups and cooperative behaviors established in the learning phase would be used to categorize new partners accordingly. Specifically, we wanted to test whether learning that a specific group is equitable while the other group is non-equitable would lead participants to cooperate more with new individuals of the former, and less with those of the latter.

In line with previous research using the trust game (Telga et al., 2018), we expected all participants to use learning strategies based on social heuristics at zero acquaintance, when there is no clue predictive of partners' behaviors. Therefore, both older and younger adults were expected to cooperate more with older than younger partners, according to age stereotypes associating older adults with more trustworthiness than younger adults (Fiske, Cuddy, Glick, & Xu, 2002; Schniter & Shields, 2014). Further, across repeated interactions, we expected younger adults to use an individuation strategy to learn about ingroup partners, as in previous studies (Telga et al., 2018). Moreover, we aimed at exploring three hypotheses regarding the learning strategies of older participants and the learning about outgroup partners: whether a) both older and younger participants show a perfect individual-based learning strategy about ingroup partners, and a rather categorical learning strategy about outgroup partners (Telga et al., 2018), or b) older participants rely more on categorical/heuristic information than younger participants (von Hippel & Henry, 2012), showing a category-based learning for both ingroup and outgroup members, or c) older adults try to adopt a role model for young, "teaching" to the younger the benefits of reciprocal exchanges (Bailey & Leon, 2019; Charness & Villeval, 2009), overall being more cooperative than younger participants.

Finally, in the transfer phase, on the basis of previous studies (Telga et al., 2018) we expected younger adults not to use the information from the learning phase to predict new partners' behaviors. That is, we expected them to cooperate in similar ways with younger and older new partners, no matter how much this group learned in the learning phase. As for older participants, the strategy of cooperation was expected to depend on the learning from the previous phase. No differences in cooperation between older and younger partners were expected in the case participants' learning had been affected by categorical information only for outgroup members. In contrast, if participants had used categorical information for both ingroup and outgroup members in the learning phase, they were expected to categorize new partners according to the associations established during the learning phase. The hypotheses and procedure of this experiment were pre-registered on Open Science Framework (https://osf.io/kqc5v).

Method

Participants. According to the pre-registered criterion, 41 older (mean age: 65.36, range: 60-86) and 41 younger (mean age: 21.78, range: 18-27) volunteers participated in the study for monetary compensation. Sensitivity analyses assuming an alpha criterion of 0.05 and a power criterion of 0.80 revealed that with this sample, the smallest effect size that could have been detected for the critical Participant Age x Group Behavior x Consistency interaction was $f = .18$. All participants reported normal or corrected to normal vision. Written informed consent was obtained from all participants. The study was part of a larger project approved by the local university ethical committee (175/CEIH/2017).

Stimuli and Materials. Computers equipped with a 24-inch monitor and E-Studio 2.0 software (Schneider, Eschman, & Zuccolotto, 2002) were used for stimuli presentation and data acquisition. The game partners were represented by 16 photographs taken from the UT Dallas Face Database (Minear & Park, 2004). All faces were presented against a grey background with a neutral emotional expression. Older partners were between 61 and 68 years while younger partners were between 20 and 25 years. Within each group of age, half of the partners were men.

Procedure. Before performing the trust game, older participants were individually administered the Mini-Mental State Examination (MMSE, Folstein, Robins, & Helzer, 1983). All participants were led to individual cubicles and provided with visual and verbal instructions about the trust game. Participants were informed that they would have to decide whether or not to cooperate with unfamiliar people in several rounds. Each round started with '€1' displayed during 200 ms indicating that participants virtually received

€1. After a 200-ms fixation point, the photograph of the partner (8.10º x 10.95) of this round appeared during 1000 ms and participants were asked to make their decision by pressing the '1' key to cooperate or the '0' to keep the euro and therefore, not to cooperate. Keeping the €1 led to the end of the round. Alternatively, if participants cooperated, their partners would receive €5 and in turn decide whether or not to reciprocate by sending back €2.50 or keeping the €5 to themselves, leaving the participant with nothing. After participants made their decision of after 1500 ms, a 500-ms fixation cross was displayed followed by the partner's decision and the final outcomes in a single screen during 1500 ms. A 1000 ms inter-trial black screen ended each round.

As shown in Figure 2.1, the first block of the experiment was the baseline, in which participants played with 4 older and 4 younger partners (2 men and 2 women in each age group). Each partner was presented 8 times, cooperating half of them, resulting in 64 trials. Next, in the learning phase, each age group was associated with a particular cooperative tendency, either equitable or non-equitable. In the equitable group, 3 partners cooperated on 75% of the trials while in the non-equitable group, 3 partners cooperated only on 25% of the trials. In each age group, the fourth partner was inconsistent with respect to the rest of their group. That is, in the equitable (non-equitable) group, one partner cooperated on 25% (75%) of the trials. The picture corresponding to the inconsistent partner within each group was counterbalanced across participants. If participants categorized their partners, they should adopt the same cooperative behavior with consistent and inconsistent partners, according to the group behavior. This should be reflected in a main effect of the group behavior variable, not qualified by the consistency. In contrast, if participants individuated their partners, they should show opposite patterns of cooperation for consistent and inconsistent partners, cooperating with each partner according to their individual behavior. This should be reflected in a Group Behavior x Consistency interaction, in which participants cooperate more with consistent partners belonging to the equitable group and inconsistent partners belonging to the non-equitable group, and less with consistent partners belonging to non-equitable and inconsistent partners belonging to the equitable group. Each partner was presented 32 times, resulting in 256 trials divided in 4 blocks of 64 trials.

Next, participants played the transfer phase with 4 older and 4 younger new partners, half women in each age group. Each partner was presented 8 times, cooperating on half of the trials, resulting in 64 trials. This phase allowed to examine whether participants categorize the new partners using the information learnt in the learning phase.

If participants categorize, they should cooperate more with partners belonging to the group that was equitable in the leaning phase, and less with the group that was non-equitable in the learning phase. Alternatively, if they individuate the new partners, they should equally cooperate with all of them, independently of their age. Finally, in Blocks 7 to 10, participants performed a final learning phase in which we counterbalanced the associations established between age groups and group behaviors. For instance, if older partners were cooperative in Block 2 to 5, they were non-equitable in Blocks 7 to 10. This allowed to have a full within-participants design regarding the variables of interest. Whether younger or older partners were equitable in Blocks 2 to 5 was counterbalanced across participants.

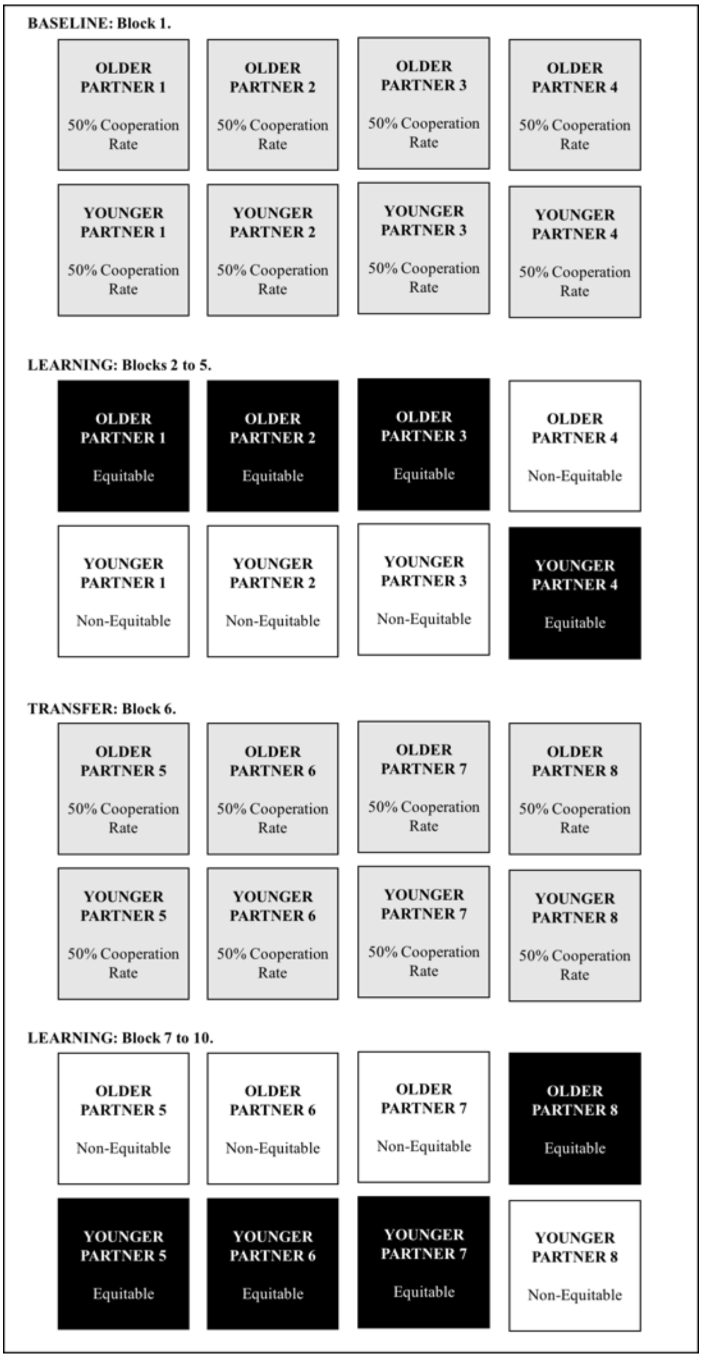

Figure 2.1. Example of the procedure employed for the trust game. Consistent partners are represented in white and inconsistent partners are represented in black. In the baseline and in the transfer phases, partners are presented for the first time and their cooperative behavior is not manipulated.

Results

All older participants met the inclusion criterion, obtaining more than 27 points in the MMSE. However, two participants were excluded from the analyses, for cooperating indiscriminately in all the trials and for abandoning the study, respectively, leaving 39 participants in the older group and 41 participants in the younger group. Trials with RT faster than 200 ms were also filtered out.

Baseline. Cooperation rates in the baseline were subjected to a mixed-design ANOVA with partner age (older vs. younger adults) as a within-participants variable and participant age (older vs. younger adults) as a between-participants factor. The main effect of participant age was not significant, $F(1, 78) = 2.00$, $p = .162$, $\eta_p^2 = .03$. Neither was the main effect of partner age $F(1, 78) = 1.30$, $p = .259$, $\eta_p^2 = .02$, nor the Partner Age x Participant Age interaction, $F(1, 78) = .71$, $p = .402$, $\eta_p^2 = .01$.

Learning. The data from Blocks 2 to 5 and Blocks 7 to 10 were collapsed such that in these analyses, the blocks are numbered from 2 to 5 from the first presentations of the partners. A mixed-design ANOVA was performed on cooperation rates in the learning phase with partner age (older vs. younger adults), group behavior (equitable vs. non-equitable), consistency (consistent vs. inconsistent) and block (2-5) as within-participants variables, and participant age as a between-participants factor. This analysis revealed a Group Behavior x Consistency x Block interaction, $F(3, 234) = 34.88$, $p < .001$, $\eta_p^2 = .31$. In the consistent partners condition, the Group Behavior x Block was significant, $F(3, 234) = 42.42$, $p < .001$, $\eta_p^2 = .35$, indicating that participants linearly increased their cooperation with equitable partners across blocks, $F(1, 78) = 27.31$, $p < .001$, $\eta_p^2 = .26$. The quadratic component was also significant, $F(1, 78) = 6.32$, $p = .014$, $\eta_p^2 = .08$, indicating that this increase reached an asymptote as shown in Figure 2.2. In contrast, participants cooperated less with non-equitable partners with cooperation linearly decreasing across blocks, $F(1, 78) = 48.46$, $p < .001$, $\eta_p^2 = .38$, until reaching an asymptote, quadratic component, $F(1, 78) = 9.43$, $p = .003$, $\eta_p^2 = .11$. In the inconsistent partners condition, the Group Behavior x Block interaction was also significant, $F(3, 234) = 13.15$, $p < .001$, $\eta_p^2 = .14$. However, in this case, the interaction indicated that participants tended to cooperate more with inconsistent partners belonging to a non-equitable group, although the linear increase in cooperation was only marginal, $F(1, 78) = 3.31$, $p = .073$, $\eta_p^2 = .04$. In contrast, they cooperated less, and linearly decreased their cooperation across blocks with inconsistent partners belonging to an equitable group, $F(1,78) = 25.20$, $p < .001$, $\eta_p^2 = .24$. Here, the quadratic component was also significant, $F(1, 78) = 8.15$, $p = .006$, $\eta_p^2 = .10$, also suggesting that the decrease in cooperation stabilized across blocks as shown in Figure 2.2. The main effect of participants' age was not significant, $F(1, 78) < .01$, $p = .993$, $\eta_p^2 < .01$, and did not modulate any other effects, larger $F(1, 78) = 2.06$, $p = .156$, $\eta_p^2 < .03$ for the Partner x Participant age interaction.

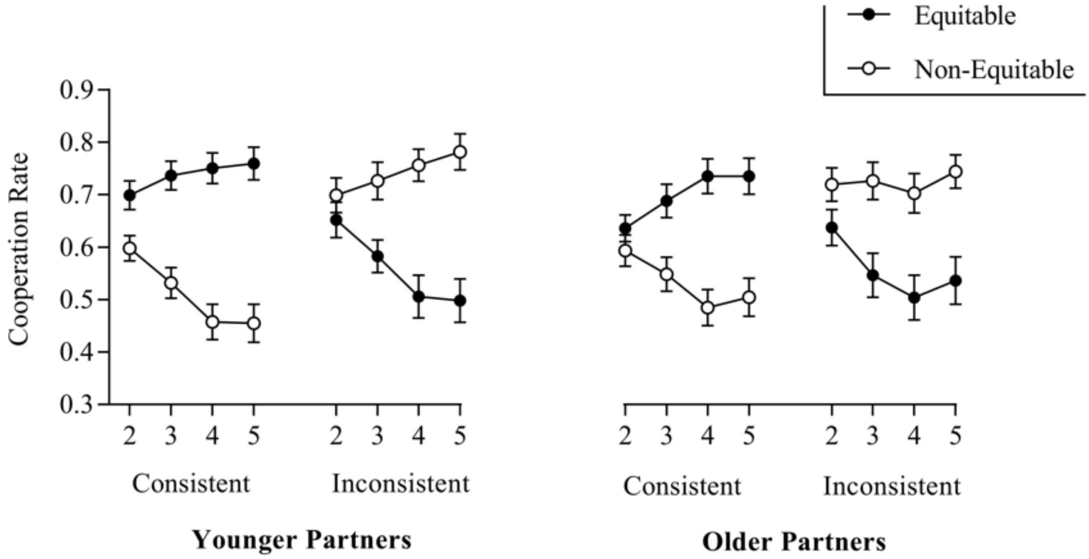

Figure 2.2. Cooperation rate in the learning phase as function of partners' consistency (consistent vs. inconsistent) and group behavior (equitable vs. non-equitable) for younger and older participants. Error bars represent the error standard of the mean.

Transfer. A mixed-design ANOVA was conducted on cooperation rates in the transfer phase with partner age (older vs. younger adults) as a within-participants variable and equitable group (older vs. younger adults) and participant age (older vs. younger adults) as between-participants factors. The Partner Age x Equitable Group interaction was marginal, $F(1, 76) = 3.15$, $p = .08$, $\eta^2_p = .04$, showing some tendency to cooperate more with new individuals of the equitable group. However, and importantly, the critical three-way interaction was not significant, $F(1, 76) = 1.21$, $p = .275$, $\eta^2_p = .02$, showing once more no differences between the two age groups. The main effect of participants' age was not significant either, $F(1, 78) = 1.58$, $p = .213$, $\eta^2_p = .02$.

Discussion

The present research aimed at examining whether older adults use more heuristic strategies than younger adults to predict others' behaviors, at zero acquaintance and across repeated interactions. To achieve this goal, we used an adaptation of the trust game in which younger and older participants had to predict the cooperative behaviors of younger and older unfamiliar partners.

Results from the baseline suggested that at zero acquaintance, older and younger participants did not differ in their cooperation decisions and do not seem to use heuristics based on age stereotypes to predict their partners' behaviors. Although this finding contrasts with results from previous research evaluating the use of gender and ethnic stereotypes in impression formation (Telga et al., 2018), it also replicates previous studies indicating that despite older adults are perceived as more trustworthy than younger adults, this does not necessarily translate at the behavioral level into higher cooperation rates with older partners (Bailey et al., 2015).

We also aimed at analyzing older adults' learning strategies in a more demanding context, in which they had to learn the cooperative behavior of 8 unfamiliar partners. We expected to replicate previous findings indicating that younger adults relied on categorical (gender and ethnicity) knowledge when learning about outgroup members. Moreover, given that older adults are more prone to use heuristic strategies than younger adults in social contexts (von Hippel & Henry, 2012), we expected older participants' to rely on categorical information at least when learning about outgroup members. However, the results from the learning phase suggested that both older and younger adults adopted an individuation strategy and accurately learned the individual behaviors of consistent and inconsistent ingroup and outgroup partners, suggesting that participants do not rely on age categories to the same extent as they rely on ethnicity or gender categorization. This interpretation seems interesting given the large body of research supporting the relevance of age in social categorization in other paradigms (Fiske & Neuberg, 1990; Martin & Macrae, 2007; Stolier & Freeman, 2016).

The results from the transfer phase again suggested that older and younger participants behave similarly and do not seem to use previous knowledge acquired during the leaning phase to categorize new partners. Although this result is no surprise regarding young participants, it keeps challenging the theorizing that older adults rely on heuristic strategies to make social decisions.

We did not find evidence that older participants trust more unfamiliar partners than younger participants, as suggested by previous studies (Bailey & Leon, 2019). The expectations to interact several times with the same partners may have biased participants' spontaneous responses.

Overall, the data from this study consistently suggest that older adults do not always rely on heuristics to a greater extent than younger adults. It is interesting to note that the use of an economic game and the fact that participants were rewarded with real economic outcomes may have enhance participants' motivation to respond accurately, promoting the engagement of the cognitive resources necessary to overcome social biases and individuate their partners (Kawakami et al., 2014; Neuberg & Fiske, 1987; Telga et al., 2018). This interpretation rises an interesting question regarding whether the use of heuristics in older adults in social settings may be overcome with the appropriate motivation.

Acknowledgments and funding

We are thankful to all the volunteers who took part in the study, and to Inmaculada Núñez Jiménez and Andrea Diezma Fernández for their help in data collection. We also would like to thank José Rodríguez Montero for his invaluable help in the recruitment of older participants.

This work was supported by the Spanish Ministry of Education, Culture and Sports, with pre-doctoral FPU fellowship FPU14/07106 to MT, and the Spanish Ministry of Economy and Competitiveness, with research projects PSI2014-52764-P and PSI2017-84926-P to JL. This research is part of MT thesis dissertation under the supervision of JL.

Correspondence concerning this article should be addressed to Maïka Telga, Departamento de Psicología Experimental, Universidad de Granada, Campus de Cartuja s/n, 18 071 Granada, Spain. Email: mtelga@ugr.es

CHAPTER 4. Social Power

EXPERIMENTAL SERIES 3

The content of this experimental series is currently submitted as:

Telga, M., & Lupiáñez, J. (2019). *Differences between men and women in social perception: An effect of social power?* [Submitted].

Different Strategies of Social Perception Between Men and Women: An Effect of Social Power?

Abstract

The current research explores the use of social categorization and individuation strategies by male and female participants to make inferences about male and female strangers. Across two experiments, we used an adaptation of the multi-round trust game in which participants had to predict the cooperative behaviors of their partners to earn monetary outcomes in three phases. In a baseline, partners were cooperative in half of the trials, independently of their gender. In a learning phase, we provided participants with reliable category-based information (by associating gender groups with opposite patterns of cooperation), and relevant individual-based information (by including 25% of inconsistent individuals within each gender group; i.e., individuals displaying a pattern of cooperation opposite to the group behavior). Finally, in a transfer phase, participants interacted with new men and women with whom they had no prior experience, to evaluate the possible use of the categorical associations established in the learning phase to predict the behaviors of unfamiliar people. Meta-analyses including data from the current and previous studies showed that male and female participants used different strategies to inform their judgments. Women showed an ingroup favoritism, cooperating more with female than with male partners in the baseline, but later attempted to individuate their partners in the learning and the transfer phases. In contrast, men engaged in individuation processes in the baseline and the learning phase, but relied on gender categories to predict their partners' behaviors in the transfer phase. We discuss these results in light of power differences between men and women.

Keywords: social categorization, individuation, gender, learning, power

Discussing politic news with an acquaintance, attending a work meeting for the first time, or trying to convince a journal editor of the relevance of our manuscript, are different social situations in which we involve with positive expectations. Although some situations may seem more challenging than others, we take part into social interactions at zero acquaintance on a daily basis, hopefully with satisfactory results. For these social interactions to be harmonious, we need to integrate numerous cues to form accurate impressions of others and to behave accordingly. A large body of research suggests that in these contexts, the cognitive tool allowing us to organize all this information efficiently is social categorization (Macrae & Bodenhausen, 2000; Quinn & Macrae, 2005).

Social categorization consists of attending the characteristics diagnostic of people's group membership to make inferences about them (Brewer, 1988; Fiske & Neuberg, 1990). Facial features related to gender, ethnicity or age are readily extracted and integrated to categorize others into their corresponding social groups (Hugenberg, Young, Bernstein, & Sacco, 2010; Stolier & Freeman, 2016). Once a person has been categorized into a familiar social category, perceivers may use their category-based knowledge or beliefs to make inferences about this particular individual (Kawakami, Amodio, & Hugenberg, 2017; Macrae & Bodenhausen, 2000).

Alternatively, perceivers may focus on the targets' unique characteristics diagnostic of their identities (Brewer, 1988; Hugenberg et al., 2010). This individuation strategy is cognitively demanding as requires a piecemeal integration of a variety of individual attributes. Therefore, it is only used when perceivers are provided with both cognitive ability and sufficient motivation to individuate (Fiske & Neuberg, 1990; Gilbert & Hixon, 1991). In fact, social categorization is the default strategy in impression formation processes, allowing to make sense of others effortlessly (Cloutier, Mason, & Macrae, 2005; Macrae & Bodenhausen, 2001; Quinn & Macrae, 2005). However, category-based judgments do not consider targets' individual and unique characteristics and, given the heterogeneity of social groups, may lead to flawed decisions.

Research in social psychology has extensively investigated the potential negative consequences of social categorization, from over-categorization to stereotypes and prejudice (Allport, 1954; Kawakami et al., 2017). Interestingly, the perception of social groups is also impacted by self-perception. When perceivers organize social information in terms of social categories (e.g., men vs. women), they easily notice the groups they belong to (i.e., ingroups) and the groups to which they do not (i.e., outgroups) (Ellemers & Haslam, 2012). This process of self-categorization affects different domains of social cognition and perception. Thus, to enhance their group identity, perceivers might engage in strategies favoring ingroup over outgroup members (Tajfel, 1978). For instance, they may develop a more positive attitude toward ingroup than outgroup members, known as ingroup favoritism (Tajfel & Turner, 1979; Turner, Brown, & Tajfel, 1979). The distinction between ingroup and outgroup members has also important implications in traits attribution as reflected in the outgroup homogeneity effect, that is, the perception of more similarity among outgroup members than among ingroup members (Freeman, Schiller, Rule, & Ambady, 2010; Tajfel & Wilkes, 1963).

Recently, Telga, de Lemus, Cañadas, Rodríguez-Bailón, and Lupiáñez (2018) reported a study documenting both ingroup favoritism and outgroup homogeneity effect in a single experiment. In this study, a sample of women played an adaptation of the trust game in which they had to predict the cooperative behavior of unfamiliar male and female partners. To earn real economic outcomes, participants had to cooperate with equitable partners and to not cooperate with non-equitable partners. This procedure provides participants with a highly motivating context with economic rewards in which they were mostly expected to individuate their partners (Neuberg & Fiske, 1987). Whether participants' decisions were impacted by category- or individual-based processes was explored in three phases: baseline, learning and transfer.

In the baseline phase, partners' behaviors were independent of their gender. However, an ingroup favoritism was observed in that female participants cooperated more with ingroup (i.e., female) than outgroup (i.e., male) partners. That is, participants categorized their partners into gender groups and used these gender categories to make their cooperation decisions, despite its lack of relevance. In the learning phase, gender was relevant to perform the task as gender groups were associated with opposite behaviors. For instance, men were equitable and women were non-equitable. Moreover, one out of four partners within each gender group displayed an individual behavior opposite to the group behavior. Following the same example, one man was non-equitable, and one woman was equitable. Participants accurately learned the individual behaviors of ingroup (i.e., female) partners and made their cooperation decisions accordingly. However, some outgroup homogeneity effect was observed. Specifically, outgroup members were perceived to behave more similarly, resulting in a poorer learning of the individual behavior of outgroup (i.e., male) partners who did not behave consistently with their group membership. Finally, in a transfer phase, participants were presented with new male and female partners with whom they had no prior experience. The aim of this transfer phase was to test whether participants would use the learning acquired during the learning phase to categorize new partners accordingly. For instance, a categorization strategy would be observed if participants cooperated more with male than female partners after learning that men were the equitable group. However, the pattern of data did not support this hypothesis.

The results of this experiment provided evidence for both the ingroup favoritism and the outgroup homogeneity effect, reflecting the consequences of categorical thinking, even in an outcome-dependency context in which participants should be motivated to individuate (Neuberg & Fiske, 1987). However, the selection of only female participants limits the generalizability of the results, especially relevant considering differences in status and power between men and women. In fact, power is defined as the control of people’s valued resources and outcomes (Dépret & Fiske, 1993; Fiske, 1993) and often correlates with status, that is, a person’s position on an unequal stratification or hierarchy (Fiske, 2010). Because of salient gender inequalities granting men higher social status than women in a large range of domains such as economic outcomes, role attributions or chance of survival (Fiske, 2010; Sen, 2003; Shen, 2013), it is likely that men, as a social group, perceive themselves as more powerful than women. Importantly, research has shown that people holding power use categorization processes to a greater extent than powerless individuals (Fiske & Neuberg, 1990; Guinote & Phillips, 2010; Rodríguez-Bailón, Moya, & Yzerbyt, 2000), paying more attention to information that confirms their expectations (Goodwin, Gubin, Fiske, & Yzerbyt, 2000). To the extent that men and women differ in their self-perception on power, they may use different social learning strategies, men relying on categorical information to a greater extent than women.

Exploring male participants’ cooperation decisions may also help to understand the processes underlying spontaneous cooperation decisions with male and female partners at zero acquaintance. In fact, Telga et al. (2018) observed that women cooperated more with female partners when their gender did not predict their behavior, and interpreted this result as an evidence for ingroup favoritism. However, an alternative explanation is that participants relied on gender stereotypes associating women with more communal behaviors (Cuddy et al., 2009; Cuddy, Fiske, & Glick, 2008), and higher trustworthiness (Dong, Liu, Jia, Li, & Li, 2018). Hence, male participants’ decisions may shed light on this matter. If men spontaneously cooperate more with male than female partners in the present experiment, ingroup favoritism would likely drive cooperation decisions at zero acquaintance. Conversely, if men also cooperate more with female than with male partners, the impact of gender stereotypes may at least partially explain the result reported by Telga et al. (2018).

The aim of the present research is to address the aforementioned limitations by comparing male and female participants' social learning strategies in a trust game. To achieve this goal, we first replicated Telga and colleagues' (2018) study with a sample of men (Experiment 1). Further, we experimentally compared men and women's use of categorization vs. individuation strategies with a between-participants manipulation by selection of participants' gender (Experiment 2). In this second study, we also increased the number of partners with whom participants played the trust game from 8 to 16 to rule out a possible ceiling effect when individuating only 8 partners.

Experiment 1

In Experiment 1, we explored male participants' cooperation decisions with unknown partners in a trust game. In the baseline, when partners' gender did not predict their cooperation tendency, we expected participants to use gender stereotypes associating women with more trustworthiness than men (Buchan, Croson, & Solnick, 2008; Dong et al., 2018), and to cooperate more with women than with men. In the learning phase, we predicted an outgroup homogeneity effect, that is, we expected participants' learning to be impaired with inconsistent outgroup (i.e., female) members. Finally, on the basis of previous studies (Telga, Cantiani, & Lupiáñez, 2019; Telga et al., 2018), we did not expect participants to use the knowledge acquired in the learning phase to categorize new partners in the transfer phase.

Method

Participants. Prior to analyses, estimation of the appropriate sample size was calculated with G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007) on the basis of the effect size reported by Telga et al., (2018) for the main effect of partner gender in the baseline $\eta_p^2 = .11$, revealing that a sample of 67 participants would be sufficient to replicate the higher cooperation with women than with men, with an alpha error of .05 and a power of .80. Sixty-eight male volunteers (mean age: 21.88, range: 17-31) participated in the study in exchange for a financial compensation proportional to their performance (€6.22 on average). Written informed consents were obtained from all the participants. The study was part of a larger project approved by the local university ethical committee (175/CEIH/2017).

Stimuli and Materials. Computers equipped with a 24-inch monitor and E-Prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2002) were used for stimuli presentation and data acquisition. The same sixteen pictures (i.e., 8 men and 8 women) with neutral emotional expressions used by Telga et al. (2018) were taken from the NimStim Set of Facial Expressions (Tottenham et al., 2009) to represent the game partners.

Procedure. The study consisted of a single session in which participants performed an adaptation of the multi-round trust game (Telga et al., 2018). Each round started with '€1' displayed during 200 ms indicating that participants virtually received €1. Next, a fixation point appeared during 500 ms, followed by the photograph of the partner (4.30º x 3.34º) of this round during 1500 ms. Participants had to indicate whether or not they decided to cooperate with this partner by pressing the '1' key to cooperate or the '0' key not to cooperate. Not cooperating and therefore keeping the €1 would conclude the round. Alternatively, if participants cooperated, their partner would receive €5 and in turn decide whether or not to reciprocate by either sending back €2.50 or keeping the €5 to themselves, leaving the participant with €0. After participants made their decision or after 1500 ms, a fixation cross was displayed for 500 ms followed by the partner's decision and the final outcomes in a single screen during 1500 ms. A 1000 ms inter-trial black screen ended each trial, as shown in Figure 3.1.

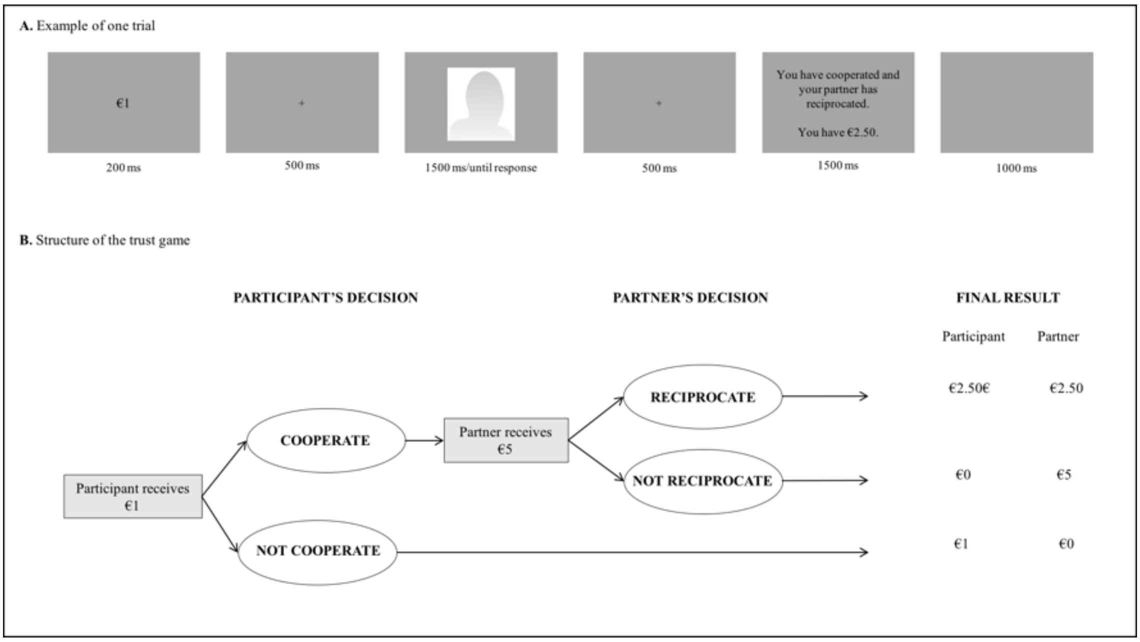

Figure 3.1. Example of one trial (A) reflecting the structure of one round of the trust game (B).

Participants performed the trust game in three different phases as shown in Figure 3.2. In the baseline, they played with 8 partners (4 men and 4 women) cooperating in half of the trials and not cooperating in the other half. Each partner was presented 8 times resulting in one block of 64 trials. This phase allowed to examine whether participants would spontaneously cooperate more with one of the gender groups.

Next, in the learning phase, the two gender groups were associated with opposite patterns of cooperation, being either equitable or non-equitable. For instance, 3 men were equitable and cooperated on 75% of the trials, while 3 women were non-equitable and cooperated only on 25% of the trials. Moreover, within each gender group, the fourth partner displayed a pattern of cooperation inconsistent with their group membership. Following the previous example, one inconsistent man was non-equitable and cooperated only on 25% of the trials, and one inconsistent woman was equitable and cooperated on 75% of the trials. With this manipulation, we aimed at analyzing whether participants used individual- or category-based information to make predictions about their partners. If they use a categorization strategy, they should make their cooperation decisions according to their partners' group membership and independently of their individual consistency. This strategy would be reflected in a main effect of the group behavior variable, not modulated by the consistency. Alternatively, if participants use an individuation strategy, they should notice that inconsistent individuals behave differently from the group, and consequently, they should display opposite patterns of cooperation with consistent and inconsistent individuals within the same gender group. This strategy would be reflected in a Group Behavior x Consistency interaction. All partners were presented 32 times, resulting in 256 trials divided in 4 blocks of 64 trials. The photographs associated with inconsistent partners were counterbalanced across participants.

Next, in a transfer phase, participants played with 8 new partners (4 men and 4 women) again cooperating in half of the trials and not cooperating in the other half. This phase allowed to verify whether participants used their learning from the previous phase to categorize new individuals. If participants categorized their partners, they should cooperate more with the gender group that was equitable in the learning phase, and less with the gender group that was non-equitable in the learning phase. Alternatively, if they individuated their new partners, they should cooperate with them independently of their gender. Each partner was presented 8 times resulting in one block of 64 trials.

Finally, after the transfer, participants performed a new learning phase in which we counterbalanced the associations established between gender group and group behavior. That is, if men were equitable in Blocks 2 to 5, new men would be non-equitable in Block 7 to 10. This procedure allowed to have a full within-participants design regarding the variables of interest. The order in which gender groups were associated with either equitable or non-equitable behaviors was counterbalanced across participants.

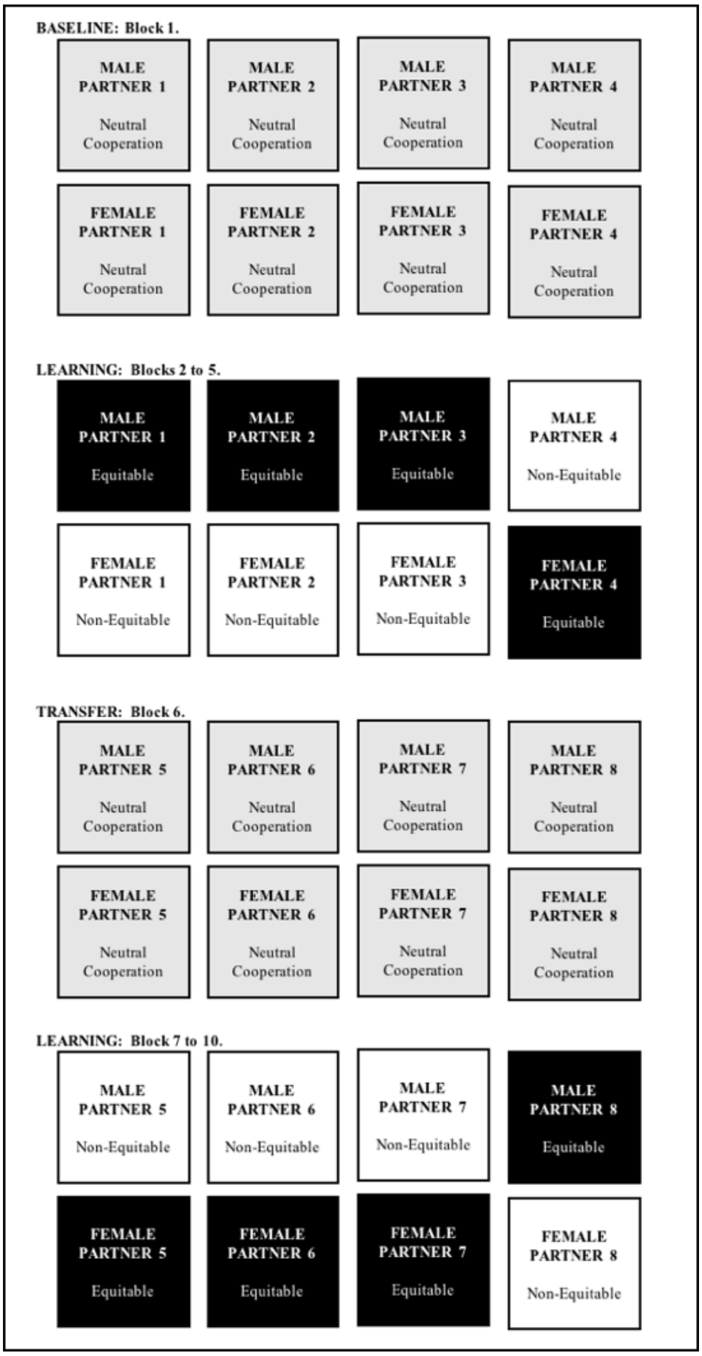

Figure 3.2. Group and individual partners' behaviors manipulated across the baseline, learning, and transfer. Equitable partners (i.e., partners cooperating on 75% of the trials) are represented in black and non-equitable partners (i.e., partners cooperating on 25% of the trials) are represented in white. Partners are represented in grey when their cooperation rate is at 50%.

Results

One participant was excluded from the analyses for having extreme cooperation rates (i.e., cooperating in more than 80% of the trials), leaving in 67 participants in the analyses. Trials with RTs shorter than 200 ms were also filtered out, as in Tortosa, Lupiáñez, and Ruz (2013).

Baseline. To examine whether participants would cooperate more with one of the gender groups at zero acquaintance, cooperation rates in the baseline were subjected to a repeated-measures ANOVA with partner gender as a within-participants variable. We found than male participants did not differ significantly in their cooperation rate between women (*M* = .67, *SD* = .18) and men (*M* = .63, *SD* = .18), $F(1, 66) = 2.85$, $p = .096$, $\eta_p^2 = .04$, thus neither the ingroup bias nor the gender stereotype hypothesis were supported.

Learning. Results from the two learning phases (Blocks 2 to 5 and Blocks 7 to 10) were collapsed and the blocks are numbered in the following analyses from the first (Block 2) to the last (Block 5) presentation of the game partners. To verify participants learning strategies across repeated interactions, a repeated-measures ANOVA with partner gender (man vs. woman), group behavior (equitable vs. non-equitable), consistency (consistent vs. inconsistent) and block (2-5) as within-participants variables was conducted on cooperation rates in the learning phase. We found a significant Group Behavior x Consistency x Block interaction, $F(3, 198) = 63.66$, $p < .001$, $\eta_p^2 = .49$, revealing that participants accurately displayed opposite patterns of cooperation for consistent and inconsistent partners. In fact, for consistent partners, the Group Behavior x Block interaction was significant, $F(3, 198) = 54.67$, $p < .001$, $\eta_p^2 = .45$. Participants linearly increased their cooperation with equitable partners from Block 2 (*M* = .73, *SD* = .15) to Block 5 (*M* = .83, *SD* = .15), $F(1, 66) = 36.27$ $p < .001$, $\eta_p^2 = .36$. In contrast, they linearly decreased their cooperation with non-equitable partners from Block 2 (*M* = .55, *SD* = .15) to Block 5 (*M* = .35, *SD* = .20), $F(1, 66) = 71.32$, $p < .001$, $\eta_p^2 = .37$. This decrease was also qualified by a significant quadratic component, $F(1, 66) = 4.08$, $p = .047$, $\eta_p^2 = .06$, indicating that it reached an asymptote as shown in Figure 3.3.

For inconsistent partners, the Group Behavior x Block interaction was also significant, $F(3, 198) = 27.36$, $p < .001$, $\eta_p^2 = .29$. Here, participants linearly increased their cooperation with inconsistent partners from the non-equitable group from Block 2 (*M* = .72, *SD* = .21) to Block 5 (*M* = .84, *SD* = .18), $F(1, 66) = 26.40$, $p < .001$, $\eta_p^2 = .29$. Conversely, they linearly decreased their cooperation with inconsistent partners from the equitable group from Block 2 (M = .54, SD = .19) to Block 5 (M = .37, SD = .23), $F(1, 66) = 38.04$, $p < .001$, $\eta_p^2 = .37$. Importantly, the Partner x Group Behavior x Consistency x Block interaction was not significant, $F(3, 198) = 1.39$, $p = .248$, $\eta_p^2 = .02$, indicating that participants learned to the same extent about ingroup and outgroup partners, as can be appreciated in Figure 3.3.

Moreover, to have a sense of participants' learning about the cooperative behaviors of their partners, we computed a learning index by subtracting cooperation rate with partners who individually cooperated from cooperation rate? with partners who individually cooperated. With this index, participants' scores with inconsistent partners are informative of their learning strategy. Negative values would reflect a categorization strategy and positive values would reflect an individuation strategy. This learning index was subjected to a repeated-measures ANOVA with consistency (consistent vs. inconsistent) as a within-participants variable. We observed that participants learned to the same extent about consistent (M = .35, SD = .16) and inconsistent (M = .34, SD = .23) partners, $F(1, 66) = .23$, $p = .633$, $\eta_p^2 < .01$, thus showing a perfect pattern of individuation, as shown in Figure 3.3.

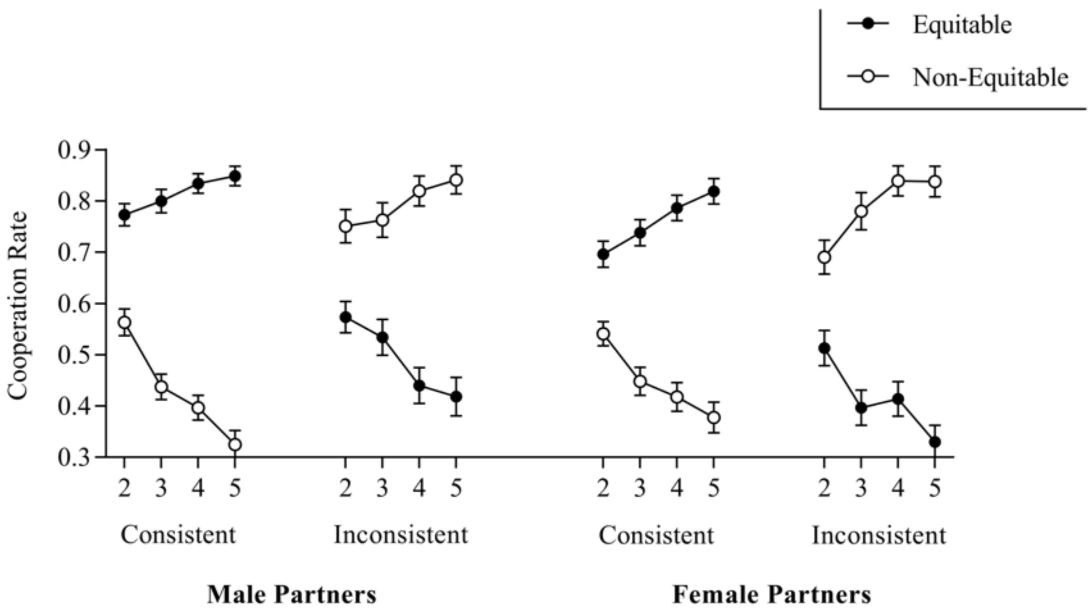

Figure 3.3. Cooperation rates in the learning phase of Experiment 1 with male and female partners according to their group behavior and individual consistency. Error bars represent the standard error of the mean.

Transfer. Cooperation rates in the transfer phase were subjected to a repeated-measures ANOVA with partner group (equitable vs. non-equitable) as a within-participants variable. We found that participants cooperated more with new partners belonging to the group that was equitable in the learning phase (M = .72, SD = .15) than with new partners from the group that was not equitable in the learning phase (M = .68, SD = .16), although this difference was only marginal, $F(1, 66) = 3.79$, $p = .056$, $\eta_p^2 = .05$.

Discussion

The aim of Experiment 1 was to explore male participants' use of category- vs. individual-based information to predict the behaviors of unknown male and female game partners. We examined men's cooperation decisions in an adaptation of the trust game across three phases: a baseline in which gender information was irrelevant, a learning phase in which participants could learn their partners' behavior according to their gender or their individual attributes, and a transfer phase in which participants played with new men and women behaving independently of their gender.

In the baseline, we predicted that participants would cooperate more with female than with male partners, according to gender stereotypes associating women with more communal behaviors than men (Cuddy et al., 2009). Neither the gender stereotype nor the ingroup favoritism hypothesis was supported by the data. However, because these data directly concern our hypothesis about the use of gender stereotype at zero acquaintance and do not allow to draw clear conclusions, we further explored them in Experiment 2.

In the learning phase, we expected male partners to rely on categorical information to predict their partners' behaviors, and therefore to show an outgroup homogeneity effect reflected in a poorer learning about inconsistent female partners. However, participants showed a pattern of perfect individuation in learning, for both male and female partners. Given the diversity of the effects of power on social perception (Fiske, 2010), it is possible that male partners did perceive themselves as more powerful than women, but the perception of power impacted social perception differently than expected. In fact, as well as the greater use of category-consistent cues (Rodríguez-Bailón et al., 2000), some studies have also shown that power facilitates prioritization and goal-consistent behaviors, and increases flexibility and inhibition of irrelevant information (Guinote, 2007a, 2007b). Therefore, it is possible that in the learning phase, in which participants could learn accurately the behavior of each one of their partners, power helps male participants to adopt a strategy allowing them to achieve the proposed goal, that is, to focus on their partners' individual attributes. In contrast, in the baseline and the transfer phases in which participants were playing with their partners for the first times, they tended to rely on categorical information.

In fact, in the transfer phase, participants seemed to use the associations established between gender and group behavior in the learning phase to categorize new individuals with whom they had no prior experience. Interestingly, in previous research, no hint of transfer effect has been found in the trust game paradigm with such a small number of partners, neither with female participants, nor manipulating other dimensions of social categorization such as ethnicity or age (Telga et al., 2018; Telga & Lupiáñez, 2019). Therefore, it is possible that the expected tendency to use categorical information was not found in the learning phase because participants adopted a goal-driven strategy in this context, but did appear in the transfer phase when they were playing at zero acquaintance. Nonetheless, the transfer effect was only marginal in Experiment 1, therefore we aimed at replicating it in Experiment 2.

Altogether, these data suggested that male participants do not use categorical information to predict others behaviors across repeated interactions, but may rely on categorical processing at zero acquaintance, as observed in the baseline and the learning phases. However, Experiment 1 did not allow to directly compare male and female participants' decisions in the trust game, as only men participated in this study, and the comparison had to be made with a different study. This issue is addressed in Experiment 2 by manipulating (by selection) participants' gender in a quasi-experimental design.

Finally, given that the reliance on categorization processes is associated with the cognitive cost of individuating (Gilbert & Hixon, 1991; Sherman, Macrae, & Bodenhausen, 2000), the fact that participants were able to perfectly individuate 8 partners raises a question regarding whether individuation was cognitively challenging with the procedure used in Experiment 1. In fact, in line with Fiske and Neuberg's (1990) model of impression formation, if individuation is easy and participants are motivated to individuate by financial rewards, therefore the present paradigm may not be sensitive enough to explore the effect of power on social perception. We address this potential issue in Experiment 2 by increasing the number of partners with whom participants play from 8 to 16.

Experiment 2

The aim of Experiment 2 was to directly compare men and women learning strategies in a between-participants manipulation by selection of participants' gender. Moreover, we aimed at improving the sensitivity of our measures by making individuation more difficult. Therefore, we increased the number of partners with whom participants played the trust game from 8 to 16. In the baseline, we expected both male and female partners to cooperate more with women than with men. In the learning phase, data from Experiment 1 would suggest an outgroup-homogeneity effect only for female participants. In the transfer phase, in line with the results from Experiment 1, we predicted that only male participants would categorize new partners according to the associations established in the learning phase.

Method

Participants. Before data collection, we decided to use the same sample size as Telga et al. (2018) and to run 40 participants in each experimental group. To avoid losing participants as in Experiment 1, we also decided to replace participants with extreme cooperation rates (less than 20% or more than 80%) before performing the analyses of interest. Eighty undergraduates (40 men, mean age: 20.09, range: 18 - 26) took part in the in experiment in exchange for financial compensation proportional to their performance (€5.68 on average). Written informed consents were obtained from all the participants. The study was part of a larger project approved by the local university ethical committee (175/CEIH/2017).

Stimuli, Materials and Procedure. Stimuli, materials and procedure were similar to the ones used in Experiment 1, except that participants were presented with 16 (8 men and 8 women) instead of 8 partners in each learning phase. We maintained the same proportion of inconsistent partners within each gender group (i.e., 25%) such that in the present experiment, 6 partners were consistent and displayed the group cooperative behavior, while 2 partners were inconsistent and displayed a pattern of cooperation opposite to the group behavior. The photographs used to represent the game partners were taken from the Karolinska Directed Emotional Faces database (Lundqvist, Flykt, & Öhman, 1998). As in the previous experiment, all partners were presented against a grey background with a neutral emotional expression.

Results

Six participants were replaced because of extreme cooperation rates, (i.e., cooperating in less than 20% or in more than 80% of the trials), and trials faster than 200 ms were filtered out.

Baseline. To analyze participants' spontaneous tendency of cooperation with men and women at zero acquaintance, cooperation rates in the baseline block were subjected to a 2(Partner gender: men vs. women) x 2(Participant gender: men vs. women) mixed-design ANOVA with partner gender as a within-participants variable and participant gender as a between-participants factor. The expected main effect of partner gender was not significant, $F(1, 78) = .30$, $p = .584$, $\eta_p^2 < .01$, indicating that participants did not significantly differ in their cooperation between male ($M = .65$, $SD = .13$) and female ($M = .64$, $SD = .16$) partners. Neither the main effect of participant gender, $F(1, 78) = .29$, $p = .590$, $\eta_p^2 < .01$, nor the Partner Gender x Participant Gender interaction were significant, $F(1, 78) = 1.86$, $p = .176$, $\eta_p^2 = .02$.

Learning. To analyze whether learning about the cooperative tendencies on unknown male and female partners was affected by individual- or category-based information, cooperation rates in the learning phase were subjected to a mixed-design ANOVA with partner gender (men vs. women), group behavior (equitable vs. non-equitable), consistency (consistent vs. inconsistent), and block (2-5) as within-participants variables, and participant gender (men vs. women) as a between-participants factor. The Group Behavior x Consistency x Block interaction was significant, $F(3, 234) = 42.71$, $p < .001$, $\eta_p^2 = .35$, indicating a mainly individuating learning strategy. This interaction was not qualified by participants' gender, $F(3, 234) = .12$, $p = .946$, $\eta_p^2 < .01$, suggesting that learning was similar for male and female participants. Neither was it qualified by partner gender, $F(3, 234) = .39$, $p = .761$, $\eta_p^2 < .01$, indicating that learning about male and female partners were similar. Importantly, the expected five-way interaction was not significant either, $F(3, 234) = .69$, $p = .560$, $\eta_p^2 < .01$. Note that the Bayes factor for the five-way interaction suggested no evidence for the difference in learning between inconsistent ingroup and outgroup members, with $BF_{10} = .06$.

For consistent partners, the Group Behavior x Block interaction was significant, $F(3, 234) = 50.29$, $p < .001$, $\eta_p^2 = .39$, indicating that participants linearly increased their cooperation from Block 2 ($M = .66$, $SD = .13$) to Block 5 ($M = .74$, $SD = .15$) with equitable partners, $F(1, 78) = 28.09$, $p < 001$, $\eta_p^2 = .27$, until reaching an asymptote, quadratic effect, $F(1, 78) = 5.61$, $p = .020$, $\eta_p^2 = .07$. In contrast, they linearly decreased their cooperation with non-equitable partners from Block 2 ($M = .56$, $SD = .14$) to Block 5 ($M = .41$, $SD = .22$), $F(1, 78) = 55.61$, $p < .001$, $\eta_p^2 = .42$, until reaching an asymptote, quadratic effect, $F(1, 78) = 9.00$, $p = .004$, $\eta_p^2 = .10$. For inconsistent partners, the Group Behavior x Block interaction was also significant, $F(3, 234) = 13.49$, $p < .001$, $\eta_p^2 = .15$. Participants linearly increased their cooperation from Block 2 ($M = .61$, $SD = .18$) to Block 5 ($M = .65$, $SD = .19$) with inconsistent partners belonging to the non-equitable group, $F(1, 78) = 6.42$, $p = .013$, $\eta_p^2 = .08$. In contrast, they linearly decreased their cooperation from Block 2 ($M = .60$, $SD = .20$) to Block 5 ($M = .47$, $SD = .26$) with inconsistent partners belonging to the equitable group, $F(1, 78) = 25.35$, $p < .001$, $\eta_p^2 = .25$, as shown in Figure 3.4.

Overall, as expected, the individuation strategy was not as perfectly achieved as in Experiment 1. In fact, in contrast to Experiment 1, participants learned more about consistent ($M = .24$, $SD = .19$) than inconsistent ($M = .09$, $SD = .23$) partners, $F(1, 79) = 27.19$, $p < .001$, $\eta_p^2 = .06$. However, and importantly, this decrease in learning for inconsistent individuals was similar for ingroup and outgroup members, indicating that although the individuating processes were reduced when playing with 16 partners, this reduction was independent of the partners' group membership.

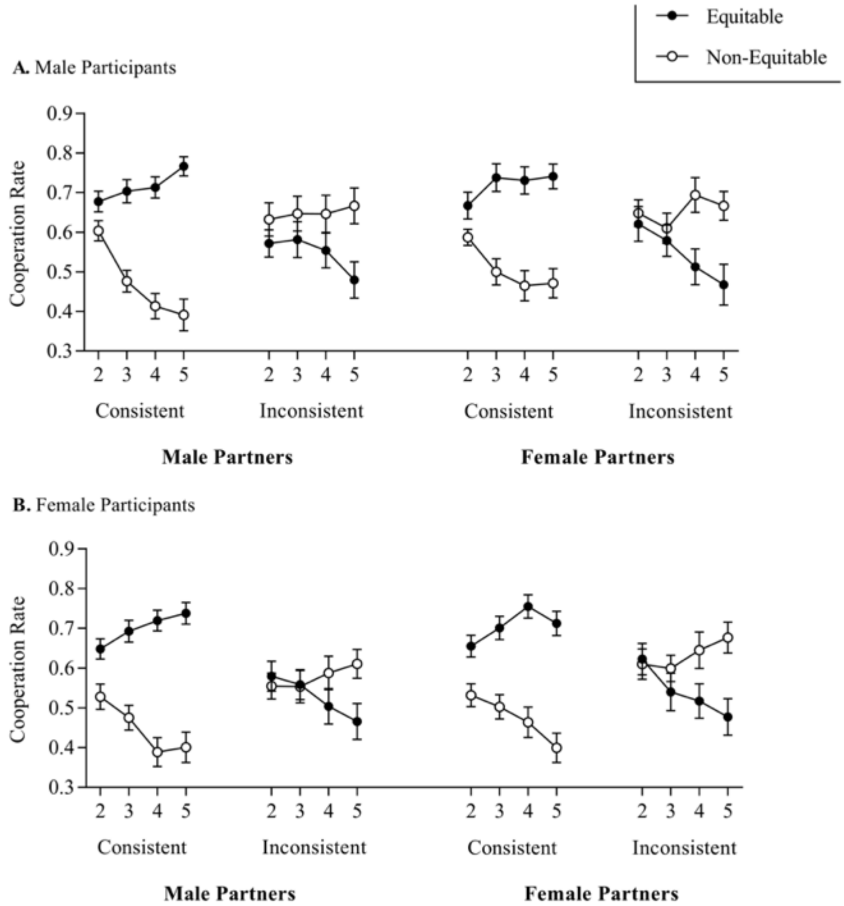

Figure 3.4. Cooperation rates in the learning phase of Experiment 2 for A) male participants and B) female participants. Cooperation rates are displayed according to partner gender, group behavior and individual consistency. Error bars represent the standard error of the mean.

Transfer. To verify whether participants used their knowledge from the learning phase to make cooperation decisions with new male and female partners, participants' cooperation rates in the transfer phase were subjected to a 2(Partner group: equitable vs. non-equitable) x 2(Participant gender: men vs. women) mixed-design ANOVA. The Partner Group x Participant Gender was significant, $F(1, 39) = 4.15$, $p = .045$, $\eta_p^2 = .15$. Male participants cooperated more with new partners belonging to the group that was equitable in the learning phase ($M = .74$, $SD = .16$) than with those belonging to the group that was not ($M = .63$, $SD = .20$), $F(1, 78) = 12.79$, $p = .001$, $\eta_p^2 = .25$. In contrast, female participants did not significantly differ in their cooperation between partners belonging to the group that was equitable ($M = .67$, $SD = .15$) and not equitable ($M = .64$, $SD = .14$) in the learning phase, $F(1, 39) = 1.49$, $p = .229$, $\eta_p^2 = .04$, as shown in Figure 3.5.

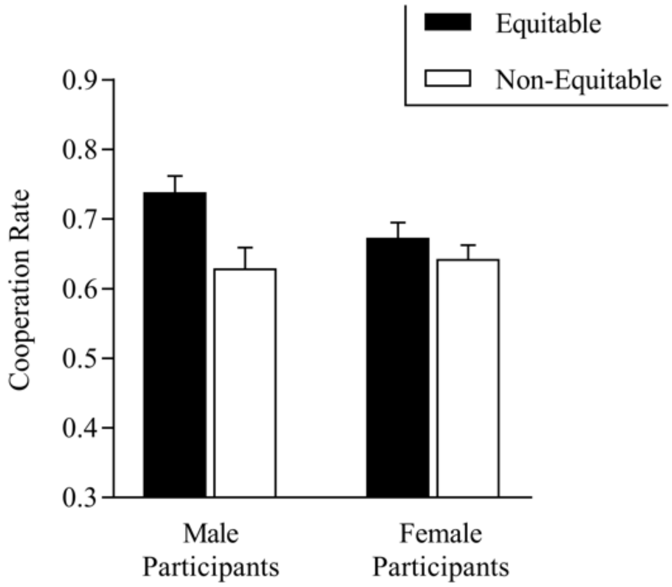

Figure 3.5. Cooperation rates in the transfer phase in Experiment 2 with partners belonging to the gender group associated to equitable vs. non-equitable behavior in the learning phase, for male and female partners. Error bars represent the standard error of the mean.

Discussion

The aim of Experiment 2 was to directly compare male and female social strategies to predict unknown partners' behavior at zero acquaintance and across repeated interactions.

In the baseline, we found that neither male nor female participants differed in their cooperation between male and female partners. These data contrast with previous studies showing a tendency to cooperate more with women than with men, in Experiment 1 for male participants, and in Telga et al. (2018) for female participants. To better understand male participants' inferences strategies at zero acquaintance, we meta-analyzed the two studies of the present manuscript in the section below.

In the learning phase, we observed that both male and female participants used an individuation strategy to predict their partners' behaviors, and learned the individual behaviors of consistent and inconsistent ingroup and outgroup partners. Regarding male participants, these data replicated the pattern observed in Experiment 1, confirming that men do not rely on categorical information when learning about partners across repeated interactions. However, these data are inconsistent with studies reporting an outgroup homogeneity effect in learning with female participants (Telga et al., 2018). The increase in the number of partners may have made learning more difficult in general, and the pattern of poorer learning observed for inconsistent outgroup members in Telga and colleagues' (2018) report has been extended to all conditions. To better understand these discrepancies, we meta-analyzed in the section below the experiments from the present research together with the experiment reported by Telga et al. (2018)

Importantly, we replicated data from Experiment 1 in the transfer phase. In fact, we found that only male participants used category-based information to predict their new partners' behaviors, cooperating more with new partners from the group that was equitable in the learning phase. This result is consistent with the literature indicating that powerholders are more prone to attend stereotypic information (Fiske & Dépret, 1996; Guinote & Phillips, 2010; Rodríguez-Bailón et al., 2000) than low power people. The fact that a category-based strategy has been observed for men only in the transfer phase is still consistent with the literature on the effect of power on social perception. Making accurate inferences about people is often related to higher outcomes for powerholders. Therefore, under some circumstances, powerful people may be highly motivated to individuate others (Overbeck & Park, 2006; Schmid Mast, Khademi, & Palese, 2020) especially when their predictions are associated with personal rewards (Keltner, Gruenfeld, & Anderson, 2003). As participants were outcome-dependent on making accurate predictions, it is possible that in the first phases of the experiment, male participants' responses were goal driven, leading them to attend their partners' individual attributes. However, in the transfer phase with new faces, when participants could not predict their partners' behaviors, only male participants relied on the categorical information learned in the previous phase of the trust game, while female participants kept attempting to make individual judgments.

Meta-analyses

To better understand the differences between men and women in impression formation and learning strategies, we meta-analyzed several studies (including the experiments from the current research) using the trust game with gender categories, weighting effect sizes by sample sizes with random effects model, separately for male and female participants.

For the **baseline**, we meta-analyzed the data from male participants in the two experiments of the current research, and data from female participants of Experiment 2 together with the data from three more experiments conducted in our lab. These meta-analyses revealed that male participants do not differ in their cooperation between men and women, $Z = .03$, $p = .974$, while female participants show an ingroup favoritism at zero acquaintance, with significantly higher cooperation with female than with male partners, $Z = 3.33$, $p = .001$, as shown in Figure 3.6.

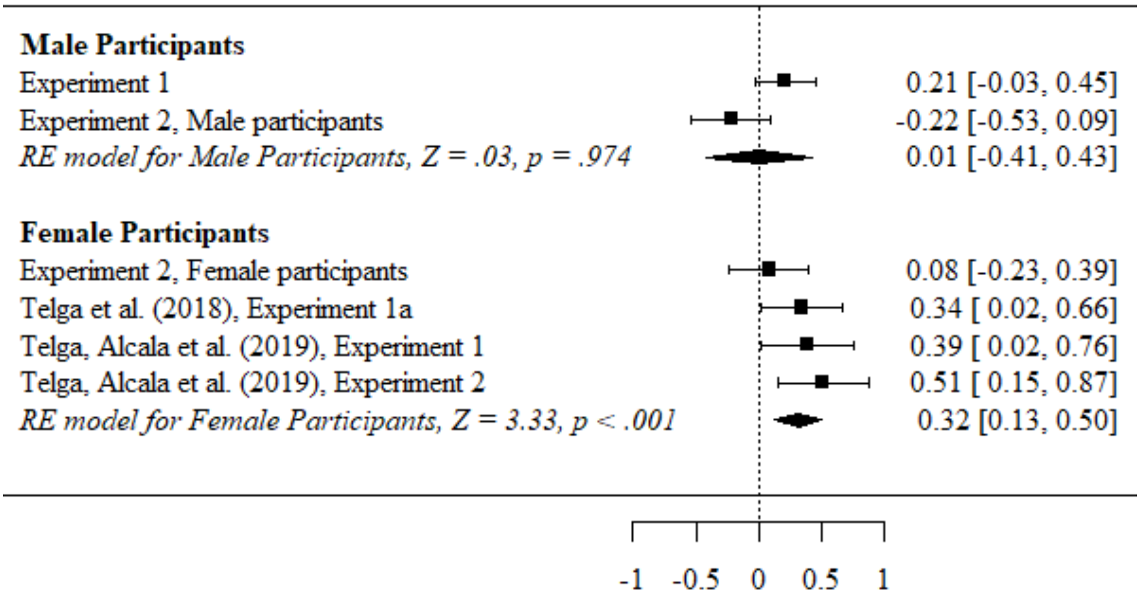

Figure 3.6. Forest plot showing cooperation rates in the baseline with male and female partners. Negative values correspond to higher cooperation with male partners, and positive values correspond to higher cooperation with female partners. For male participants, the meta-analysis includes data from Experiment 1 ($N = 67$) and data from the male participants of Experiment 2 ($N = 40$). For female participants, the meta-analysis includes data from female participants in Experiment 2 ($N = 40$), from Telga et al.'s (2018) Experiment 1a ($N = 40$), and two experiments from a manuscript in preparation (Telga, Alcalá, & Lupiáñez, 2019). In these two experiments, data from a few male participants were filtered out leaving in 31 participants in Experiment 2, and 34 participants in Experiment 3.

For the **learning phase**, we meta-analyzed the data from male participants in the two experiments of the current research, and data from female participants of Experiment 2 of the present work together with the data from Telga et al. (2018) Experiment 1a, as only those experiments employed a full within-subjects design allowing to compare learning about ingroup and outgroup members. We observed that neither men, $Z = -1.48$, $p = .139$, nor women, $Z = 1.51$, $p = .131$, consistently display an outgroup homogeneity effect, as shown in Figure 3.7. Therefore, it is possible that the data reported by Telga et al. (2018) were the results of a type I error and need to be considered with caution in light of the present meta-analysis.

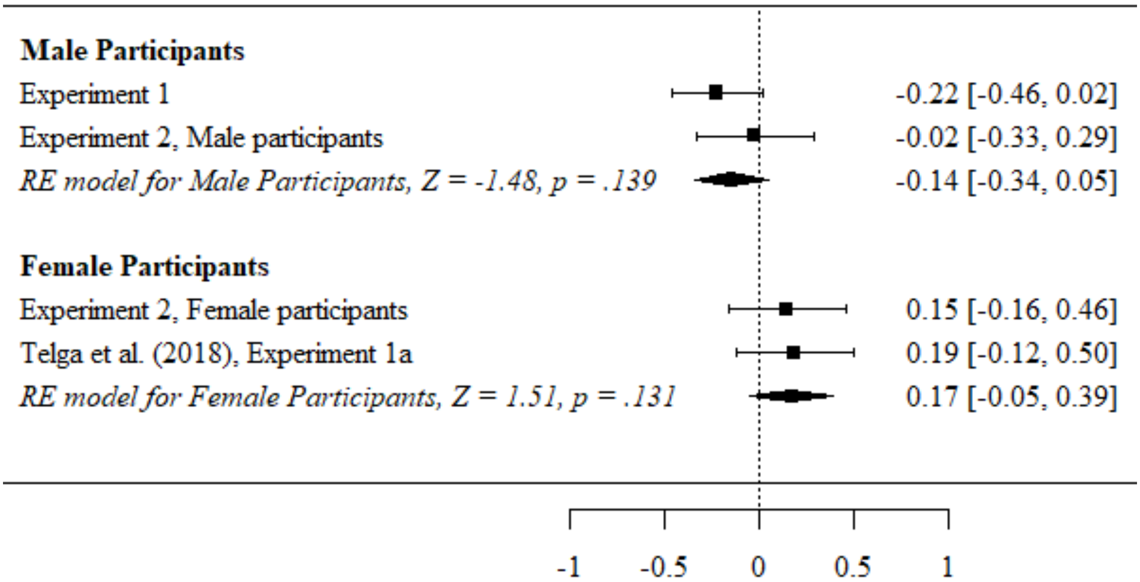

Figure 3.7. Forest plot showing learning about inconsistent ingroup and outgroup partners. Negative values indicate larger learning about outgroup members and positive values indicate larger learning about ingroup members. For male participants, the meta-analysis includes data from Experiment 1 ($N = 67$) and data from the male participants of Experiment 2 ($N = 40$). For female participants, the meta-analysis includes data from female participants in Experiment 2 ($N = 40$) and data from Telga et al.'s (2018) Experiment 1a ($N = 40$).

Finally, for the **transfer phase**, we also meta-analyzed the two experiments of the present research for male participants, and data from women in Experiment 2 of the current work together with four more experiments using gender categories in the trust game for female participants, (Telga, Alcalá, et al., 2019; Telga, Cantiani, et al., 2019; Telga et al., 2018), as shown in Figure 3.8. The meta-analyses confirmed that male participants rely on social categories to make predictions in the transfer phase, $Z = 2.35$, $p = .019$, cooperating significantly more with the group that was equitable during the learning phase, while female participants do not differ in their cooperation with new partners from the equitable and non-equitable groups, $Z = -0.16$, $p = .876$.

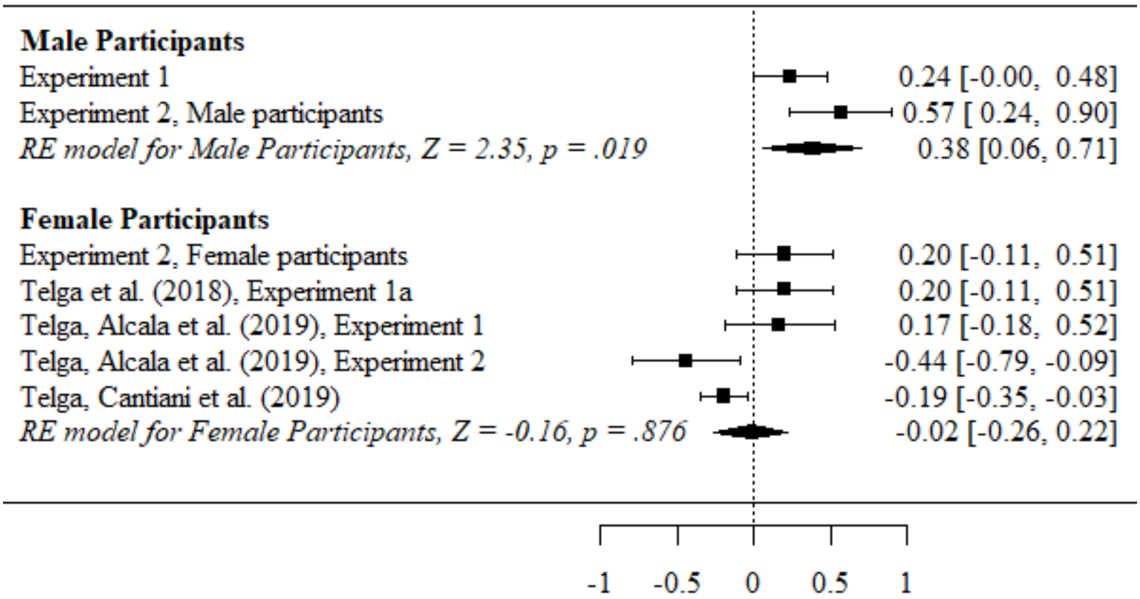

Figure 3.8. Forest plot showing cooperation rates in the transfer with partners from the groups that were equitable vs. non-equitable in the learning phase. Negative values correspond to higher cooperation with non-equitable partners and positive values correspond to higher cooperation with equitable partners. For male participants, the meta-analysis includes data from Experiment 1 (N = 67) and data from the male participants of Experiment 2 (N = 40). For female participants, the meta-analysis includes data from female participants in Experiment 2 (N = 40), from Telga et al.'s (2018) Experiment 1a (N = 40), and three experiments from two manuscripts currently in preparation (Telga, Alcalá, et al., 2019; Telga, Cantiani, et al., 2019). In the two experiments from Telga, Alcalá, et al. (2019), data from a few male participants were filtered out leaving in 31 participants in Experiment 2, and 34 participants in Experiment 3. In contrast, all participants were female in Telga, Cantiani, et al. (2019) such that all 160 participants were included in the meta-analysis.

General discussion

Across two experiments, we explored whether categorical thinking differently impacts social perception of male and female participants. Specifically, in an adaptation of the trust game, we tested whether men (Experiments 1 and 2) and women (Experiment 1) responses were impacted by gender stereotypes, ingroup favoritism or outgroup homogeneity effect, all being well documented consequences of social categorization.

To understand gender differences in cooperation decisions at zero acquaintance, we meta-analyzed the two experiments of the present research for male participants, and several studies using the same paradigms for female participants. The results broadly indicated that women's responses are driven by an ingroup favoritism, while men do not use gender categories to decide whether or not to cooperate with unknown partners. Gender differences related to status may explain these data. In fact, because ingroup favoritism is a strategy to enhance social identity (Brewer, 2001), low-status groups may feel a greater need to adopt such strategy than high-status groups, resulting in larger intergroup biases in the former (Brewer, 1979; Mullen, Brown, & Smith, 1992). The economic settings of the task may have made status differences particularly salient, as women are generally less confident than men in math and economic-related fields (Bordalo, Coffman, Gennaioli, & Shleifer, 2019). Therefore, in this context, self-perception of low status may promote ingroup favoritism to a greater extent for female than for male participants.

A second meta-analysis revealed that male and female participants individuated their partners in the learning phase, independently of their group membership. Despite increasing the number of partners from 8 to 16 decreases learning about inconsistent individuals, this decrease impacted to the same extent learning about inconsistent ingroup and outgroup partners. Importantly, these analyses suggest that the data reported by Telga et al. (2018) regarding outgroup homogeneity in gender are not reliable. Overall, it seems that the outgroup homogeneity effect during learning is observed with ethnic categories (Telga et al., 2018), but not with gender, as shown in the present research, or age (Telga & Lupiáñez, 2019). Despite the salience of gender, ethnicity and age categories in social perception (Stolier & Freeman, 2016), age and gender are different from ethnic categories in that age- and gender-related cross-categories interactions occur on a daily basis, even within the family nucleus, but are far less likely with regard to ethnicity. This is especially true in the homogenous ethnic context in which the present research has been conducted. Therefore, it is possible that greater familiarity with age and gender outgroups favors more individual learning, while it is not the case with ethnic outgroups (see for instance Tanaka, Kiefer, & Bukach, 2004).

Finally, data from the transfer phase support the hypothesis that powerholders rely more on category-related information than powerless people (Fiske & Neuberg, 1990; Guinote & Phillips, 2010). In fact, we observed that only male participants used the associations established between gender categories and cooperative behaviors to make inferences about men and women with whom they had no prior experience. It is interesting to note that across repeated interactions (i.e., in the learning phase), male participants did engage in individuation processes to make accurate inferences about their partners, likely because repeated interactions provide a suitable context for learning and achieving monetary outcomes. Given that power is associated with greater attention to rewards (Keltner et al., 2003), the goal of making accurate inferences to earn economic outcomes may have allowed male participants to overcome category-based judgements in the learning phase, promoting instead individuated accurate responses (Guinote, 2007a, 2007b). However, in spite of their non-categorical perception of men and women faces during the learning phase, men (but not women) categorized new male and female partners in the transfer phase.

In spite of the congruency of the results observed with previous literature investigating the effect of power on social perception, future studies are needed to replicate and extend the current research. For instance, despite of interracial inequalities granting higher status to white people (Fiske, 2010), Telga et al. (2018) did not report a greater reliance on ethnic category information in the transfer phase, but did observe an outgroup homogeneity effect in the learning phase. In the same vein, age-related status differences grants higher status to young versus old adults (Cuddy, Norton, & Fiske, 2005; Fiske, 2010). However, Telga and Lupiáñez (2019) showed that older and younger adults do not differ in their learning strategies about older and younger partners. Therefore, it is possible that different status contrasts result in different consequences for powerholders (Fiske, 2010), and differentially impact social perception. This question could be explored in further investigation by examining different dimensions of social hierarchies, or experimentally manipulating participants' power.

Overall, the current research suggests that, in intergroup settings, male and female participants used different strategies to predict others' behaviors. Women first show an ingroup favoritism, which is later overcome with individuation processes. In contrast, men first adopt an individuation strategy, until they disengage and rely on categorical information. Interestingly, such categorical judgments occur even in contexts in which participants are cognitively able to individuate, and aware of the greater accuracy of individual predictions, suggesting that motivational factors withdraw attention from individual to categorical attributes.

Acknowledgments and funding

We are tankful to all the volunteers who took part in the experiments, to the research group HUM-642 from the University of Jaén for their help in the recruitment of participants, and to Andrea Diezma Fernández for her help in data collection.

This work was supported by the Spanish Ministry of Education, Culture and Sports, with pre-doctoral FPU fellowship FPU14/07106 to MT, and the Spanish Ministry of Economy and Competitiveness, with research projects PSI2014-52764-P and PSI2017-84926-P to JL. This research is part of MT's thesis dissertation under the supervision of JL.

Correspondence concerning this article should be addressed to Maïka Telga, Departamento de Psicología Experimental, Universidad de Granada, Campus de Cartuja s/n, 18 071 Granada, Spain. Email: mtelga@ugr.es

CHAPTER 5. Cognitive resources

EXPERIMENTAL SERIES 4

The content of this experimental series is in preparation as:

Telga, M., Cantiani, A., & Lupiáñez, J. (2019). *Individuation at any cost: Motivation to individuate prevents social categorization in cognitively demanding contexts.* [Manuscript in preparation].

INDIVIDUATION AT ANY COST: MOTIVATION TO INDIVIDUATE PREVENTS SOCIAL CATEGORIZATION IN COGNITIVELY DEMANDING CONTEXTS

Abstract

Social categorization is the cognitive tool allowing us to make sense of the social world effortlessly, by attending targets' attributes diagnostic of their group membership instead of their individual characteristics. Because of its resources-saving function, social categorization has been argued to be the default strategy in impression formation processes. In the present research, we adapted the trust game to explore how the cognitive cost of individuation impacts the differential use of categorization or individuation strategies. In a modified trust game, four experimental groups had to make predictions about the cooperative behaviors of unfamiliar partners to earn economic rewards. Such cooperative behaviors were manipulated at the group and at the individual level, allowing to examine whether participants individuated or categorized their partners. Moreover, across the four experimental groups, we increased the number of partners from 8 to 16, 32 and finally 64 individuals. We predicted that individuation would be favored over categorization as participants were outcome-dependent on their predictions, but when individuation would be too costly (i.e., when playing with 64 individuals), participants would switch to a social categorization strategy. However, we observed that individuation was the main strategy in all conditions, independently of its cognitive cost and its hindering impact on learning, highlighting the role of motivation over cognitive economy in social perception.

Keywords: individuation, social categorization, learning, cognitive resources

Living in complex social groups has allowed us to develop several abilities related to social competence (Dunbar, 1998; Dunbar & Shultz, 2007). Social relationships are cognitively costly (Dunbar, 2009), and yet we have to manage social interactions on a daily basis. Among the demands of social interactions, we may need to form accurate impressions of others under different circumstances. Imagine, for instance, that you have to find a replacement for an important job position. This may occur in a context in which you are rushed by time pressure and overwhelmed by a large number of applications. Alternatively, you may have only a few applications to evaluate, and enough time to organize a face-to-face interview with all the candidates. In these contexts, you will need to organize and integrate all the information you could get from applicants efficiently. Research in social psychology and social cognition has demonstrated that the processes used to perceive and organize social information lie on a continuum stretching from social categorization to individuation (Fiske, Lin, & Neuberg, 1999; Fiske & Neuberg, 1990).

On one end of the continuum, social categorization is the default strategy to make sense of people (Cloutier, Mason, & Macrae, 2005; Kawakami, Amodio, & Hugenberg, 2017; Macrae & Bodenhausen, 2000; Quinn & Macrae, 2005). In first encounters, we readily encode facial and bodily cues diagnostic of a person's category membership (Bijlstra, Holland, Dotsch, & Wigboldus, 2018; Freeman & Ambady, 2011; Hugenberg, Young, Bernstein, & Sacco, 2010; Stolier & Freeman, 2016). For instance, when searching for the ideal candidate among several job applicants, you could easily and effortlessly categorize them into relevant social categories such as men, women, elderly or young. Once a person has been categorized into a familiar social category, our knowledge or beliefs about this category may inform our judgments (Brewer, 1988; Liberman, Woodward, & Kinzler, 2017; Macrae & Bodenhausen, 2000). For instance, if the position to fill is related with finances and you believe that men are more competent than women in financial decision-making (Bordalo, Coffman, Gennaioli, & Shleifer, 2019), you may be more likely to hire a man than a female counterpart. When the target is relevant enough to us, we may start to attend their individual features, either confirming the initial categorization, or attempting a re-categorization if we perceive discrepancies between individual and categorical attributes. When the re-categorization is impossible or unsatisfactory, we will finally reach the other end of the continuum and achieve a complete individuation. By a piecemeal integration of the target's unique characteristics, we may achieve a fully individual assessment of this person, beyond his or her category membership.

Because individuation requires a deeper analysis and integration of individual attributes, we further our level of analysis to individuating processes only when we have sufficient attentional resources (Gilbert & Hixon, 1991; Macrae, Milne, & Bodenhausen, 1994). Thus, contextual demands consuming cognitive resources such as concurrent task, ego depletion or time pressure promote a greater reliance on category-based information (Fiske & Neuberg, 1990; Govorun & Payne, 2006; Rivers, Sherman, Rees, Reichardt, & Klauer, 2019; Sherman, Macrae, & Bodenhausen, 2000). Conversely, some circumstances may rather enhance motivation and promote individuation processes, as for instance low prejudice level (Lepore & Brown, 1997) or interdependence (Bukowski, Moya, de Lemus, & Szmajke, 2009). For instance, when our accuracy on knowing someone yields economic outcomes, we may favor individual attributes over categorical information to inform our judgements (Kawakami et al., 2014; Neuberg & Fiske, 1987).

Outcome dependency may be recreated in the lab by financially rewarding participants as they make accurate predictions about unfamiliar individuals. Such a procedure has been implemented in the multi-round trust game (King-Casas et al., 2005; Telga, de Lemus, Cañadas, Rodríguez-Bailón, & Lupiáñez, 2018). In this paradigm, participants play several rounds with unfamiliar game partners. At the beginning of each round, participants are endowed with a certain amount of money (e.g., €1) and have to decide whether or not to cooperate with the partner of this round. In case participants do not cooperate, they keep the €1 and move to the next round. In case they cooperate, their partner receives €5 and in turn decides whether or not to reciprocate by either sending back €2.50 to the participant, or keeping the €5 for him or herself. Importantly, in this setting, figuring out the cooperative behavior of each partner is a key strategy to maximize benefits, by cooperating with equitable partners (i.e., those who send back the €2.50 in a reciprocal way) and avoiding cooperation with non-equitable partners (i.e., those who keep the €5 for themselves). To achieve this goal, participants may monitor the individual behaviors of their partners across time, that is, they may adopt an individuation strategy. However, this strategy may only be efficient if they have enough resources to individuate. If they play with a small number of partners, participants should be able to individuate all of them accurately. In contrast, with a large number of partners, individuation may be too demanding and participants may rather rely on categorical information to predict their partner's behaviors.

The aim of the present study was to investigate the impact of the number of partners in a trust game on the differential use of category- vs. individual-based information at zero acquaintance, across repeated interactions, and on later memory about partners. Specifically, in an adaptation of the trust game, participants had to predict the cooperative behaviors of either 8, 16, 32 or 64 partners, to earn economic rewards. In the learning phase, partners' cooperative behaviors were manipulated both at the group and at the individual level. At the group level, we associated two gender categories (i.e., men and women) with a particular group behavior. That is, men were equitable and cooperated on most trials, and women were non-equitable and did not cooperate on most trials. We intentionally established counter-stereotypical associations between gender and cooperation tendency to ensure that participants need to learn them, rather than simply rely on gender stereotypes associating women with more trustworthiness than men (Buchan, Croson, & Solnick, 2008) to perform the task. Moreover, at the individual level, we manipulated the individual consistency of each partner with respect to their group membership. Within each gender group, 75% of the partners were consistent and displayed the group behavior, while 25% of the partners were inconsistent and displayed a pattern of cooperation opposite to the group behavior. Namely, 25% of women (i.e., inconsistent) were equitable and 25% of men (i.e., inconsistent) were non-equitable.

With this procedure, participants' cooperation pattern with inconsistent partners is informative of their learning strategies. If participants use categorical information, they should show the same pattern of cooperation with consistent and inconsistent partners within the same gender group, according to the group behavior. Alternatively, if participants use individual information, they should show opposite patterns of cooperation with consistent versus inconsistent partners within each gender group. With the equitable group, participants should cooperate more with consistent partners and less with inconsistent partners. In contrast within the non-equitable group, participants should cooperate more with inconsistent partners and less with consistent partners.

Furthermore, we analyzed whether participants' memory of their partners was impacted by their group's behavior and individual consistency during the trust game. Finally, in a transfer phase, we tested participants' cooperation decisions with new men and women with whom they had no prior experience. If their decisions to cooperate with these new partners were impacted by categorical thinking, they should make their decisions according to the associations established in the learning phase. Alternatively, if they adopt an individuation strategy, they should cooperate with new partners independently of their gender.

We broadly expected participants to individuate their game partners in the learning phase, motivated by the guarantee of an economic reward proportional to their performance (Fiske & Neuberg, 1990; Neuberg & Fiske, 1987). Importantly, however, individuating processes should be efficient only when playing with a manageable number of partners. Therefore, participants playing with 8 partners were expected to fully individuate their partners, learning accurately their partners' cooperation trends and cooperating with consistent and inconsistent partners according to their individual behavior. Participants playing with 16 and 32 partners were also expected to make an individuation attempt. Nevertheless, learning was expected to linearly decrease across these experimental groups as the cost of individuating increased, for both consistent and inconsistent partners, in line with the increase in the number of partners.

Importantly, we expected this individuation attempt to take place only up to a point after which a categorization strategy would dominate. Thus, the leap from 32 to 64 partners was expected to result in a shift in strategy, from an attempt to individuate in the group playing with 32 partners, to a categorization strategy in the group playing with 64 partners. In this case, participants playing with 64 partners should use category-based information to a greater extent than any other group, learning accurately about consistent individuals and cooperating with them according to the group behavior better than the other experimental groups interacting with fewer partners. Therefore, for consistent partners, we predicted a similar magnitude in learning in the groups playing with 8 and 64 partners, but driven by different mechanisms. While a large learning in the group playing with 8 partners should be the result of a perfect individuation, learning in the group playing with 64 partners should be large in the consistent condition because of a greater reliance on the predictive categorical information. Furthermore, in the group playing with 64 partners, the categorization strategy should result in inaccurate decisions with inconsistent partners, as shown in Figure 4.1, so that cooperation with inconsistent partners, rather than depending on their individual cooperative behavior, would entirely depend on the group behavior.

In the transfer phase, we also predicted an increased reliance on categorical information as the number of partners increased. Participants playing with 8 individuals were expected to individuate new partners with whom they had no prior experience, but a tendency to categorize would emerge in the group playing with 16 participants and linearly increased across the experimental groups up to a full categorization in the group playing with 64 partners. The hypotheses, procedure and analyses of this experiment were pre-registered on Open Science Framework (https://osf.io/4xhwg).

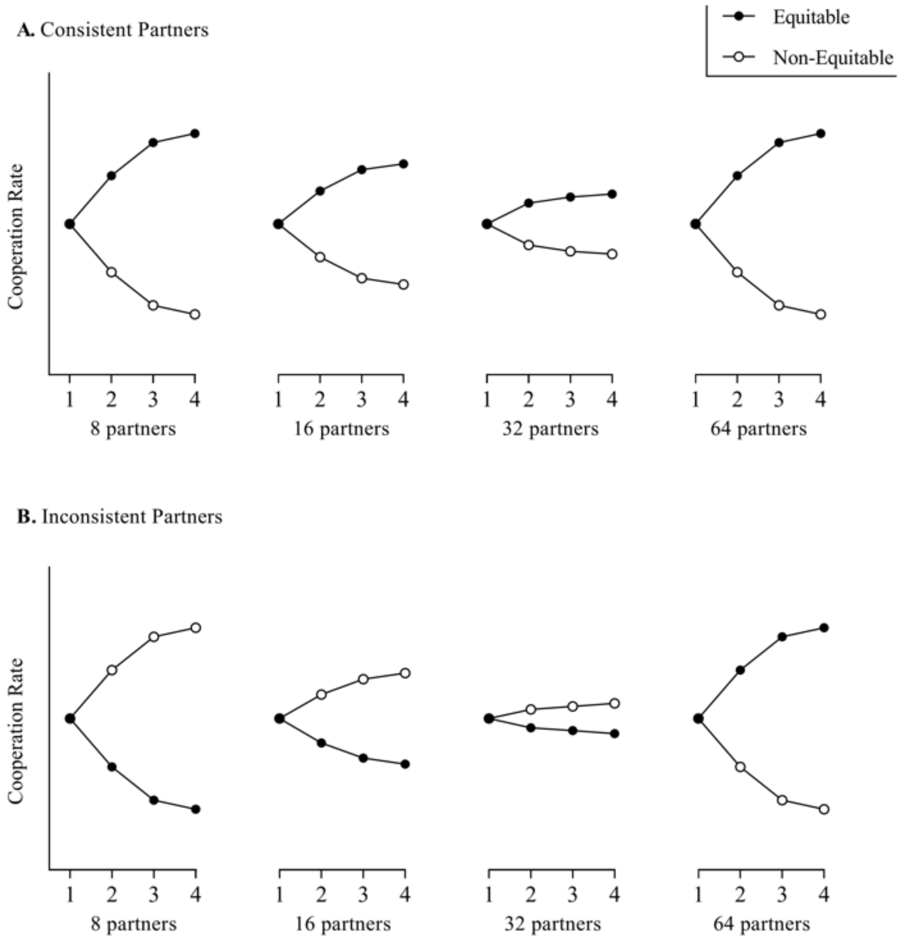

Figure 4.1. Expected results in the learning phase for (A) consistent and (B) inconsistent partners. The groups playing with 8, 16 and 32 partners were expected to attend the individual attributes of their partners, resulting in a pattern of cooperation in line with their individual cooperative trend in both consistent and inconsistent conditions (i.e., opposite patterns of cooperation for consistent and inconsistent partners). The group playing with 64 partners, however, was expected to use a categorization strategy, resulting in accurate cooperation decisions with consistent partners and inaccurate cooperation decisions with inconsistent partners, as driven by categorical information.

Method

Participants. On the basis of previous studies (Telga et al., 2018), we decided prior to data collection to run 40 participants in each experimental group. Sensitivity analyses assuming an alpha criterion of 0.05 and a power criterion of 0.80 revealed that with this sample, the smallest effect size that could have been detected for the critical Number of Partners x Group Behavior x Consistency interaction was $f = .11$. Therefore, 160 female volunteers (average age: 21.57 years, range: 18-30 years old) took part into the experiment in exchange for financial compensation, depending on their accuracy in the task (€6.50 on average). Written informed consent was obtained from all participants. This study is part of a larger project approved by the local ethical committee

(175/CEIH/2017).

Apparatus, materials and stimuli. E-prime 2 software (Schneider, Eschman, & Zuccolotto, 2002) was used for stimuli presentation and data acquisition. Stimuli were displayed on a 24-inch monitor placed at 60 cm from participants. Eighty photographs of 40 men and 40 women with neutral emotional expression were taken from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015) to represent the game partners. Out of the 80 faces, 16 were used as fillers for the memory and transfer phases, whereas the remaining faces were used in the learning phase.

Procedure. Participants were randomly assigned to one of the 4 experimental conditions, playing with either 8, 16, 32 or 64 partners across different phases, as shown in Figure 4.2.

Each trial began with "€1" for 190 ms in the center of the screen, indicating that participants virtually received €1. Next, a fixation cross appeared for 500 ms, followed by the photograph of a partner (8.10º x 10.95º) during 1500 ms. Participants had to press the '1' key if they decided to cooperate, or the '0' key if they chose not to cooperate with the partner. When participants chose not to cooperate, they kept the initial amount (i.e., €1) and their partner would receive €0. When participants decided to cooperate, their game partners received the initial amount multiplied by 5 (i.e., €5). Then, the partner could in turn decide whether to be equitable, returning half of the profit (i.e., € 2.50), or non-equitable by keeping the €5 received. Once a decision was made, or after 1500 ms, another fixation cross appeared for 500 ms, followed by a 1000 ms display with the information about the partner's decision and the consequent participants' outcome (e.g., "Your partner decided to keep the €5, you receive €0"). Participants played the trust game in three phases.

First, in the learning phase, each gender group was associated with a cooperative group behavior, either equitable or non-equitable. Specifically, men were equitable, cooperating on 75% of the trials, while women were non-equitable, cooperating on 25% of the trials. Moreover, within each group, we also manipulated the individual consistency of each partner with respect the rest of the group. Consistent partners displayed the group behavior while a small proportion of inconsistent partners displayed a pattern of cooperation opposite the group behavior. That is, 25% of men were non-equitable, cooperating on 25% of the trials, while 25% of women were equitable, cooperating on 75% of the trials.

To maintain the length of the task equal across the four experimental groups, the number of presentations of each partner varied depending on the number of partners with whom participants played the trust game. Partners were presented 64 times in the group playing with 8 partners, 32 times in the group playing with 16 partners, 16 times in the group playing with 32 partners, and 8 times in the group playing with 64 partners, resulting in 512 trials in the learning phase for all groups. This manipulation was consistent with our goal to increase the difficulty of individuating partners across experimental groups. Auto-administered breaks were allowed every 64 trials. The set of photographs used to represent partners was counterbalanced in the groups playing with 8, 16 and 32 partners such that across participants, all the photographs used in the group playing with 64 partners were also used in the three other experimental groups. The photographs associated with the inconsistent partners were also counterbalanced across participants.

Next, participants performed a 5-minute cancellation task in which they were asked to mark every letter “Q” found among a large array of letters “O”. This task served as a distraction for the upcoming memory test.

The third phase corresponded to the memory test, in which participants were presented with 8 game partners from the learning phase and 16 new individuals (8 men and 8 women) with whom they had no prior experience. Participants were asked to press the “1” key when they recognized the target as a previous game partner, or to press the “0” when they did not. When participants recognized the individual as a previous partner, they were then asked to press the “1” key if they thought the individual usually cooperated in the learning phase, or the “0” key if they thought the target usually did not cooperate during the learning phase. The set of 8 pictures taken from the learning phase was counterbalanced across the participants in the groups playing with 16, 32 and 64 partners.

Finally, the last phase of the experiment was the transfer, in which all participants played the trust game with 16 new partners. Eight men and 8 women were presented twice, once being equitable and once being non-equitable resulting in 32 trials. The pictures that served as fillers in the memory phase were used to represent the new partners in the transfer phase. Overall, the experiment lasted around 60 minutes.

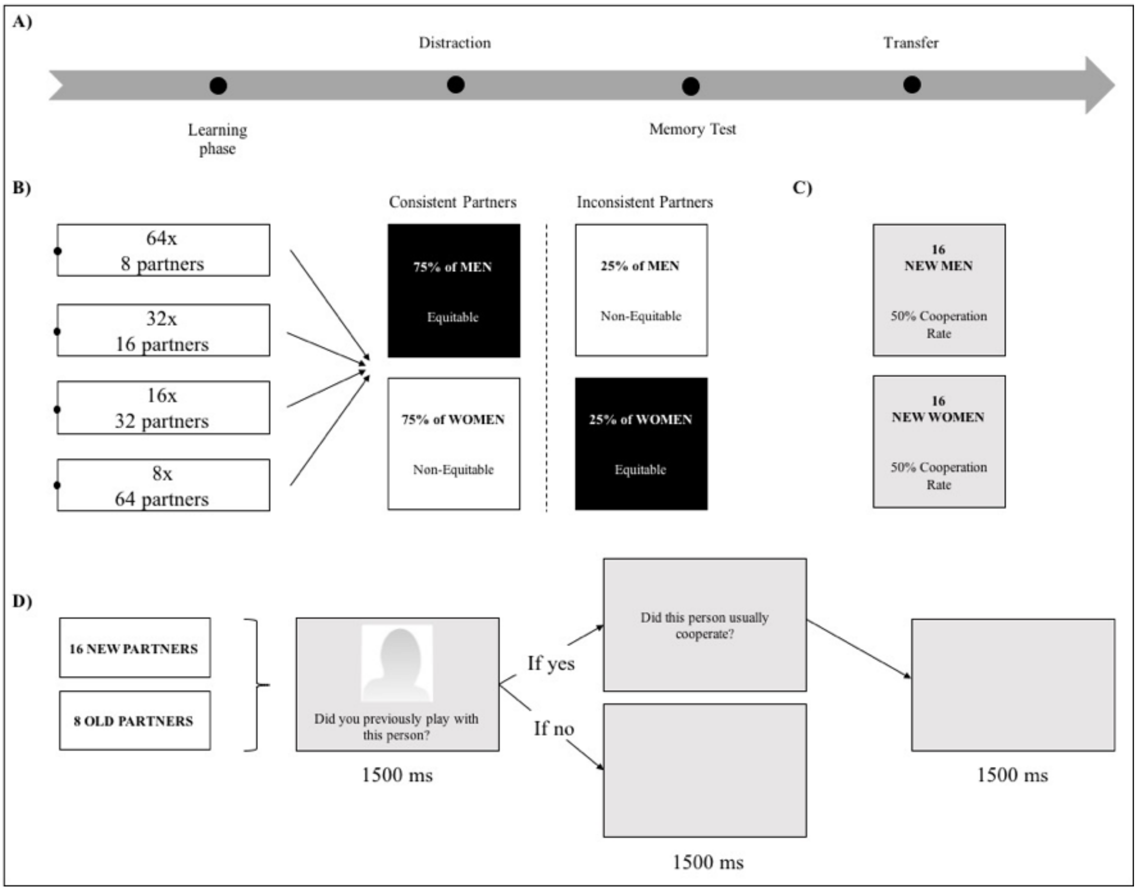

Figure 4.2. Timeline of the experiment (A) and specific procedures introduced in the learning (B), transfer (C) and memory test (D). Equitable partners (i.e., those who cooperate on 75% of the trials) are represented in black and non-equitable partners (i.e., those who cooperated on 25% of the trials) are represented in white. Partners are represented in grey when their cooperation rate is left at 50%.

Design and dependent variables.

Learning phase. For the analyses of the learning phase, the 512 trials were grouped in four blocks of 128 trials. Therefore, the learning phase involves a 2(Group Behavior: equitable vs. non-equitable) x 2(Consistency: consistent vs. inconsistent) x 4(Block: 1-4) x 4(Number of partners: 8, 16, 32, 64) mixed-design with group behavior, consistency and blocks as within-participants variables, and number of partners as a between-participants factor. The dependent variable (DV) was participants' cooperation rate. With this design, a main effect of group behavior, not qualified by the consistency variable, would indicate a complete categorization strategy. Alternatively, a Group Behavior x Consistency interaction would reveal that participants behave differently with consistent vs. inconsistent partners within the same group, and therefore that they attend to their individual attributes. A completely opposite pattern of cooperation for consistent and inconsistent partners would reveal that participants fully individuate their partners. An opposite pattern for inconsistent partners but of a reduced size would indicate inaccurate individuation. The magnitude of the difference in cooperation between inconsistent individually equitable and inconsistent individually non-equitable partners, compared to the same difference for consistent partners - and therefore the effect size of the interaction - indicates the strength of the individuation pattern. The more similar is this magnitude between consistent and inconsistent individuals, the more participants individuate their partners.

To have a sense of participants' learning in this phase, we also computed the aforementioned magnitude in a learning index by subtracting participants' cooperation rate with partners who were individually non-equitable partners from their cooperation rate with individually equitable partners. With this DV, we could directly compare individual learning about consistent vs. inconsistent individuals: the larger the learning, the higher the scores. This learning involves a 2(Consistency: consistent vs. inconsistent) x 4(Block: 1-4) x 4(Number of Partners: 8, 16, 32, 64) mixed-design, with consistency and block as within-participants variables and the number of partners as a between-participants factor. The absence of a consistency effect would indicate that participants equally learn about consistent and inconsistent partners, and therefore fully individuate their partners. Alternatively, an effect of consistency would reveal that participants do not learn to the same extent about consistent and inconsistent individuals, and therefore fail to individuate. Finally, opposite learning effects for consistent and inconsistent partners would indicate the categorization of partners.

Memory. In the recognition phase, for new partners, we used a univariate design with number of partners (8, 16, 32, 64) as a between-participants variable to analyze participants' false alarms (i.e., mistaken recognition of new partners -those with whom they have not played in the learning phase- as old partners –those with whom they have played in the learning phase). For old partners, we analyzed how their group behavior and individual consistency impacted memory. Therefore, we used a 2(Group Behavior: equitable vs. non-equitable) x 2(Consistency: consistent vs. inconsistent) x 4(Number of partners: 8, 16, 32, 64) mixed design, the last variable being a between-participants factor, for both the recognition and the recall phase. The DV was participant's accuracy, corresponding to the rate of old partners correctly recognized as old for the recognition phase, and the rate of correct judgments of old partners' cooperation tendency for the recall phase.

Transfer. In the transfer, we aimed at verifying whether new partners' group membership (i.e., whether they belonged to the group that was equitable or non-equitable in the learning phase) impacted participants' cooperation decisions with the new partners. Therefore, we used a 2(Partner group: equitable vs. non-equitable) x 4(Number of partners: 8, 16, 32, 64) mixed-design, with partner group as a within-participants variable, and number of partners as a between-participants factor. The DV was participants' cooperation rate. A categorization strategy would be reflected in a main effect of the partner group variable, if participants cooperate more with partners from the group that was equitable in the learning phase, and less with partners from the group that was not equitable in the learning phase. Alternatively, an individuation strategy would be observed if participants cooperate with new partners independently of their gender.

Results

Consistently with the pre-registered criteria, nine participants with cooperation rate above 80% or under 20% were replaced to maintain 40 participants in each experimental group, and trials with RTs shorter than 200 ms (3%) were also filtered out.

Learning. To examine whether participants' learning about their partners was impacted by categorization or individuation strategies, cooperation rates in the learning phase were subjected to a 2(Group Behavior: equitable vs. non-equitable) x 2(Consistency: consistent vs. inconsistent) x 4(Block: 1, 2, 3, 4) x 4(Number of Partners: 8, 16, 32, 64) mixed-design ANOVA.

The predicted Group Behavior × Consistency × Number of Partners interaction was significant, $F(3, 156) = 30.78$, $p < .001$, $\eta_p^2 = .37$. Although the Group Behavior x Consistency interaction indicating an individuation strategy was significant in the four experimental groups, $F(1, 39) = 158.44$, $p < .001$, $\eta_p^2 = .80$, for 8 partners, $F(1, 39) = 129.78$, $p < .001$, ηp2 = .77, for 16 partners, $F(1, 39) = 126.92$, $p < .001$, ηp2 = .77, for 32 partners, and $F(1, 39) = 40$, $p < .001$, $\eta_p^2 = .51$, for 64 partners, the smaller the number of partners, the stronger the pattern of individuation. This interaction was significantly modulated by the block variable, with the four-way interaction, $F(9, 468) = 2.11$, $p = .027$, $\eta_p^2 = .04$, indicating that the pattern of individuation increased across blocks as a result of learning, as shown in Figure 4.3. Additional analyses on the individual learning index revealed that the main effect of consistency was not significant in the group playing with 8 partners, $F(1, 39) < .01$, $p = .952$, $\eta_p^2 < .01$, but was significant in each of the other three groups, smaller $F(1, 39) = 5.95$, $p = .019$, $\eta_p^2 = .13$, in the group playing with 16 partners. Therefore, participants playing with 8 partners learned to the same extent the cooperative behaviors of consistent and inconsistent partners. In contrast, participants playing with 16, 32 and 64 partners learned more the cooperative tendencies of consistent than inconsistent individuals.

To have a sense of the differences in learning between the four experimental groups, we analyzed participants' cooperation decisions separately for consistent and inconsistent partners. Separate mixed-design ANOVAs were conducted on cooperation with consistent and inconsistent targets, with group behavior (equitable vs. non-equitable) and block (1-4) as within-participants variables, and number of partners (8, 16, 32, 64) as a between-participants variable.

For **consistent partners**, the Group Behavior x Number of Partners interaction was significant, $F(3, 156) = 7.34$, $p < .001$, $\eta_p^2 = .12$, indicating that although participants in the four experimental groups cooperated more with consistent equitable than with consistent non-equitable partners, smallest $F(1, 39) = 86.66$, $p < .001$, $\eta_p^2 = .69$ in the group playing with 64 partners, learning (i.e., the difference between equitable and non-equitable) about consistent individuals linearly decreased across groups as the number of partners increased, $p < 001$. However, in contrast to our hypothesis, we observed no hint of a quadratic trend in learning, $p = 854$. Therefore, the predicted qualitative change in the group interacting with 64 partners towards larger categorical learning was not observed. It is important to note that, although all groups were evenly-matched regarding the categorical information, with consistent men cooperating 75% of the times, and women only in 25%, the larger the number of partners the less participants learned about the gender information as shown in Figure 4.3.

For **inconsistent partners**, the Group Behavior x Number of Partners interaction was also significant, $F(3, 156) = 32.49$, $p < .001$, $\eta p2 = .39$. Here, the main effect of group behavior indicated that participants cooperated more with inconsistent partners belonging to the non-equitable group than with those belonging to the equitable group in the groups playing with 8, 16 and 32 partners, smallest $F(1, 39) = 12.35$, $p = .001$, $\eta_p^2 = .24$ in the group playing with 32 partners. However, when playing with 64 partners, participants did not significantly differ in their cooperation between equitable and non-equitable individuals, $F(1, 39) = 1.75$, $p = .194$, $\eta_p^2 = .04$. Thus, participants learned about the individual cooperation rate of inconsistent partners, and this learning linearly decreased as the number of partners increased, $p < .001$, up to a null learning when interacting with 64 partners, as shown in Figure 4.3. Again, although learning decreased with the number of partners, the predicted shift to a categorization strategy with 64 partners was clearly absent. Instead, learning about inconsistent individuals decreased across the four experimental groups and seemed to fall on some stage of the categorization-individuation continuum (Fiske & Neuberg, 1990), reflecting the extent to which participants were cognitively *able* to assess the individual cooperative behaviors of their partners, up to a null learning when interacting with 64 partners.

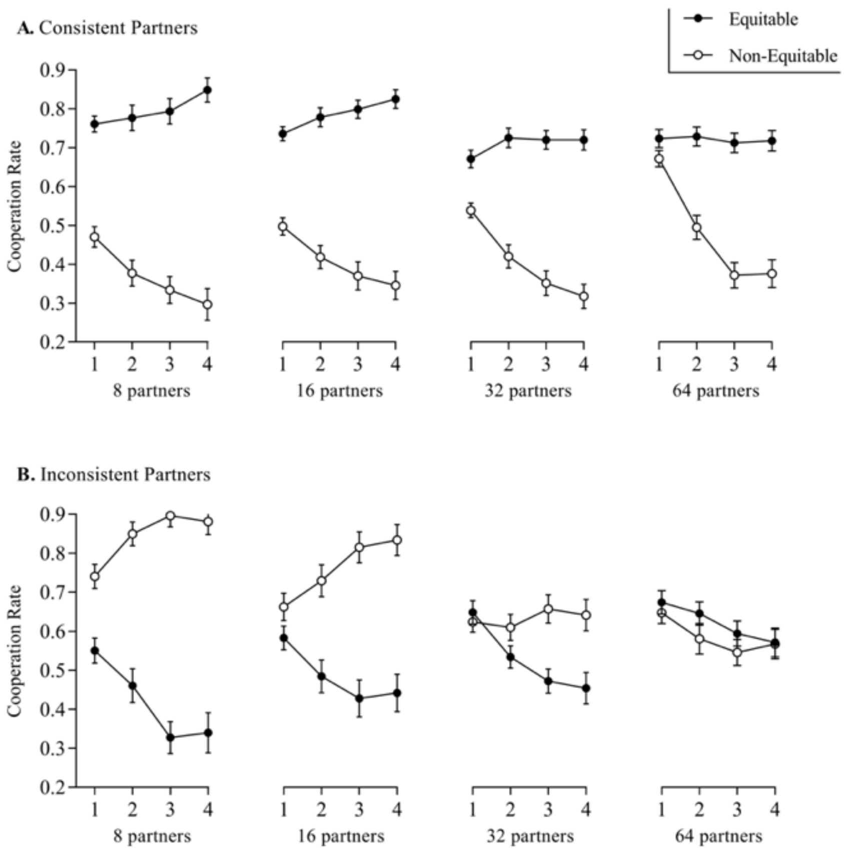

Figure 4.3. Cooperation rates with (A) consistent and (B) inconsistent partners across blocks in the four experimental group. Error bars represent the standard error of the mean.

Memory. To examine how the number of partners and participants' learning strategies impacted their memory of the partners, we analyzed memory scores in two stages. Recognition scores informed about participants' accuracy at identifying the partners they had interacted with in the learning phase among new individuals with whom they had no prior experience. Recall scores reflected participants' memory of the specific behavior of their partners (i.e., whether they were equitable or non-equitable) in the learning phase.

Recognition performance. For **new partners**, a univariate ANOVA with the number of partners (8, 16, 32, 64) as a between-participants variable was conducted on false alarms (i.e., the mistaken recognition of new partners as old partners**)**, revealing a significant main effect of the number of partners, $F(3, 156) = 12.16, p < .001, \eta_p^2 = .19$. Post hoc comparisons using Bonferroni corrections indicated that new partners were incorrectly taken as old partners more often in the group playing with 64 partners than in any other group, all *p*s < 001. In contrast, almost no false alarms were committed in the groups playing with 8, 16 and 32 partners, with no significant differences between groups, all *p*s > .999, as shown in Table 1.

Recognition for **old partners** were subjected to a mixed-design ANOVA with group behavior (equitable vs. non-equitable) and consistency (consistent vs. inconsistent) as within-participants factors and number of partners (8, 16, 32, 64) as a between-participants variable. The Group Behavior x Consistency x Number of Partners interaction was significant, $F(3, 149) = 3.07, p = .030$, η*p2* = .06. In fact, the Group Behavior x Consistency interaction was significant in the group playing with 8 partners, $F(1, 37) = 8.14, p = .007, \eta_p^2 = .18$ and in the group playing with 16 partners, $F(1, 37) = 13.55, p = .001, \eta_p^2 = .27$, indicating that participants recognized better partners who were individually equitable (consistent partners from the equitable group and inconsistent partners from non-equitable group) than partners who were not (inconsistent partners in the equitable group and consistent partners in the non-equitable group) as shown in Table 1. In contrast, this two-way interaction was not significant in the groups playing with 32 and 64 partners, $F(1, 39) = .89, p = .352, \eta_p^2 = .02$, and $F(1, 39) = .39, p = .536, \eta_p^2 = .01$, respectively, suggesting that the individuation attempt in the learning was successful mostly for the groups playing with 8 and 16 partners.

Recall performance. A mixed-design ANOVA with group behavior (equitable vs. non-equitable) and consistency (consistent vs. inconsistent) as within-participants factors and number of partners (8, 16, 32, 64) as a between-participants variable was conducted on recall accuracy (i.e., the percentage of correctly recalled cooperation tendencies). We found a significant main effect of consistency, $F(1, 105) = 17.80, p < .001, \eta_p^2 = .15$, indicating that participants recalled better the cooperative behavior of consistent than inconsistent individuals, as shown in Table 4.1. The main effect of group behavior was also significant, $F(1, 105) = 4.61, p = .034, \eta_p^2 = .04$, revealing that participants recalled better the cooperative behavior of non-equitable than equitable individuals. Finally, the main effect of the number of partners was also significant, *F*(3,

105) = 3.04, $p < .032$, $\eta_p^2 = .08$. Post hoc comparisons using Bonferroni corrections revealed that recall accuracy was significantly lower in the group playing with 64 partners than in the group playing with 8 partners, $p = .032$.

Table 4.1. Mean scores (and standard deviations) in the memory test.

Recognition		8 partners	16 partners	32 partners	64 partners
Equitable	Consistent	.97 (.11)	.98 (.07)	.98 (.07)	.93 (.19)
	Inconsistent	.95 (.22)	.85 (.37)	.87 (.34)	.95 (.19)
Non-Equitable	Consistent	.88 (.19)	.84 (.21)	.93 (.14)	.93 (.14)
	Inconsistent	1 (.00)	.97 (.16)	.90 (.30)	.90 (.31)
False alarms		.02 (.07)	.01 (.02)	.02 (.03)	.11 (.16)
Recall					
Equitable	Consistent	.87 (.21)	.86 (.22)	.78 (.21)	.67 (.32)
	Inconsistent	.75 (.44)	.73 (.45)	.78 (.42)	.55 (.51)
Non-Equitable	Consistent	.91 (.24)	.94 (.14)	.89 (.22)	.92 (.15)
	Inconsistent	.92 (.28)	.76 (.43)	.63 (.49)	.61 (.50)

Transfer. To verify whether participants used their knowledge from the learning phase to categorize new individuals with whom they had no prior experience, we conducted a mixed-design ANOVA on cooperation rates in the transfer phase with partner group (equitable vs. non-equitable) as a within-participants factor, and number of partners (8, 16, 32, 64) as a between-participants variable. The Partner Group x Number of Partners interaction was significant, $F(3, 156) = 2.75$, $p = .044$, $\eta_p^2 = .05$. Participants playing with 8 or 16 individuals cooperated more with partners belonging to the group that was non-equitable in the learning phase than with partners from the group that was equitable in the learning phase, $F(1, 39) = 6.10$, $p = .018$, $\eta_p^2 = .14$, and $F(1, 39) = 5.78$, $p = .021$, $\eta_p^2 = .13$, respectively, as shown in Figure 4.4. However, participants playing with 32 and 64 partners did not significantly differ in their cooperation between partners from the equitable group and partners from the non-equitable group, $F(1, 39) = .27$, $p = .609$, $\eta_p^2 = .01$, and $F(1, 39) < .01$, $p = .979$, $\eta_p^2 < .01$, respectively.

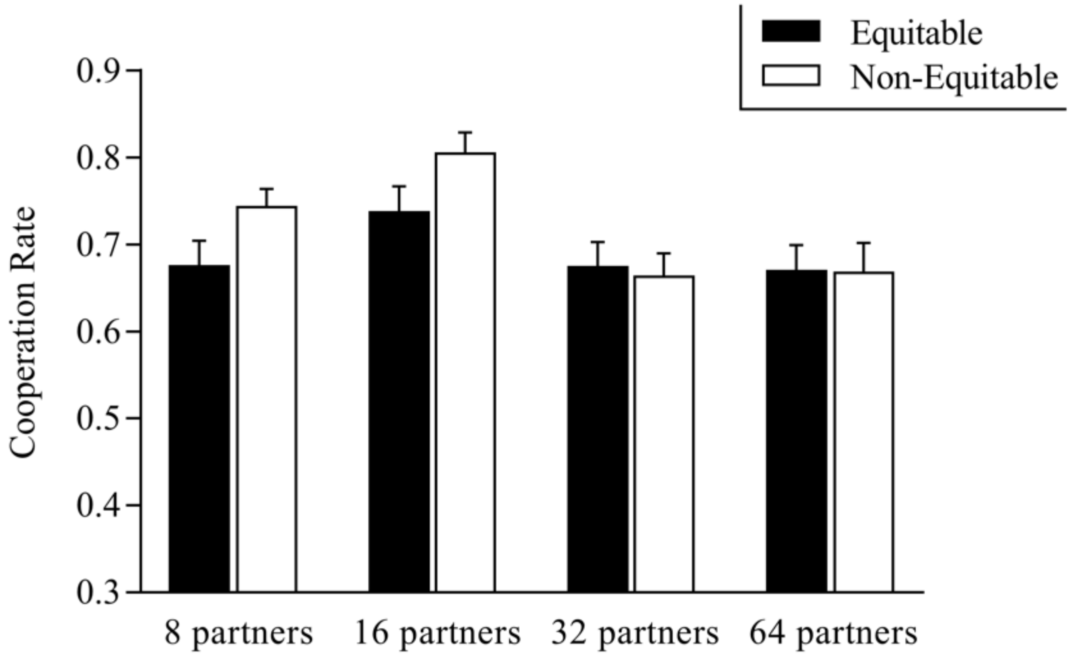

Figure 4.4. Cooperation rates in the transfer phase with new partners from the equitable vs. non-equitable groups in the learning phase, in the four experimental groups. Error bars represent the standard error of the mean.

Discussion

The present research explored how the attentional resources needed to individuate targets impact social learning strategies and the subsequent memory of these individuals. In an adaptation of the trust game, participants had to predict the cooperation tendencies of either 8, 16, 32 or 64 partners to earn economic rewards. Because participants were outcome-dependent on the accuracy of their predictions, we expected this context to promote an individuation strategy. However, given the attentional demands of individuating processes, we expected individual learning about partners to linearly decrease across groups as the number of partners increased. Importantly, in the group playing with 64 partners, we expected a qualitative change in learning strategies from an attempt to individuate in the group playing with 32 partners to a complete categorization strategy in the group playing with 64 partners, resulting in accurate predictions about consistent partners but incorrect decisions with inconsistent partners. This prediction was not verified and the data rather suggested a persistent prioritization of individual information in contexts of outcome dependency, even when individuation was impossible.

First, we did observe that most participants (except the group interacting only with 8 partners) learned more about consistent than inconsistent partners, suggesting that they attended their partners' categorical attributes in the first stages of learning. On the basis of this categorical information, confirmatory categorization of consistent partners was easier than learning the individual behaviors of inconsistent partners, resulting in a larger learning for consistent than inconsistent individuals (Fiske & Neuberg, 1990).

However, and importantly, all participants displayed opposite patterns of cooperation with consistent and inconsistent individuals within the same gender group, indicating that all participants attended their partners' individual attributes and noticed that inconsistent partners' behavior deviated from the group's. The economic rewards associated with learning may have motivated participants to further their level of analysis beyond category membership (Kawakami et al., 2014; Neuberg & Fiske, 1987). Nonetheless, individuating processes were not always achievable. Despite all participants may have been motivated to individuate their partners, not all of them were actually able to individuate successfully. The increase in the number of partners across the experimental groups was associated with increased attentional resources needed to individuate. Because participants persisted in using an individuation strategy, its hindering impact on learning increased across the four experimental groups.

In fact, the decrease in learning about both consistent and inconsistent partners indicates a dominant individuation strategy in all groups. Importantly, only the group playing with 8 partners did not differ in learning from consistent vs. inconsistent partners, indicating that only those participants were cognitively able to achieve a fully individual assessment of their partners' behaviors. However, not even the group playing with 64 partners was discouraged from using individual information in spite of the cognitive cost of such strategy. Playing with 64 partners implied a qualitative leap compared to the three other experimental groups, as reflected in the poorer memory performance of this group in the recognition test. Still, participants playing with 64 partners did not show the predicted greater reliance on categorical information than the other groups, which would have resulted in larger learning from consistent partners. This result strongly suggests that, just as participants playing with 8, 16 and 32 partners, participants playing 64 partners were still motivated to leave out category-related information in very early stages of learning, even in this situation in which they were not able to determine the individual behaviors of inconsistent individuals, broadly resulting in an impaired learning.

It is noteworthy that with our procedure, the number of interactions with each individual partner linearly decreased across groups, as the number of partners increased. Taking into account that learning about both consistent and inconsistent partners linearly decreased across groups, one may argue that our results may be simply explained by the between group differences in the number of presentations of each partner. To test this alternative explanation, we analyzed the cooperation rates in the four experimental groups including only the first 8 interactions between participants and their partners. Thus, we compared, for instance, the first half of Block 1 in the group interacting with 8 partners, with the whole data set of the group interacting with 64 partners. We observed that the critical Group Behavior x Consistency x Number of Partners interaction was significant when all the groups were evenly-matched on the number of interactions with each partner, $F(3, 156) = 4.08$, $p = .008$, $\eta_p^2 = .07$, indicating that different patterns of learning between groups already emerged in the first 8 interactions with partners, as shown in Figure 4.5. All these findings consistently indicate that participants tried to individuate their partners, but this individuating learning linearly decreased across groups, as the number of partners increased.

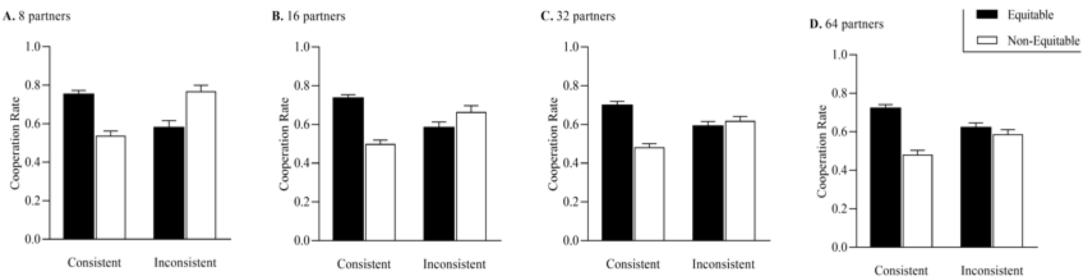

Figure 4.5. Cooperation rates with partners as a function of the group behavior and individual consistency in the first 8 interactions for the three experimental groups. Note that this analysis includes data from the first half of the first block of trials for participants playing with 8 partners, the first block of trials for participants playing with 16 partners, the two first blocks for participants playing with 32 partners, and the whole learning phase for participants playing with 64 partners.

The goal of individuating partners was also reflected in the transfer phase. In fact, the first and evident result from this phase is that there was no hint of categorization of new partners according to the associations established in the learning phase. This result is coherent with previous studies showing that female participants do not transfer categorical information to new individuals (see Telga & Lupiáñez, 2019 for a meta-analysis). Therefore, in the present study, given that women were the non-equitable group during the learning phase, once all new partners cooperated at chance levels during the transfer phase, participants showed instead an ingroup bias cooperating more with women than with men, commonly found in women's first interactions with unknown game partners when they cannot predict their behaviors (Telga, Alcalá, & Lupiáñez, 2019;

Telga et al., 2018; Telga & Lupiáñez, 2019). This ingroup bias was observed only in the groups playing with 8 and 16 partners, which were the groups who achieved an accurate individuation. Therefore, participants from these groups were not impacted by the categorical associations established in the learning phase, and treated these new individuals from the transfer phase as they would before learning, influenced by ingroup favoritism. However, participants playing with 32 or 64 partners seem to have been impacted by categorical information during learning to a greater extent than the groups playing with a smaller number of partners, because they could not achieve an individuated impression of their partners. This small categorical learning, opposite to the ingroup favoritism, may have been sufficient for participants to overcome the spontaneous ingroup favoritism, resulting in no bias in the transfer phase.

Finally, in the memory phase, when participants were no longer outcome-dependent on making accurate behaviors attribution to their partners, their pattern of responses reflected a predominance of categorical information in all groups. In fact, partners from the non-equitable group were better recalled than the ones from the equitable group, in line with previous studies reporting that reputational memory for uncooperative individuals is better than for cooperative targets (Hechler, Neyer, & Kessler, 2016). Moreover, consistent individuals' behaviors were better recalled than those of inconsistent individuals. It is interesting to note that these two effects reflecting a categorical recall were found even in the group playing with 8 partners, who achieved a complete individuation of their partners in the learning phase. Importantly, this result suggests that the economic rewards associated with performance in the trust game may have been highly responsible for participants' high motivation to individuate, and once this motivation disappeared, so did the cognitive effort allowing individuation processes. Alternatively, and according to the literature suggesting that category-related information is better remembered than individual characteristics in demanding context (Macrae, Hewstone, & Griffiths, 1993; Pendry & Macrae, 1999), this result may imply that individuating 8 partners was sufficient to challenge participants' cognitive resources. Nevertheless, we consider this highly unlikely given the magnitude of learning during the learning phase.

Altogether, these results indicate that when participants were motivated to individuate their partners by economic outcomes, all groups tried to assess their partners' behaviors on the basis of their individual attributes, even those participants who played with a large number of partners. However, this strategy resulted in different degrees of learning because of the increasing attentional demands of individuating across the four experimental groups. Interestingly, when the outcome-dependency disappeared, participants' responses were mostly driven by categorical attributes, reflecting a decrease in the motivation to individuate.

The present research demonstrates that the amount of information to process impacts social perception in different instances, but also raises some questions to be addressed in future research. For instance, with the present experimental design, we cannot explore the consequences of using counter-intuitive associations between gender and group behavior in the learning phase on learning, memory and transfer. This issue may be addressed in future design by counterbalancing the associations between gender and group behavior in the learning phase as in previous research (Telga et al., 2018). Moreover, we do not have a full understanding of the implications of presenting the same photographs in the memory and in the transfer phases to represent new partners. Because single exposition to a face may impact attitudes toward this target (Zebrowitz, White, & Wieneke, 2008), it may be interesting to verify whether a mere exposure effect has impacted the data in the transfer phase. Given that the four experimental groups were evenly-matched in terms of pre-exposure before performing the transfer phase, it is unlikely that this possible issue affected our main conclusions regarding the between group differences. Nonetheless, a replication of this study using different photographs in the memory and in the transfer phases seems to be the most reasonable way to ensure it.

Overall, this study suggests that when people are motivated to individuate targets, as for instance, in outcomes-dependency situations (Kawakami et al., 2014; Neuberg & Fiske, 1987), they may be ready to engage all the cognitive resources needed to achieve this goal, even in highly demanding contexts. To some extent, these results challenge the argument of a resources-saving function of social categorization (Allport, 1954; Macrae & Bodenhausen, 2001; Macrae et al., 1994; Sherman et al., 2000) in a highly motivating context. In fact, if social categorization is a tool to achieve a greater efficiency in information processing (Bodenhausen & Macrae, 1998; Kutzner & Fiedler, 2017; Quinn & Macrae, 2005), then it should be used whenever individuating processes negatively impacts performance (i.e., in cognitively demanding task environments). However, the results from this study rather indicated that participants tried to individuate their partners at any cost, even though a categorization strategy would have been more efficient and easier to apply. While attentional resources and motivation have been argued to be the principal determinants of social perception strategies (Fiske & Neuberg, 1990), the present research confers a more than considerable weight to motivational factors.

Acknowledgments and funding

We would like to thank to all the volunteers who participated in the study. This work was supported by the Spanish Ministry of Education, Culture and Sports, with pre-doctoral FPU fellowship FPU14/07106 to MT, and the Spanish Ministry of Economy and Competitiveness, with research projects PSI2014-52764-P and PSI2017-84926-P to JL. This research is part of MT's thesis dissertation under the supervision of JL.

Correspondence concerning this article should be addressed to Maïka Telga, Departamento de Psicología Experimental, Universidad de Granada, Campus de Cartuja s/n, 18 071 Granada, Spain. Email: mtelga@ugr.es

CHAPTER 6. Motivational factors

EXPERIMENTAL SERIES 5

The content of this experimental series is in preparation as:

Telga, M., Alcalá, J. A., & Lupiáñez, J. (2019). *Motivated social and non-social categorization in decision making and learning*. [Manuscript in preparation].

Motivated Social and Non-Social Categorization in Decision Making and Learning

Abstract

Categorical processes allow us to make sense of the environment effortlessly, by grouping together stimuli sharing relevant features or attributes. Although these processes occur in both social and non-social contexts, motivational, affective and epistemic factors specific to the social world may motivate individuation over categorization of social stimuli to a greater extent than non-social stimuli. This hypothesis is tested in the present research, by analyzing the use of categorization vs. individuation strategies of social and non-social targets in an adaptation of the trust game. Three experimental groups of participants had to predict the economic outcomes associated with either humans (i.e., social stimuli), artificial races (i.e., social-like stimuli), or paintings (non-social stimuli) to earn economic rewards. In two studies, different patterns of responses were observed in the first presentations of the targets and during learning across repeated interactions. The differences between the three experimental groups suggests that motivation to individuate enhances learning and to some extent, may compensate for the lack of perceptual expertise.

Keywords: learning, social categorization, individuation, motivation

Classifying objects into familiar categories is a basic process known as categorization, allowing perceivers to bring coherence in the diversity of stimuli features and attributes, efficiently and effortlessly (Simon, 1993). To the extent that objects share relevant features such as shape, color, or function, they can be grouped together and distinguished from other objects that do not possess these features. For instance, perceivers may categorize seats with four legs and a back as chairs, and use them adequately without paying much attention to the differences between all exemplars of chairs. Interestingly, categorization is also a key process in social perception. Perceivers readily classify others on the basis of the features related to their group membership (Macrae & Bodenhausen, 2000; Macrae & Bodenhausen, 2001). In fact, attributes informing of people's gender, ethnicity or age are identified in early stages of face perception (Bruce & Young, 1986; Hugenberg, Young, Bernstein, & Sacco, 2010; Stolier & Freeman, 2016), and used to make complex inferences about them (Bodenhausen, Kang, & Peery, 2012; Kawakami, Amodio, & Hugenberg, 2017).

Imagine, for instance, that a perceiver believes that women are more trustworthy than men. Therefore, in social encounters, once a new target has been categorized as a woman, this perceiver may likely adopt cooperative behaviors with this particular target, to a greater extent than he or she would with a man. Importantly, past and recent theorizing in social psychology have demonstrated that social categorization is the most prominent strategy in impressions formation (Cloutier, Mason, & Macrae, 2005; Quinn & Macrae, 2005), with important consequences on a large range of cognitive processes such as traits attribution, learning and memory (Huart, Corneille, & Becquart, 2005; Linville, Salovey, & Fischer, 1986; Rothbart, Evans, & Fulero, 1979; Telga, de Lemus, Cañadas, Rodríguez-Bailón, & Lupiáñez, 2018).

Under some circumstances, perceivers may further their level of analysis beyond category-related information by attending individual attributes informative of a person's individual identity (Fiske & Neuberg, 1990). For instance, they may monitor a target's behaviors individually, and decide whether or not this person may be trusted independently of his or her gender. For this individuating processes to be successful, perceivers need to be cognitively able to process the individual characteristics of targets (Gilbert & Hixon, 1991; Macrae, Milne, & Bodenhausen, 1994). Importantly, being individuation more effortful, cognitive ability is not sufficient to engage in individuation strategies. Motivational factors including power differences (Guinote & Phillips, 2010), or outcome-dependency (Neuberg & Fiske, 1987), may promote either a social categorization or an individuation strategy (Fiske & Neuberg, 1990).

Despite social and non-social categorizations respond to the same need of cognitive efficiency, they are also different on several aspects. First, once social categories are established according to a particular dimension, perceivers immediately identify the groups they fall into (i.e., ingroups) and the groups they do not belong to (i.e., outgroups). The perception of shared characteristics between targets and oneself impact social perception with important implications at the motivational and affective levels (Ellemers, 2012). For instance, self-categorization may trigger cognitive strategies promoting enhanced self-perception, which in turn impacts perception and attitude toward social targets (Tajfel, 1978). A clear example of such strategy is the ingroup favoritism, consisting of a more positive attitude towards ingroup than outgroup members even at zero acquaintance (Tajfel & Turner, 1979).

Moreover, from the first years of life, humans are able to adopt others' perspective and empathize with them, that is, to understand their cognitive and affective states (Frith & Frith, 1999; Shamay-Tsoory, 2011). These processes also impact social perception by decreasing attention to category-related and stereotypic features (Galinsky & Moskowitz, 2000). Such personal involvement seems far less likely with nonsocial stimuli. Even though non-social stimuli may convey some affective and motivational motives, perceivers' capacity to identify with any human seems much more questionable when it comes to non-social stimuli.

Finally, it is noteworthy that social and non-social categorization may differ with regards to the processes underlying the perception of the social world. In fact, social perception is influenced by epistemic motivation, that is, the desire to understand others and the complexity of social relationships. In this sense, specific motives may determine inter-individual differences regarding preferences on how this social knowledge is acquired (Bodenhausen, Todd, & Becker, 2006; Ford & Kruglanski, 1995). Importantly, self-perception, affective factors and epistemic concerns converge to predict a high motivation to understand people, but not objects, as unique individuals.

The aim of the present research was to test whether social vs. non-social stimuli would trigger different motivational strategies of inferences and learning. Specifically, we predicted that participants would be more motivated to learn about social than non-social targets, resulting in individuation processes for the former, and category-based learning for the latter. To test this hypothesis, we adapted the multi-round trust game to investigate how categorization or individuation strategies would impact participants' learning of the associations established between targets and real economic outcomes in three conditions, either manipulating a social, a social-like or a non-social dimension. In all three conditions, participants had to make predictions about 16 exemplars extracted from two categories. On the social dimension, human targets belonging to two gender groups could be categorized as men or women. On the social-like dimension, individuals sharing facial features with humans, could be categorized into one of two artificial races: Lunaris and Taiyos. Finally, on the non-social dimension, targets were paintings from two different artists with unequivocally different styles: Wassily Kandinsky and Jaison Cianelli.

In the present work, social (i.e, humans), social-like (i.e., artificial races), and non-social (i.e., paintings) conditions will be referred to as "dimensions". Within these dimensions, men vs. women, Lunaris vs. Taiyos, and Kandinsky vs. Cianelli will be referred to as "categories", and exemplars within these categories will be referred to as "individuals" or "targets". With these three dimensions, we aim at exploring participants' decisions to invest money on specific targets to earn economic outcomes. Specifically, participants had to verify which targets were worth investing in, allowing to earn monetary rewards, and which targets were not, making them lose the invested money, in three different phases.

First, in a baseline, all targets were worth investing in on half of the trials, and not worth investing on the other half. With this manipulation, we aimed at testing whether participants were biased toward one of the two categories presented, spontaneously investing more with one of them. Participants were expected to invest more with women than with men on the social dimension, because of gender stereotypes associating women with more trustworthiness and cooperative behaviors than men (Dong, Liu, Jia, Li, & Li, 2018). However, on the social-like and non-social dimensions, participants should have no prior knowledge about artificial races' behaviors or painting values, and therefore were expected to show a pattern of investment independent of categories.

In a learning phase, categories were associated with opposite outcomes, such that on each dimension, one category was worth investing in, allowing participants to earn money on most trials, and the other category was not worth investing in, making participants lose money on most trials. For instance, men were worth investing in, whereas women were not. Moreover, within each category, 4 out of 16 targets were inconsistent with respect to the group they belonged to. Following the same example, four women were worth investing in and four men were not. With this procedure, we could verify whether participants' learning is affected by categorization or individuation processes. Similar learning about consistent and inconsistent partners would reflect an individuation strategy. Moreover, the magnitude of learning about inconsistent targets reflects the potential impact of categorization processes. Poorer learning about inconsistent than consistent targets would reflect some category-based influence on learning, while no learning at all about inconsistent targets would reflect a total categorization strategy. In this phase, we expected human targets to be individuated, artificial races to be slightly categorized and paintings to be categorized to a greater extent.

Finally, in a transfer phase, participants were presented with 32 new targets (16 from each category on either dimension) worth investing in on half of the trials, and not worth not worth investing in on the other half. With this manipulation, we aimed at testing whether participants would use the associations established in the learning phase between categories and specific outcomes to make decisions about new targets in a categorical way. In this phase, a categorization strategy would be observed if participants who learned that one category was worth investing in and the other was not spontaneously invest more with new targets of the former category. Because participants played with a fairly large number of targets (i.e., 32), we expected some expression of categorical learning to be reflected in a tendency to invest more on the category that was worth the investment in the learning phase. The hypotheses and analyses were pre-registered on Open Science Framework (https://osf.io/gqc3b for Experiment 1 and https://osf.io/jqpvn for Experiment 2).

Experiment 1

The aim of Experiment 1 was to test our hypotheses on differentiated motivational influences promoting individual processing for social stimuli, categorization processes for non-social stimuli, and an in-between strategy for social-like stimuli. Therefore, participants were randomly assigned to one of three experimental groups, playing with either men and women, Lunaris and Taiyos, or Cianelli's and Kandinsky's paintings, in three phases. In the learning phase, when categories did not predict the outcome associated with each target, we expected participants playing with humans to cooperate more with women than with men, according to gender stereotypes associating women with warmth and communal behaviors, including trustworthiness (Cuddy et al., 2009;

Cuddy, Fiske, & Glick, 2008; Dong et al., 2018). However, we did not expect any differences in participants' cooperation with social-like (artificial races) and non-social (paintings) stimuli. In the learning phase, when specific contingencies were established between targets and economic outcomes, we expected participants to individuate humans to a greater extent than artificial races, which in turn would be individuated more than paintings. Finally, in the transfer phase, because of the rather large number of partners in the trust game, we expected all participants to use categorical knowledge to make investment decisions about new targets.

Method

Participants. One hundred twenty volunteers (9 males) took part in the study (mean age: 20.09, range: 18-31) and received economical rewards proportional to their accuracy in the task (€5.60 on average). Sensitivity analyses revealed that with a sample of 40 participants in each group, the smallest effect that could have been detected for the critical Consistency x Block interaction with an alpha criterion of .05 and a power of .80 was $f = .23$.

Stimuli selection. Several experiments were conducted with different sets of stimuli to select the targets to be used in the trust game. Across these pilot studies, we aimed at controlling that within each dimension, the categorization and the discrimination of the exemplars would be similar. Moreover, we also aimed at having a baseline of the potential differences in categorization and discrimination between the dimensions. More details on stimuli selections and analyses are available in Appendix I.

We selected 32 photographs of men and women from the Karolinska Direct Emotional Faces (Lundqvist, Flykt, & Öhman, 1998) for the group playing with the social dimension (i.e., humans), 32 pictures of Lunaris and Taiyos ceded by the Object Perception Lab of Vanderbilt University for the group playing with social-like targets (i.e., artificial races), and 32 images of paintings from Wassily Kandinsky and Jaison Cianelli for the group playing with the non-social stimuli.

With these stimuli, we observed that there was no difference between categories on categorization accuracy within any dimension, larger $F(1, 19) = 1.65$, $p = .215$, $\eta_p^2 = .08$, and memory accuracy for the exemplars from these categories, larger $F(1, 19) = 2.90$, $p = .105$, $\eta_p^2 = .13$, both for participants playing with paintings. We also observed that the categorization of humans was less accurate than the categorization of the two other sets of stimuli, $F(2, 57) = 5.65$, $p = .006$, $\eta_p^2 = .17$. In the same vein, the memory of human targets was more accurate than in the other two groups, larger $p = .002$, and memory of paintings was better than the memory of artificial races, $p < .001$, $F(2, 57) = 117.65$, $p < .001$, $\eta_p^2 = .81$, as shown in Figure 5.1. Because of the pre-existing differences between the dimensions manipulate, we decided to explore in the subsequent trust game experiments the specific pattern of investments within each experiment groups, instead of directly comparing them, and to consider the results of these trust game in light of these differences.

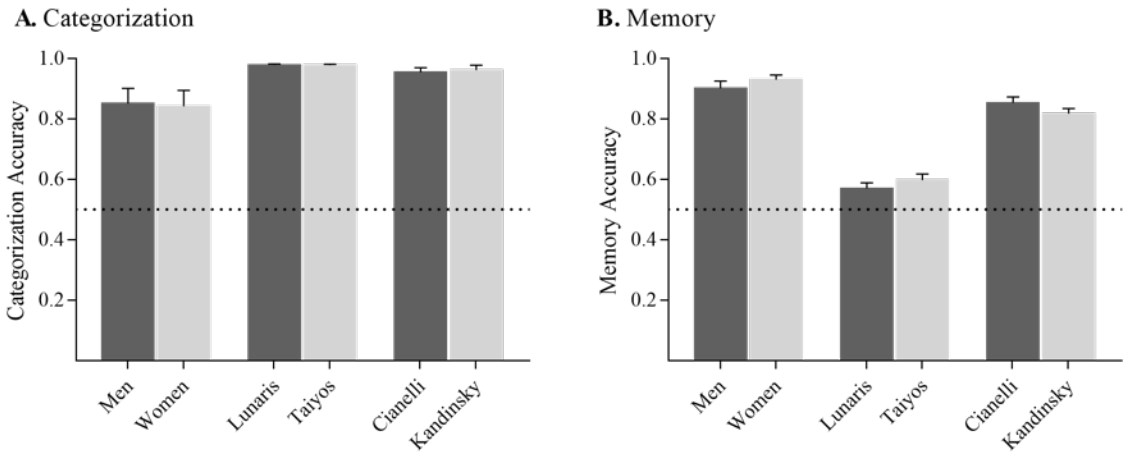

Figure 5.1. Accuracy scores in the categorization task and memory test for humans, artificial races and paintings.

Materials and Procedure. The stimuli described in the section above were used to represent the targets in the trust game, which was framed differently in the three experimental groups while maintaining the same structure. The critical differences between the group was implemented in the instructions participants received at the beginning of the experiment, but did not alter the sequence of trials. Each trial started with the euro symbol "€" (1.43º x 1.63º) for 190 ms representing that participants virtually received €1, followed by a fixation point for 500 ms in the center of the screen. Next, the picture of the target of this trial appeared for 1500 ms (5.68º x 7.77º) and participants had to decide whether or not to invest €1 in this trial by pressing '1' to invest or '0' not to invest. In case they did not invest the €1, they would keep the money to themselves and move to the next trial. The critical change between the three experimental groups regarded the case in which participants decided to invest the €1. Specifically, participants playing with humans and artificial races were told that their partners would receive €5 and in turn decide either to reciprocate giving back €2.50 to the participant, or to keep the whole money for themselves, thus the participant receiving €0, and therefore losing the invested

€1. Alternatively, partners playing with paintings were told that investing the €1 allowed them to enter into the art market which would reveal the real value of the painting up to €5. If the painting were of a high value, participants would earn the benefit of their investment shared with an art agency (i.e., €2.50). If the painting were of a low value, they would lose the invested €1. This way, the structure of the trust game was equally coherent and convincing for participants playing with social and non-social stimuli. After making their decision, participants received visual feedback on their final outcomes during 1000 ms. The task consisted of a trust game divided in 3 phases followed by a memory test.

First, in the baseline phase, participants played with 16 targets from each group (i.e., either 16 men and 16 women, 16 Lunaris and 16 Taiyos, or 16 Cianelli's and 16 Kandinsky's paintings), all of them being worth the investment in half of the trials, and not worth investing in the other half. This manipulation allowed to verify whether participants were biased towards one of the groups investing more money with one of the group before learning.

Next, in a learning phase, the two categories were associated with different outcomes. One category (e.g., men) was worth investing in and allowed to earn money on 75% of the trials while the other category (e.g., women) was not worth investing in and made participants lose money on 75% of the trials. Moreover, within each group 25% of the exemplars were inconsistent, that is, they were associated with a pattern of outcomes opposite to the pattern displayed by the group. For instance, if men were worth investing in and women were not, 12 men would be consistent and cooperate on 75% of the trials while the 4 other men would be inconsistent and cooperate only in 25% of the trials. In contrast, the 12 consistent women would cooperate on only 25% of the trials, and the 4 inconsistent women would cooperate on 75% of the trials, as shown in Figure 5.2. The associations between categories and specific outcomes were counterbalanced across participants.

Finally, in the transfer phase, participants were presented with 32 new targets with whom they had not interacted, all of them being worth investing in on half of the trials and not on the other half. This phase allowed to examine whether the learning of the category worthiness during the learning phase would be transferred to new targets from the same categories.

After the trust game participants completed a memory test in which they were sequentially presented with all 32 targets from the learning phase and were asked to indicate whether or not this exemplar was worth investing in during the trust game by pressing '1' if they remembered they were, '0' if they remembered they were not, or the spacebar if they did not remember the stimulus at all. The memory consisted of 32 trials. The experiment lasted around 50 minutes.

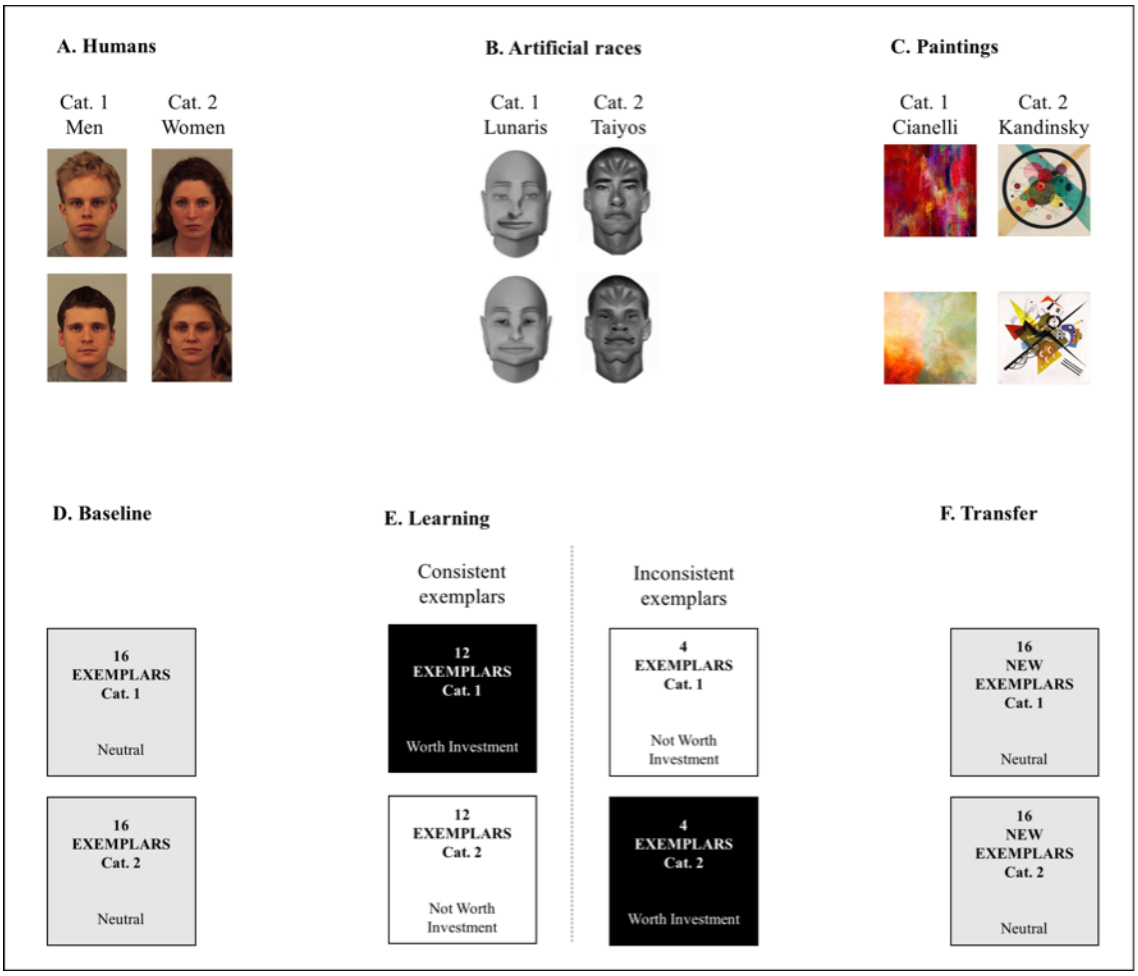

Figure 5.2. Example of stimuli used in the groups playing with humans (A), artificial races (B) and paintings (C), and procedure employed in the baseline (D), learning (E) and transfer (F) phases. Exemplars are represented in black when they are worth investing in (i.e., they allow participants to earn economic outcomes on 75% of the trials), in white when they are not worth (i.e., they make participants lose money on 75% of the trials) and in grey when they are associated to positive economic outcomes on half of the trials, and to economic loss on the other half.

Results

Baseline. To analyze spontaneous investments before learning, investment rates were subjected to repeated-measures ANOVAs separately for the three experimental groups (i.e., humans, artificial races and paintings) with category as a within-participants variable.

In the group playing with men and women, the main effect of category was significant, $F(1, 39) = 7.24$, $p = .010$, $= \eta^2_p$ .16, indicating that participants invested more with women (M = .70, SD = .15) than with men (M = .61, SD = .16). In contrast, participants did not significantly differ in their investment between Lunaris (M = .62, SD = .16) and Taiyos (M = .58, SD = .16), $F(1, 39) = 1.27$, $p = .266$, = η_p^2 .03. Neither did they between Cianelli (M = .60, SD = .14) and Kandinsky (M = .60, SD = .21) paintings, $F(1, 39) < .01$, $p = .989$, $\eta_p^2 < .01$.

Learning. To analyze whether individuation or categorization strategies impacted learning, we computed a learning index by subtracting participants' investment rates with partners that were worth investing in from their investment rate with partners that were not. This learning index was subjected to separate 2 (Consistency: consistent vs. inconsistent) x 4 (Block: 2-5) repeated-measures ANOVA for each experimental group. With this analysis, the effects of consistency and block in are informative of participants' learning strategy. If participants perfectly individuate targets, there should be no effect of the consistency variable (and therefore, the Consistency x Block interaction should not be significant), indicating similar for consistent and inconsistent targets. However, a significant effect of consistency and its variation across blocks, would indicate that learning is impacted by categorical information. If learning is reduced for inconsistent as compared to consistent targets, this would mean that an individuation strategy was attempted but not successful. In contrast, if the patterns of learning are in opposite directions for consistent and inconsistent targets, then learning would be mostly impacted by categorization processes.

In the group playing with humans, the main effect of consistency was significant, $F(1, 39) = 40.31$, $p < .001$, $\eta_p^2 = .51$, indicating larger learning about consistent (M = .27, SD = .17) than inconsistent (M = -.01, SD = .20) partners. The main effect of block was also significant, $F(3, 117) = 12.23$, $p < .001$, $\eta_p^2 = .24$, revealing that learning linearly increased from Block 2 (M = .04, SD = .10) to Block 5 (M = .17, SD = .20), $F(1, 39) = 18.57$, $p < .001$, $\eta_p^2 = .32$,. This learning was also qualified by a significant quadratic component, $F(1, 39) = 8.51$, $p = .006$, $\eta_p^2 = .18$, , indicating that this learning reached an asymptote as shown in Figure 5.3. The Consistency x Block interactions was not significant, $F(3, 117) = .35$, $p = .792$, $\eta_p^2 < .01$, suggesting that the increase in learning across blocks was similar for consistent and inconsistent partners.

In the group playing with artificial races[5], the main effect of consistency was significant, $F(1, 39) = 38.92$, $p < .001$, $\eta_p^2 = .50$, and qualified by block, $F(3, 117) = 6.06$, $p = .001$, $\eta_p^2 = .14$. For consistent partners, the main effect of block was significant, $F(3, 117) = 9.61$, $p < .001$, $\eta_p^2 = .20$, indicating that learning linearly increased from Block 2 ($M = .13$, $SD = .17$) to Block 5 ($M = .29$, $SD = .25$), $F(1, 39) = 18.45$, $p < .001$, $\eta_p^2 = .32$. The quadratic component was also significant, $F(1, 39) = 4.33$, $p < .044$, $\eta_p^2 = .10$, indicating that this learning reached an asymptote as shown in Figure 5.3. For inconsistent partners, the main effect of block was not significant, $F(3, 117) = 1.87$, $p < .138$, $\eta_p^2 = .05$, suggesting that in contrast to the group playing with humans, participants playing with artificial races did not learn across blocks about inconsistent partners.

In the group playing with paintings, we also found a significant effect of block, $F(1, 39) = 44.04$, $p < .001$, $\eta_p^2 = .53$, and a significant Consistency x Block interaction, $F(3, 117) = 4.03$, $p = .009$, $\eta_p^2 = .09$. For consistent exemplars, the main effect of block was significant, $F(3, 117) = 13.04$, $p < .001$, $\eta_p^2 = .25$, indicating that learning linearly increased from Block 2 ($M = .20$, $SD = .25$) to Block 5 ($M = .40$, $SD = .29$), $F(1, 39) = 19.94$, $p < .001$, $\eta_p^2 = .34$. In contrast, for inconsistent exemplars, the main effect of block was not significant, $F(3, 117) = .36$, $p = .781$, $\eta_p^2 < .01$, suggesting that just as the group playing with artificial races, participants playing with paintings did not learn across blocks about inconsistent exemplars.

[5] Note that in this and the following experiment, we observed a Consistency x Worth Investing in Category interaction in the group playing with artificial races, smaller $F(1, 38) = 5.88$, $p = .020$, $\eta_p^2 = .13$ in Experiment 1, indicating that the effect of consistency was larger when Taiyos were worth investing in than when Lunaris were. However, because the specific differences between Lunaris and Taiyos go beyond the scope of this paper, this interaction is not further analyzed.

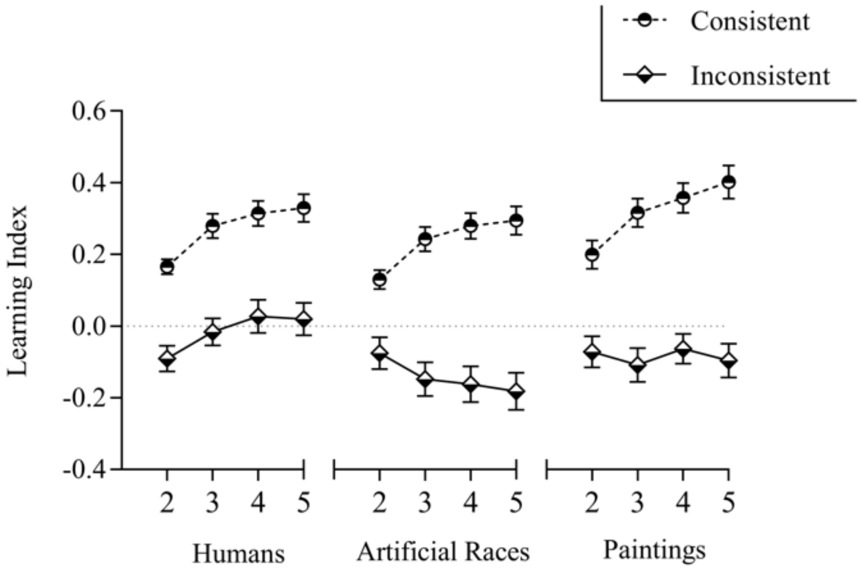

Figure 5.3. Learning about consistent and inconsistent exemplars across blocks in each experimental group in Experiment 1. Error bars represent the standard error of the mean.

Transfer. To analyze whether the associations between categories and worthiness in the leaning phase impacted investments with new exemplars from these categories, investment rates in the transfer phase were subjected to a mixed-design ANOVA with worth investing in category (men vs. women, or Lunaris vs. Taiyos or Cianelli vs. Kandinsky) as a between-participants factor and target category (men vs. women, or Lunaris vs. Taiyos or Cianelli vs. Kandinsky) as a within-participants variable, separately for each experimental group.

In the group playing with humans, the main effect of Target Category was significant, $F(1, 38) = 8.31$, $p = .006$, $\eta_p^2 = .18$, indicating that participants invested more with female ($M = .67$, $SD = .15$) than with male ($M = .59$, $SD = .19$) partners. The critical Target Category x Worth Investing in Category interaction was not significant, $F(1, 38) = 2.25$, $p = .142$, $\eta_p^2 = .06$, suggesting participants did not use their knowledge from the learning phase to categorize new humans, but instead used the same heuristics that guided their decisions in the baseline.

In the group playing with artificial races, we found a significant Target x Worth Investing in Category interaction, $F(1, 38) = 16.74$, $p < .001$, $\eta_p^2 = .31$. When in the learning phase, Lunaris were the group worth investing in, participants invested more with new Lunaris ($M = .71$, $SD = .20$) than with new Taiyos ($M = .53$, $SD = .26$) in the transfer phase, $F(1, 38) = 9.31$, $p = .007$, $\eta_p^2 = .33$. In contrast, when Taiyos were the group worth investing in, participants invested more with new Taiyos ($M = .67$, $SD = .22$)

than with new Lunaris (M = .50, SD = .29) in the transfer phase, $F(1, 38) = 7.55$, $p = .013$, $\eta_p^2 = .28$.

In the group playing with paintings, the Target x Worth Investing in Category interaction was also significant, $F(1, 38) = 22.70$, $p < .001$, $\eta_p^2 = .37$, just as in the group playing with artificial races. When in the learning phase Cianelli's painting were worth investing in, participants invested more on new Cianelli's (M = .72, SD = .14) than on new Kandinsky's (M = .48, SD = .24) paintings in the transfer phase, $F(1, 38) = 17.24$, $p = .001$, $\eta_p^2 = .48$. Conversely, when participants learned that Kandinsky's paintings were worth investing in, they invested more on new Kandinsky's (M = .70, SD = .21) than on new Cianelli's (M = .54, SD = .22) paintings in the transfer phase, $F(1,38) = 6.77$, $p = .018$, $\eta_p^2 = .26$.

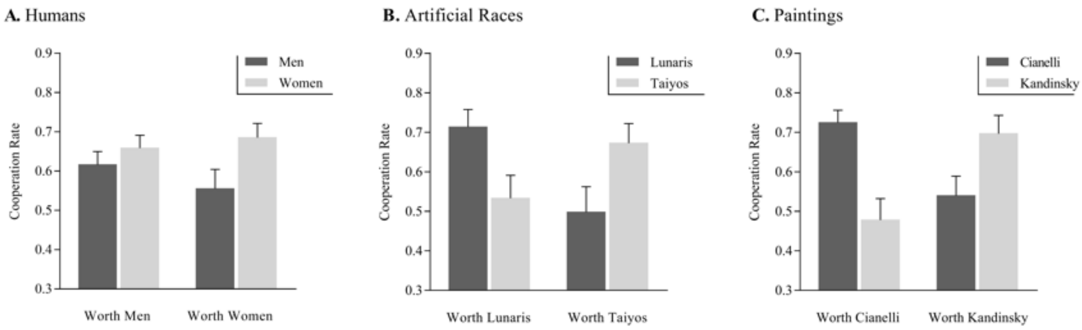

Figure 5.4. Cooperation rate with new targets in the transfer phase according to the category worth the investment in the learning phase for the groups playing with (A) humans, (B) artificial races, and (C) paintings in Experiment 1. Error bars represent the standard error of the mean.

Memory. To analyze participants' recall of the outcomes associated with the exemplars presented during the learning phase, we subjected recall scores (i.e., participants' accuracy rate at recalling the outcome associated with each exemplar) to separate repeated-measures ANOVAs with worthiness (worthy vs. unworthy investment) and consistency (consistent vs. inconsistent) as within-participants variables, in each experimental group.

In the group playing with humans, the main effect of consistency was significant, $F(1, 39) = 33.80$, $p < .001$, $\eta_p^2 = .46$, indicating that participants recalled better consistent (M = .75, SD = .12) than inconsistent partners (M = .56, SD = .20).

The same pattern of data was observed in the group playing with artificial races as the main effect of consistency was significant, $F(1, 39) = 35.84$, $p < .001$, $\eta_p^2 = .48$, again indicating that participants recalled better consistent (*M* = .69, *SD* = .16) than inconsistent (*M* = .38, *SD* = .22) partners.

In the group playing with paintings, the main effect of consistency was also significant, $F(1, 39) = 56.80$, $p < .001$, $\eta_p^2 = .59$, and qualified by the worthiness variable, $F(1, 39) = 4.30$, $p = .045$, $\eta_p^2 = .10$. For both the worth investing in, $F(1, 39) = 17.92$, $p = .001$, $\eta_p^2 = .32$, and the non-worth investing in, $F(1, 39) = 47.83$, $p = .001$, $\eta_p^2 = .55$, categories, participants recalled better consistent than inconsistent exemplars, although the difference was larger for the non-worth investing in (*M* = .80, *SD* = .17, vs. *M* = .43, *SD* = .27) than the worth investing in category (*M* = .76, *SD* = .18 vs. *M* = .51, *SD* = .29).

Discussion

The aim of Experiment 1 was to explore how categorization and individuation strategies impacted participants' investment decisions with social, social-like, and non-social stimuli. Specifically, we aimed at testing the hypotheses that social stimuli would trigger an individuation strategy, while non-social stimuli would rather trigger a categorization strategy.

Overall, participants' learning was largely impaired with inconsistent individuals, suggesting that category-related information was better acquired than individual attributes in all experimental groups. This interpretation was also supported in the memory test, in which consistent individuals were better recognized than inconsistent ones in the three experimental groups. However, in line with our predictions, learning about inconsistent humans increased across blocks, reflecting some motivation to learn in an individual fashion, while such increase in learning was not observed with inconsistent paintings or artificial races. Therefore, as one would expect, playing with a rather large number of partners made individuation complicated (Telga, Cantiani, & Lupiáñez, 2019), but only the group playing with humans was motivated enough to engage the cognitive resources necessary to learn across blocks about targets deviating from the group, especially in the learning phase in which they were outcome-dependent on their performance (Neuberg & Fiske, 1987).

Interestingly, at zero acquaintance, participants relied on categorical information in both baseline and transfer phases when playing with humans, using gender stereotypes to cooperate more with women than with women. These decisions were therefore based on previous knowledge acquired outside the lab, and were maintained even after we introduced specific associations between targets and economic outcomes in the learning phase. However, consistently with our predictions, participants did not differ in their investments with artificial races and paintings in the baseline, as they did not have any prior knowledge about the categories used; conversely, they immediately used the knowledge acquired during the learning phase to categorize new targets in the transfer phase. In this sense, it is interesting to note that the associations established in the learning phase did not overcome participants' prior knowledge when playing with humans, but were sufficient to guide predictions about new targets when they were not social.

Overall, the pattern of data observed is consistent with the prediction that social stimuli motivate the individuation strategy to a greater extent than non-social stimuli, in both the learning and the transfer phases. However, it is noteworthy that similar patterns were observed with social-like and non-social stimuli. One may argue that both artificial races and paintings have been perceived as non-social to the same extent, and that despite artificial races shared facial features with participants, they were not sufficient for participants to personally get involved in interactions with these targets, being overall considered as mere objects just as paintings. However, and interestingly, it is also possible that participants were more motivated to individuate artificial races than paintings as predicted, but because they have a poorer capacity of discrimination of those targets (see Figure 5.1), their learning about artificial races were similar to learning about paintings that were easier to discriminate. In other words, similar learning about paintings and artificial races may be explained by different factors: a higher motivation but a poorer capacity of discrimination for artificial races, and a smaller motivation but a higher capacity of discrimination for paintings. This interpretation is consistent with the fact that in the transfer phase, although new targets were categorized in both the social-like and non-social dimensions, effect sizes of categorization were larger in the group playing with paintings than in the group playing with artificial races. We decided to explore this explanation in Experiment 2 by using the memory performance as an indicator of participants' capacity of discrimination between targets, and exploring their pattern of investment in the transfer phase while controlling their capacity of discrimination.

Experiment 2

The goal of Experiment 2 was to replicate Experiment 1 and to verify whether differences in investment decisions were explained by motivational factors, by controlling the results by participants' capacity of discrimination. To achieve this goal, we used the same procedure as in Experiment 2, but introduced the memory phase after the learning phase and before the transfer phase. With this change in the procedure, we could use the score in the memory test as a co-variate when analyzing investments in the transfer phase. We predicted that even after controlling by participants' capacity of discrimination, participants would show a larger categorization for paintings than artificial races, using the associations established in the learning phase to predict the outcomes associated with each category.

Method

Participants. One hundred twenty volunteers (6 men) took part in the study (mean age: 22.28, range: 18-45) and received economical rewards proportional to their accuracy in the task (€5.85 on average). As we used the same sample as in Experiment 1, the outcome of the sensitivity analyses remains at $f = .23$.

Material, stimuli and procedure. The baseline and learning phases were identical to Experiment 1 for the three experimental groups. Notably, in Experiment 2, participants performed the memory test before the transfer phase. After the baseline and the learning phases, participants performed a 5-minute distraction task consisting of identifying the letter "Q" among arrays of letters "O". Next, they performed the memory test in which they were presented with all the 32 stimuli from the learning phase, and 32 new stimuli with whom they had no prior experience (16 from each category). Participants' task was to indicate whether or not they had already been presented with those stimuli during the learning phase by pressing the '1' key if they recognized them, or the '0' key if they did not. With this new manipulation, we aimed at exploring participants' ability to discriminate individual targets from the categories used in the trust game. In case they indicated that they had been presented with a target stimulus, they were asked to indicated whether or not it was worth investing in by pressing the '1' key if they thought it was, and the '0' if they thought it was not, allowing to analyze their recall of the stimuli presented in the trust game, just as in Experiment 2. After the memory test, participants realized the transfer phase as in Experiment 2. The 32 fillers used in the memory test were presented as new individuals in the transfer phase. The experiment lasted around 60 minutes.

Results

Baseline. Investment rates in the baseline were subjected to the same analyses as in Experiment 1, separately for each experimental group. Again, participants invested more with women (M = .70, SD = .12) than with men (M = .64, SD = .15), F(1, 39) = 6.34, p = .016, η_p^2 = .14. In contrast, participants did not significantly differ in their investment rates between Lunaris (M = .63, SD = .15) and Taiyos (M = .64, SD = .15), F(1, 39) < .01, p = .769, η_p^2 < .01. Neither did they between paintings from Cianelli (M = .57, SD = .19) and Kandinsky (M = .64, SD = .18), F(1, 39) = 2.97, p = .093, η_p^2 = .07.

Learning. Learning indexes were subjected to separate repeated-measures ANOVAs with consistency (consistent vs. inconsistent), and blocks (2-4) as within-participants variables, for each experimental group.

In the group playing with humans, the main effect of consistency was significant, F(1, 39) = 32.85, p < .001, η_p^2 = .46, revealing larger learning for consistent (M = .28, SD = .21) as compared to inconsistent (M = .00, SD = .22) partners. The main effect of block was also significant, F(3, 117) = 27.03, p < .001, η_p^2 = .41, indicating that learning linearly increased from Block 2 (M = .02, SD = .11) to Block 5 (M = .26, SD = .23), F(1, 39) = 49.27, p < .001, η_p^2 = .56. The Consistency x Block interaction was only marginal, F(3, 117) = 2.51, p = .062, η_p^2 = .06.

In the group playing with artificial races, the main effect of consistency was significant, F(1, 39) = 38.44, p < .001, η_p^2 = .50, as well as the Consistency x Block interaction, F(3, 117) = 10.54, p < .001, η_p^2 = .21. For consistent partners, the significant main effect of block, F(3, 117) = 14.93, p < .001, η_p^2 = .28, indicated that participants' learning linearly increased from Block 2 (M = .15, SD = .18) to Block 5 (M = .38, SD = .31), F(1, 39) = 22.35, p < .001, η_p^2 = .36, until reaching an asymptote, as suggested by the significant quadratic trend, F(1, 39) = 5.42, p = .025, η_p^2 = .12, shown in Figure 5.5. For inconsistent partners, the main effect of block was also significant, F(3, 117) = 4.15, p = .008, η_p^2 = .10, but indicated a linear decrease in learning from Block 2 (M = -.05, SD = .26) to Block 5 (M = -.21, SD = .40), F(1, 39) = 7.84, p = .008, η_p^2 = .17.

In the group playing with paintings, the main effect of consistency was significant, F(1, 39) = 34.92, p < .001, η_p^2 = .47, and qualified by the Consistency x Block interaction, F(3, 117) = 5.01, p = .001, η_p^2 = .11. For consistent exemplars, the main effect of block was significant, F(3, 117) = 19.15, p = .001, η_p^2 = .33, as learning linearly increased from Block 2 (M = .21, SD = .28) to Block 5 (M = .49, SD = .32), F(1, 39) = 23.12, p < .001, η_p^2 = .37, until reaching an asymptote, as endorsed by the quadratic trend, F(1, 39) = 10.89, p = .002, η_p^2 = .22. In contrast, for inconsistent exemplars, the main effect of block was not significant, $F(3, 117) = 1.03$, $p = .381$, $\eta_p^2 = .03$, suggesting that participants did not learn across block about inconsistent paintings.

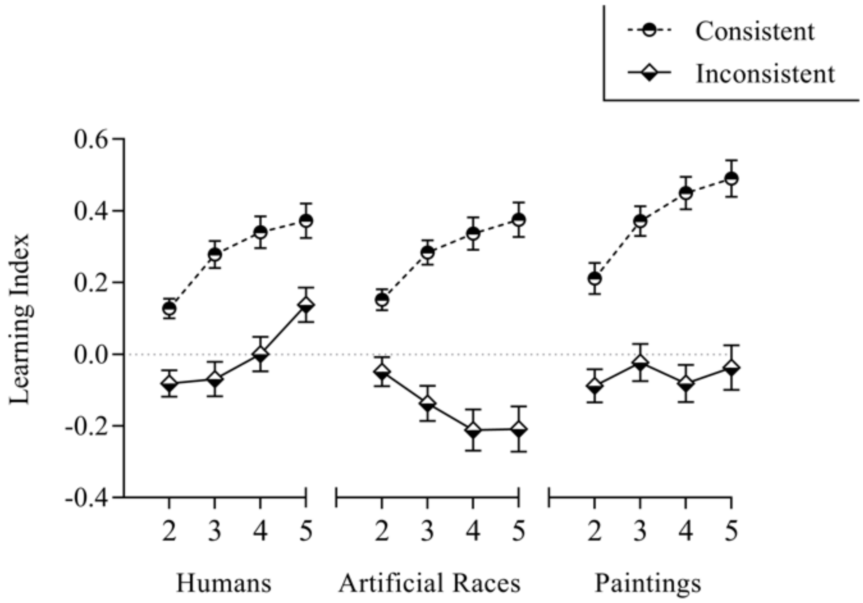

Figure 5.5. Learning about consistent and inconsistent exemplars across blocks in each experimental group in Experiment 2. The error bars represent the standard error of the mean.

Memory.

Recognition. Separate one-sample T-tests were conducted on accuracy in the recognition phase with new exemplars (i.e., correct identification of new exemplars as such) in each experimental group, to verify that participants correctly identify new exemplars above chance. We observed that participants' accuracy was significantly above chance in the groups playing with humans, $t(39) = 68.85$, $p < .001$, and paintings, t(39) = 29.20, $p < .001$, but not in the group playing with artificial races, $t(39) = 1.61$, $p = .114$.

Furthermore, to analyze recognition of the targets, separate repeated-measures ANOVAs were conducted on recognition scores with worthiness (worthy vs. not unworthy investment) and consistency (consistent vs. inconsistent) as within-participants variables in each experimental group. In the group playing with humans, the Worthiness x Consistency interaction was significant, $F(1, 39) = 5.01$, $p = .031$, $\eta_p^2 = .11$. Although the pattern of data indicated that participants seemed to recognize better partners who individually reciprocated, (i.e., consistent partners in the group worth investing in and inconsistent partners in the group that was not) the post hoc comparisons of recognition scores for consistent vs. inconsistent partners were not significant, largest $F(1, 39) = 2.09$, $p = .157$, $\eta_p^2 = .05$, in the category that was not worth investing in.

A similar pattern was observed in the group playing with artificial races. The Worthiness x Consistency interaction was significant, $F(1, 39) = 6.52$, $p = .015$, $\eta_p^2 = .14$. Again, participants seemed to recognize better partners that were individually worth investing in, although the post hoc comparisons of recognition scores between consistent and inconsistent partners were not significant, largest $F(1, 39) = 3.49$, $p = .069$, $\eta_p^2 = .08$, in the category worth investing in.

Again, in the group playing with paintings, a similar pattern was found. The significant Worthiness x Consistency interaction, $F(1, 39) = 15.09$, $p < .001$, $\eta_p^2 = .28$, indicated that higher recognition scores for exemplars that were individually worth investing in. Within the category worth investing in, recognition scores were higher for consistent ($M = .93$, $SD = .11$) than for inconsistent ($M = .84$, $SD = .21$) paintings, $F(1, 39) = 10.38$, $p = .003$, $\eta_p^2 = .21$. Conversely, within the category that was not worth investing in, recognition scores were higher for inconsistent ($M = .89$, $SD = .20$) than for consistent ($M = .83$, $SD = .14$) exemplars, $F(1, 39) = 4.20$, $p = .047$, $\eta_p^2 = .10$.

Recall. A different repeated-measures ANOVA was conducted on recall scores with worthiness (worthy vs. unworthy investment) and consistency (consistent vs. inconsistent) as within-participants variables in each experimental group.

In the group playing with humans, we found a significant main effect of consistency, $F(1, 37) = 24.77$, $p < .001$, $\eta_p^2 = .40$, indicating that participants recalled better the cooperative behaviors of consistent ($M = .78$, $SD = .13$) than inconsistent ($M = .61$, $SD = .20$) partners.

In the group playing with artificial races, the main effect of consistency was also significant, $F(1, 34) = 32.14$, $p < .001$, $\eta_p^2 = .49$, and qualified by the worthiness variable, $F(1, 34) = 31.14$, $p < .035$, $\eta_p^2 = .12$. Participants recalled better consistent than inconsistent partners both within the category worth investing in, $F(1, 34) = 11.87$, $p < .001$, $\eta_p^2 = .24$, and the category not worth investing in, $F(1, 34) = 47.51$, $p < .001$, $\eta_p^2 = .58$, although the difference between consistent and inconsistent was larger in the non-worth investing in category ($M = .77$, $SD = .23$ vs. $M = .28$, $SD = .31$) than in the worth investing in category ($M = .74$, $SD = .21$ vs. $M = .44$, $SD = .40$).

A similar pattern of data was observed in the group playing with paintings. The main effect of consistency was significant, $F(1, 39) = 14.03$, $p < .001$, $\eta_p^2 = .27$, and qualified by the worthiness variable, two-way interaction, $F(1, 39) = 7.93$, $p = .008$, $\eta_p^2 =$

.17. Participants significantly recalled better consistent (M = .82, SD = .24) than inconsistent (M = .51, SD = .31) paintings within the category that was not worth investing in, $F(1, 39) = 22.71$, $p < .001$, $\eta_p^2 = .37$. However, in the category worth investing in, participants showed similar recall of consistent (M = .72, SD = .19) and inconsistent (M = .61, SD = .34) paintings, $F(1, 39) = 2.67$, $p = .110$, $\eta_p^2 = .06$.

Transfer phase. Investment rates in the transfer phase were subjected to mixed-design ANOVAs with worth investing group (men vs. women, or Lunaris vs. Taiyos or Cianelli vs. Kandinsky) as a between-participants factor and category (men vs. women, or Lunaris vs. Taiyos or Cianelli vs. Kandinsky) as a within-participants variable, separately for each experimental group.

As shown in Figure 5.6, in the group playing with humans, the main effect of target was significant, indicating that participants invested more with women (M = .66, SD = .17) than with men (M = .58, SD = .19), as observed in Experiment 1, $F(1, 38) = 7.67$, $p = .009$, $\eta_p^2 = .17$. Moreover, the critical Target x Worth Investing in Category interaction was significant, $F(1, 38) = 34.92$, $p < .001$, $\eta_p^2 = .47$. When in the learning women were the group worth investing in, participants invested more with new women (M = .69, SD = .16) than with new men (M = .51, SD = .19) in the transfer phase, $F(1, 19) = 11.61$, $p = .003$, $\eta_p^2 = .38$. However, when in the learning phase men were the group worth investing in, participants did not significantly differ in their investment between men (M = .63, SD = .18) and women (M = .64, SD = .17), $F(1, 19) = .26$, $p = .614$, $\eta_p^2 = .01$.

In the group playing with artificial races, the Target x Worth Investing in Category interaction was significant, $F(1, 38) = 16.20$, $p < .001$, $\eta_p^2 = .30$. When in the learning phase Lunaris were the group worth investing in, participants invested more with new Lunaris (M = .63, SD = .19) than with new Taiyos (M = .45, SD = .25) in the transfer phase, $F(1, 19) = 7.58$, $p = .01$, $\eta_p^2 = .29$. When in the learning phase Taiyos were the group worth investing in, participants invested more with new Taiyos (M = .74, SD = .18) than with new Lunaris (M = .55, SD = .24) in the transfer phase, $F(1, 19) = 8.64$, $p = .008$, $\eta_p^2 = .31$.

A similar pattern was observed in the group playing with paintings. The Target x Worth investing category interaction was significant, $F(1, 38) = 12.24$, $p = .002$, $\eta_p^2 = .23$. When in the leanring phase Cianelli's painting were worth investing in, participants invested more with new Cianelli's (M = .63, SD = .23) than with new Kandinsky's (M =

.43, *SD* = .24) paintings in the transfer phase, $F(1, 19) = 5.41$, $p = .031$, $\eta_p^2 = .22$. Conversely, when in the learning phase Kandinsky's paintings were worth investing in, participants invested more with Kandinsky's (*M* = .63, *SD* = .23) than with Cianelli's (*M* = .42, *SD* = .25) paintings in the transfer phase, $F(1, 19) = 5.83$, $p = .026$, $\eta_p^2 = .24$.

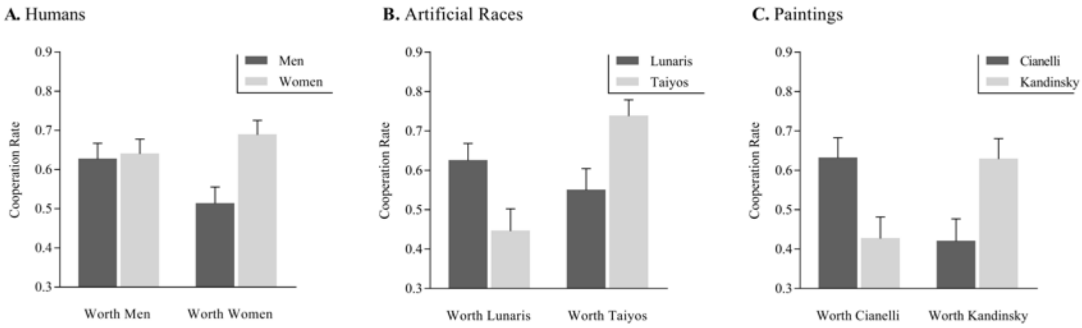

Figure 5.6. Cooperation rate with new targets in the transfer phase according to the category was worth the investment in the learning phase for the groups playing with A) humans, B) artificial races, and C) paintings, in Experiment 2. Error bars represent the standard error of the mean.

To directly test the differences between the three experimental groups (and subsequently control these effects by participants' capacity of discrimination), we computed a transfer index by subtracting participants' investment with the group not worth investing in from their investment with the group worth investing in during the transfer phase. With this index, higher scores indicate more categorization, that is, more investment with the category that was worth investing in during the learning phase. This index was subjected to a univariate ANOVA with dimension as a between-participants factor. Although the categorization effect seemed to be smaller in the group playing with humans (*M* = .10, *SD* = .20) than in the group playing with artificial races (*M* = .18, *SD* = .28) or the group playing with paintings (*M* = .21, *SD* = .38), the expected main effect of dimension was not significant, $F(2, 117) = 1.53$, $p = .220$, $\eta_p^2 = .03$. After introducing the recognition score as a co-variate in the analyses, the between-group differences seemed slightly larger, although still not significant, $F(1, 19) = 1.70$, $p = .190$, $\eta_p^2 = .03$.

Discussion

In Experiment 2, we replicated Experiment 1 comparing investment decisions and learning about social vs. non-social stimuli. Overall, the data confirmed the larger weight of category-related against individual information in the three experimental groups, both during learning and in the memory test. However, we observed an individuation pattern only for humans, in line with our hypotheses. It is also noteworthy that during learning, artificial races seemed to have been categorized across blocks to a greater extent than paintings. This potential difference is further explored in the next section.

Regarding participants' first investment decisions, the data confirmed that with humans, participants used gender stereotypes to make decisions at zero acquaintance, in both baseline and transfer phases. In fact, in Experiment 2, participants cooperated more with women than with men in the baseline, and used category-based associations from the learning phase in the transfer phase only when they were congruent with gender stereotypes (i.e., when women were the group worth investing in). Moreover, as in Experiment 2, participants were not biased toward social-like and non-social categories in the baseline, but decisions made in the learning phase with only 16 exemplars from each category were sufficient to trigger categorization processes with paintings and artificial races in the transfer phase. Despite the direct between-group comparisons in the transfer phase revealed no difference between the three experimental groups, it is noteworthy that each group displays a different pattern of categorization in the transfer phase. Participants rely to a greater extent on previous knowledge when they possess it (i.e., when playing with humans) but use the associations created within the trust game to categorize new exemplars when they have no prior knowledge to rely on (i.e., when playing with paintings and artificial races).

Between-groups comparisons

To have a sense of the differences between the three experimental groups in the learning phase, we conducted a mixed design ANOVA on learning indexes from Experiments 1 and 2 with consistency (consistent vs. inconsistent) and block (2-4) as within-participants variables and experiment (Experiment 1 vs. Experiment 2) and Dimension (artificial races vs. humans vs. paintings) as between-participants factors. The experiment variable did not modulate, $F(6, 702) = .63$, $p = .709$, $\eta_p^2 < .01$, the significant Consistency x Block x Dimension interaction, $F(6, 702) = 3.66$, $p = .001$, $\eta_p^2 = .03$. Regarding learning about consistent exemplars, we observed a main effect of dimension, $F(2, 237) = 4.06$, $p = .019$, $\eta_p^2 = .03$, which was clearly not modulated by Block, $F(6, 711) = .37$, $p = .898$, $\eta_p^2 < .01$, as shown in Figure 5.7. Post hoc comparisons with Bonferroni correction revealed that participants learned more about consistent paintings than about consistent artificial races, $p = .026$, and slightly more about consistent paintings than about consistent humans, $p =$

.080, overall indicating a larger use of category-related information in the group playing with paintings.

More interestingly, learning about inconsistent exemplars clearly showed a very different pattern. Here the Dimension x Block interaction was significant, $F(6, 702) = 7.49$, $p < .001$, $\eta_p^2 = .06$. Post hoc Bonferroni comparisons revealed that in the last block of trials, when learning was well established, learning was significantly larger in the group playing with humans than in the group playing with paintings, $p = .022$, and in the group playing with artificial races, $p < .001$. However, the difference in learning between the group playing with paintings and the group playing with artificial races was rather ambiguous, $p = .051$. Bayesian analyses comparing learning about artificial races and learning about paintings suggest no evidence for the difference between the two groups $BF_{10} = 1.84$.

Overall, learning about inconsistent paintings $t(79) = 2.50$, $p = 014$, and inconsistent artificial races $t(79) = 5.16$, $p < .001$, were significantly below 0, indicating a significant application of categorical attributes to inconsistent non-social exemplars, which was not the case not with human targets $t(79) = -.38$, $p = .705$.

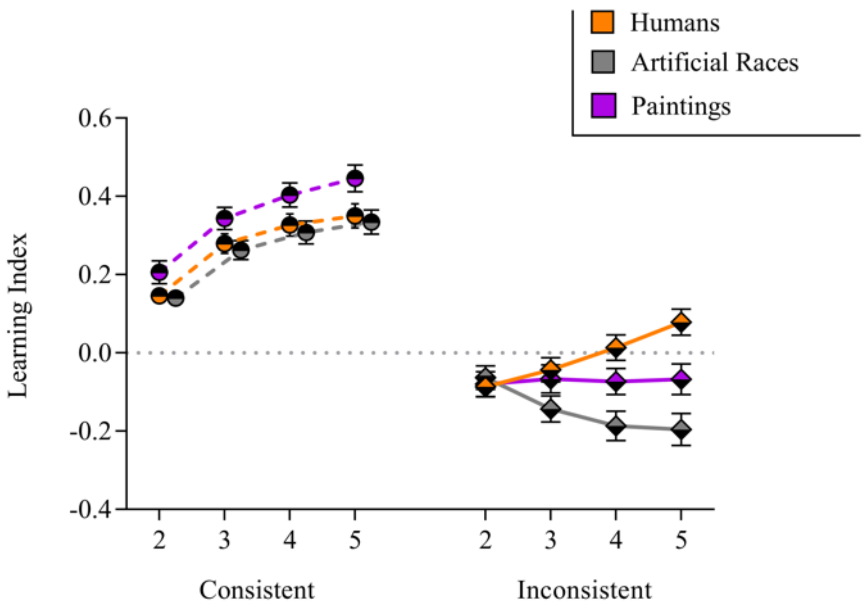

Figure 5.7. Learning about consistent and inconsistent exemplars grouping together Experiments 1 and 2.

General discussion

Across two experiments, we investigated how individuation and categorization strategies impact learning about social, social-like and non-social targets. Specifically, we adapted a multi-round trust game in which participants had to predict the economic outcomes associated with different targets.

Because the data from the pilot studies indicated that participants' capacity of discrimination and categorization were different for humans, artificial races and paintings, we decided to analyze specific patterns of learning within each dimension, and broadly observed that humans were individuated while paintings and artificial races were rather categorized. Importantly, we also compared the three experimental groups on learning about consistent and inconsistent targets. Taking into account the aforementioned between-group differences, the comparison between the three experimental groups is not informative on its own, but need to be considered in light of the data from the pilot studies. Notably, in the pilot studies, we directly measured participants' capacity to categorize and discriminate targets, and observed that artificial races were the best categorized group, followed by painting and humans. Conversely, humans were better individuated followed by paintings and artificial races. The poorer discrimination of artificial races measured in the pilot studies was later confirmed in the recognition test of Experiment 2, in which accuracy about new exemplars did not differ from chance.

In the trust game, however, we analyzed categorization and individuation processes in trait attributions, inferring from participants' investment decisions whether their behaviors were impacted by categorical or individual knowledge. This measure differs from the measures collected in the pilot studies, in that participants' decision was not solely led by perceptual expertise but also by motivation to use category- or individual-based information in the process of impression formation. Specifically, in the trust game, learning about consistent exemplars reflected the extent to which participants used categorical information, and learning about inconsistent individuals informed of participants' use of individuation processes, to form impressions of the targets. Considering these differences in the two types of measures, if we had found the exact same patterns of individuation and categorization in the pilot studies and in the trust game, we would have concluded that the differences observed in the trust game are hardly explained by the experimental manipulation of motivational factors and its impact of behavior. However, a few differences between the pilot studies and Experiment 1 and 2 granted some weight to motivational factors, and indicated that social, social-like and non-social stimuli are not equally processed.

First, learning about consistent targets suggest that participants made a greater use of categorical information for non-social (i.e., paintings) than social-like (i.e., artificial races) and, to some extent, social (i.e., humans) stimuli. The fact that categorical learning was larger for paintings than for artificial races is noteworthy. If participants had not been motivated to process artificial races in an individual fashion, they would have used up their capacity of categorization to make accurate investment for consistent exemplars on the basis of their categorical features. However, for artificial races, the categorical learning may have been also impacted by individual features (as for humans), resulting in a poorer learning about consistent artificial races than about consistent paintings, as participants took a greater advantage from categorical information regarding paintings than artificial races.

Second, learning about inconsistent targets revealed that humans were better individuated than the two other sets of stimuli, but the difference between artificial races and paintings was not clear. This pattern suggests that in spite of participants' greater ability at discriminating among exemplars of paintings, motivational factors may have encouraged more categorical processes with paintings, resulting in a performance similar to the one observed with artificial races, which were poorly discriminated. Alternatively, participants playing with artificial races may have been more motivated to individuate than their counterparts playing with paintings, such that they overcame their poor capacity of discriminability and achieve a pattern of learning similar to the one observed with non-social stimuli easier to discriminate. This interpretation is consistent with the argument that cognitive ability is not sufficient to elicit individuation, which highly depends on perceivers' motivation (Fiske & Neuberg, 1990; Neuberg & Fiske, 1987).

Overall, the fact that the three experimental groups were not evenly-matched on discrimination and categorization ability makes the interpretation of the data complex, and some methodological improvements may help to disentangle the role of motivation and cognitive ability in the differentiated use of individuation and categorization strategies. An ideal starting point would be to match the three experimental groups on discrimination and categorization ability prior to the trust game, which may be very complex for several reasons including the specificity of human faces as stimuli *per se* (Hugenberg & Wilson, 2013), but also the necessary varying degrees of stimuli familiarity across dimensions.

An alternative solution may be to measure participants' capacity of discrimination among exemplars prior to the trust game and to use this measure as a control variable when exploring investment in the learning and the transfer phase of the trust game. Instead of using the memory scores (which is likely influenced by learning strategies) as an indicator of participants' ability to discriminate exemplars, participants could realize the experiment in two sessions: the first one would be similar to the pilot studies (see Appendix I), providing an objective measure of participants' ability to categorize and discriminate stimuli, and the second one would consist of a trust game with the same design as in Experiments 1 and 2.

Finally, a potential explanation for the lack of differences between the three experimental groups is that participants were rewarded with real monetary outcomes, which may have increased their motivation in the three experimental groups (Neuberg & Fiske, 1987) to a level exceeding the specific motivational motives associated with the manipulated dimensions. This possibility may be explored in further research by adapting the paradigm to make rewards less (if not) salient.

Overall the current research gives some insight on the power of motivation in social learning. In a context in which individuation is rather complicated, the tendency to categorize targets may be overcome if they are relevant enough to us. The degree to which stimuli are perceived as social or non-social impacts the perceived relevance.

Acknowledgments and funding

We are tankful to all the volunteers who participated in the studies, and to the Object Lab Perception of Vanderbilt University for ceding part of the stimuli used in the present research.

This work was supported by the Spanish Ministry of Education, Culture and Sports, with pre-doctoral FPU fellowship FPU14/07106 to MT, and the Spanish Ministry of Economy and Competitiveness, with research projects PSI2014-52764-P and PSI2017-84926-P to JL. Note that JAA is now working at the University of Leicester. This research is part of MT thesis dissertation under the supervision of JL.

Correspondence concerning this article should be addressed to Maïka Telga, Departamento de Psicología Experimental, Universidad de Granada, Campus de Cartuja s/n, 18 071 Granada, Spain. Email: mtelga@ugr.es

APPENDIX I: STIMULI SELECTION

Different experiments were conducted to test the stimuli to be used as targets in the trust game. Importantly, we aimed at verifying that within each dimension (i.e., humans, artificial races, paintings), participants would be able to categorize the targets into the two categories of interest, and to discriminate among the exemplars of each category. Here, we report the results for the stimuli that were finally selected for the trust game.

Method

Participants. Across the pilot studies, different groups of participants evaluated different sets of stimuli. The data reported in the present research were collected in 3 different experiments, either testing stimuli for the human, the artificial races or the painting dimensions. Twenty volunteers participated in each one of these experiments in exchange for credit course. Written informed consents were obtained from all participants in this and the following experiments, following the guidelines approved by the local ethic committee (175/CEIH/2017).

Apparatus and stimuli. Stimuli were displayed on a 27’’ screen placed at 60 cm from participants. E-Prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2002) was used for stimuli presentation and data acquisition. For the social condition, we tested participants’ categorization based on gender groups: men and women. We used the pictures of 32 men and 32 women against a grey background with neutral facial expressions taken from The Karolinska Directed Emotional Faces (KDEF) (Lundqvist, Flykt, & Öhman, 1998). For the social-like condition, we verified that participants distinguished between two artificial races created in the lab with human-like attributes: Lunaris and Taiyos. In fact, research has shown that people can learn to individuate Lunaris and Taiyos and may process artificial races’ faces holistically, just as human faces (Chua, Richler, & Gauthier, 2014). Thirty-two pictures of Lunaris and 32 pictures of Taiyos with bottom and top variations were ceded by the Object Perception Lab of Vanderbilt University. Finally, for the non-social condition, we expected participants to distinguish paintings from two artists in a categorical way, as observed in previous studies (dos Santos Ferreira et al., 2018). In fact, paintings are fairly complex stimuli and allowed to build a credible cover story for the trust game, as described in Experiment 2. After various pilot studies testing paintings from different artists, we finally selected 32 paintings from Wassily Kandinsky and 32 paintings from Jaison Cianelli, with unequivocally different styles. While Kandinsky's paintings were characterized by the use of overlapping geometrical figures, Cianelli's paintings rather consisted of shapeless melted colors. Kandinsky's paintings were obtained via a Google search while Cianelli's paintings were downloaded from the artist's website. All the selected paintings were abstract to prevent participants from associating the contents of the paintings with social components. The pictures of humans and artificial races were displayed at 8.10º x 10.95º, with a total area of 88.70º. The paintings were resized such that their area were as close as possible to 88.70º, while maintaining their original shape.

Procedure. Participants performed the task with either humans (men and women), artificial races (Lunaris and Taiyos) or paintings (Cianelli's and Kandinsky's). First, they completed a categorization task in which they had to classify 32 stimuli (i.e., 16 exemplars from each category) into the categories of interest. Visual instructions about the categorization task included one example for each category without label or any explicit mention to the dimensions of categorization manipulated. Participants first performed 8 practice trials consisting of 2 presentations of 4 exemplars from each category (different from the 32 stimuli of the experimental block), in which they received visual feedback. Afterwards, they started the experimental phase consisting of 6 presentations of the 32 stimuli. Each trial started with a fixation cross displayed during 1 second followed by the picture of an exemplar from one of the categories manipulated during 1.5 seconds or until response. No feedback was provided during the experimental phase. The inter-trial interval was 1 second. Auto-administered breaks were allowed every 64 trials. Upon the completion of the categorization task, participants performed a memory test in which they were presented with the 32 stimuli from the categorization task and 32 new stimuli from the same categories. Each stimulus was displayed for 1.5 seconds and participants had to indicate whether or not the picture had been presented during the categorization phase. Overall, the experiment lasted around 35 minutes.

Results and discussion

To test whether participants equally categorize the exemplars from the two categories manipulated in each dimension, the accuracy scores in the categorization task were subjected to separate repeated-measures ANOVAs with the target (men vs. women, Lunari vs. Taiyo, or Cianelli vs. Kandinsky) as a within-participants variable. This analysis revealed that within each experiments, participants did not significantly differ in their categorization of the two targets of interest, larger $F(1, 19) = 1.65$, $p = .215$, $\eta_p^2 = .08$, in the experiment with paintings. Further, to verify whether participants categorize the targets above chance level, the same accuracy scores were subjected to a one-sample T-test, revealing that exemplars from all categories were categorized above chance, smaller $t(19) = 6.50$, $p < .001$, for men categorization. Overall, these data revealed that for each dimension, the targets were accurately categorized in the two categories of interest, and that the process of categorization was similar for both categories.

To test whether participants equally discriminated the targets within the categories of interest, accuracy scores from the memory test were subjected to a separate repeated-measures ANOVA with the target as within-participants variable. Again, in all experiments, participants did not significantly differ in their recognition of the targets from the two categories manipulated, larger $F(1, 19) = 2.90$, $p = .105$, $\eta_p^2 = .13$, for participants playing with paintings. We also verified that participants discriminated the target from all categories above chance. A one-sample T-test on accuracy in the memory test confirmed that targets from all categories were recognized above chance, smaller $t(19) = 3.67$, $p < .002$, for Lunaris' recognition. Overall, these data suggested that participants discriminated the targets accurately, and that discrimination was similar for the two categories of interest within each experimental group.

Between-experiments analyses were conducted with a multi-variate ANOVA on memory and categorization scores with dimension (humans vs. artificial races vs. paintings) as a between-participants factor, to have some sense of potential between-group differences in categorization and discrimination capacity. This analysis revealed a significant effect of category on memory scores, $F(2, 57) = 117.65$, $p < .001$, $\eta_p^2 = .81$. Post-hoc comparisons with Bonferroni corrections revealed that recognition of humans ($M = .92$, $SD = .08$) was significantly better than recognition of both paintings ($M = .84$, $SD = .06$), $p = .002$, and artificial races ($M = .58$, $SD = .07$), $p < .001$. Recognition of paintings was also significantly better than recognition of artificial races, $p < .001$. The effect of category was also significant regarding the categorization scores, $F(2, 57) = 5.65$, $p = .006$, $\eta_p^2 = .17$. Post hoc comparisons using Bonferroni corrections indicating that participants categorized humans ($M = .85$, $SD = .22$) to a lesser extent than both paintings ($M = .96$, $SD = .07$), $p = .031$, and artificial races ($M = .98$, $SD = .02$), $p = .009$.

CHAPTER 7. Emotion and Social Perception

EXPERIMENTAL SERIES 6

The content of this experimental series is in preparation as:

Telga, M., Kawakami, K., & Lupiáñez, J. (2019). *Individuation is favored over emotion categorization in social learning*. [Manuscript in preparation].

INDIVIDUATION IS FAVORED OVER EMOTION CATEGORIZATION IN SOCIAL LEARNING

Abstract

Individual identities and emotional expressions are two relevant cues to make predictions about others' behaviors. Across three experiments, we provided participants with the faces of individual targets who varied in emotional expressions in a trust game paradigm. Emotions were categorically manipulated, such that partners displaying the same expression (i.e., happy or angry) were associated with the same behavior, either cooperative or non-cooperative. We also manipulated that participants played either with 4 or 32 partners. To make accurate predictions about their partners, the most efficient strategy was for participants to attend to emotions. While a large body of research suggests categorization occur automatically and is the default process, the present research suggests an attentional bias to focus on individual identities can prevail in some situations, even when categorical processing and not individuation is functional to performing the task.

Keywords: social categorization, individuation, emotion, learning, trust

Impression formation processes are impacted by numerous social cues that we may extract from individuals' faces (Hugenberg & Wilson, 2013; Todorov, Said, Engell, & Oosterhof, 2008) to inform our judgements and determine how to interact with others. For instance, we may rely on features associated with trustworthiness or dominance to guide our perceptions (Oosterhof & Todorov, 2008). Alternatively, we may monitor an individual's behavior over time. If a person consistently acts in a particular manner, this may also impact impressions (King-Casas et al., 2005). For instance, if we notice that a person tends to be cooperative across different contexts, we may respond in turn by being cooperative.

Our impressions, however, we may not only be influenced by an individual's characteristics and behaviors but also to his or her category memberships. In the early stages of face processing, we attend to features related to social categories such as gender, race, or age (Hugenberg & Wilson, 2013; Hugenberg, Young, Bernstein, & Sacco, 2010; Stolier & Freeman, 2016). Although much of the research has focused on stable social groups, we may also process others according to more transient categories such as emotions (Bruce & Young, 1986), and these categories may also guide our perceptions and interactions (Ames & Johar, 2009; Tortosa, Strizhko, Capizzi, & Ruz, 2013). For example, if we assume that a happy person is related to more positive outcomes such as cooperation, we may be more likely to interact with that person in a reciprocal cooperative way. Just like with individuals, we may monitor category members' behaviors over time, and based on those behaviors also form impressions.

Notably, research has indicated that our initial strategy to form impressions of others is to categorize them (Cloutier, Mason, & Macrae, 2005; Macrae & Bodenhausen, 2000; Macrae & Bodenhausen, 2001). However, this inclination may be overturned if we are significantly motivated to understand the person and are able to process them as individuals (Fiske & Neuberg, 1990; Wilson & Brekke, 1994). One instance in which we may be motivated to individuate others is when we are outcome dependent on them (Neuberg & Fiske, 1987). Specifically, when our reliant on knowing on others for specific outcomes (e.g., money, promotions, etc.), we will focus on their individual attributes. However, we may only be successful in this process if we have sufficient cognitive resources (Gilbert & Hixon, 1991). When we have limited cognitive resources or when there is an abundance of information to process (e.g., if the number of targets to process is large), we may be reliant on categorical processing (Macrae, Milne, & Bodenhausen, 1994; Pendry & Macrae, 1999). Conversely, if we are motivated to individuate others and if there are available cognitive resources or the number of targets to process is small, we may be able to process them as individuals.

In the present research, we explored the use of individual identities and emotional expressions when predicting others' behaviors and examined whether emotions would be used to process people in categorical ways. Furthermore, in a context in which participants' outcome depends on knowing others and so are highly motivated to individuate them (Kawakami et al., 2014; Neuberg & Fiske, 1987; Telga, de Lemus, Cañadas, Rodríguez-Bailón, & Lupiáñez, 2018), we investigated whether the use of individual versus category-based information depended on the attentional resources.

To achieve these goals, we conducted three experiments using a modified trust game (Tortosa et al., 2013), in which participants had to predict the cooperative behavior of their partners to earn monetary outcomes. First, in the baseline phase, partners' emotional expressions were independent of their cooperation tendency, allowing us to verify whether positive emotions would cue positive outcomes (Tortosa et al., 2013), and therefore prompt cooperation. Next, in the learning phase, we manipulated partners' behaviors by establishing counterintuitive cooperation contingencies between partners' emotional expression and their cooperative trend. Specifically, the same partners were high in cooperation when portraying an angry expression, and low in cooperation when portraying a happy expression. Therefore, if participants attended to individual identities, learning would be impaired. If, however, they attended to emotions, they should learn over time about the cooperation trends of their partners. With this procedure, the magnitude of participants' learning is informative of whether they use individual identities or emotional information to predict their partners' behavior. We further manipulated the number of partners with whom participants played the trust game to verify if attentional resources modulated learning.

Notably, in Experiment 1, participants played with male and female partners being cooperative when they displayed an angry emotional expression, and not cooperative when they displayed a happy emotional expression. In Experiment 2, we narrowed the potential predictors of partners' behavior to emotional expressions and individual identities by removing gender information so that all participants only played with women. Finally, in Experiment 3, both individual identities and emotional expressions were predictive of partners' behavior, as each particular partner displayed only one emotional expression in the entire task. In all experiments, we experimentally manipulated the cognitive cost of individuation by introducing a between-groups manipulation of the number of partners in the trust game, such that participants played with either 4 or 32 partners.

We broadly predicted that in the baseline, participants would cooperate more with partners when they displayed a happy emotional expression. Moreover, we expected that participants playing with a smaller number of partners would learn better the associations established between emotional expression and behavior in the learning phase.

Experiment 1

The goal of Experiment 1 was to investigate how emotions and individual identities are used to predict others' behaviors. Notably, participants played the trust game with either 4 or 32 partners displaying angry and happy emotional expressions. In the baseline, in which emotion was not predictive of partners' behaviors, we expected that happiness compared to anger would spontaneously cue positive outcomes and therefore participants would cooperate more with partners portraying happy as compared to angry expressions (Tortosa et al., 2013). Further, in the learning phase in which partners were cooperative when they displayed an angry emotional expression, and not cooperative when they displayed a happy emotional expression, we predicted different patterns of learning for the two experimental groups. In line with impression formation theories (Fiske & Neuberg, 1990), we expected participants to be highly motivated to individuate their partners, and therefore to attend their individual identities. However, the cognitive resources needed to individuate were expected to impact learning differentially. Specifically, when playing with 4 partners, we expected participants to learn over time that individual identity was not predictive of partners' behaviors, thus focusing and learning about the associations between emotional expression and behavior. Instead, when playing with 32 partners, we also expected participants to attempt to focus on individual identities, but given the large number of partners, this strategy would divert attention away from the relevance of emotions and impair learning (Jiménez & Méndez, 1999). The hypotheses and procedures of this experiment were preregistered before data collection on Open Science Framework (osf.io/62gwk).

Method

Participants. Thirty-two undergraduates (12 men, mean age of 25.03 years), volunteered to take part in the experiment. A sensitivity power analysis assuming an alpha criterion of 0.05 and a power criterion of 0.80 revealed that with the present sample of 32 participants, the smallest effect size that could have been detected for the critical Emotion x Number of Partners interaction was $f = .36$, for this and the following experiment. In all of the reported experiments, participants had normal or corrected to normal vision and were naïve to the purposes of the study. At the end of the experiment, participants were rewarded with an economic compensation proportional to their performance in the task (€4.41 on average). This research is part of a larger research project approved by the local university ethical committee (175/CEIH/2017) on the use of human participants in research.

Procedure. In a modified multi-round trust game (Tortosa et al., 2013), participants were presented with the photographs of partners expressing two distinct emotions, anger and happiness. Furthermore, to manipulate the ability to individuate others (Fiske & Neuberg, 1990; Gilbert & Hixon, 1991), participants were randomly assigned to play the trust game with either the same 4 partners or with 32 different partners.

In the baseline phase (one block of 64 trials), neither partners' individual identity nor their emotional expressions were predictive of their cooperation behavior. That is, partners cooperated on half of the trials regardless of whether they portrayed a happy or angry expression. In the learning phase (5 blocks of 64 trials), the same partners appeared once again with both happy or angry expressions, however, now counter-intuitive association between emotional expressions and partners' behavior were introduced. Specifically, when partners portrayed an angry expression they cooperated on 75% of the trials and when they portrayed a happy expression, they cooperated on 25% of the trials. Across the trials, the same partners appeared with both happy or angry expressions, thereby reflecting the unstable property of emotions. Given that in this experiment, individual identities were not relevant, we ensured that in each block, happiness and anger were portrayed in 50% of the trials, and displayed by men in half of the trials, and women in the other half, regardless of their individual identities. To succeed in this task, participants had to learn the counterintuitive contingencies established between expressions and cooperation tendency and ignore the individual identities. Before performing these 6 blocks of trials, participants performed 8 practice trials with 2 men and 2 women presented twice, once being cooperative and once being uncooperative.

Stimuli and materials. Stimuli were displayed on a 24'' monitor placed at 60 cm from participants. E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) was used for stimuli presentation and data acquisition. Thirty-two partners (half women) portraying either happiness or anger (sixty-four pictures in total), were taken from the Karolinska Directed Emotional Faces (KDEF) database (Lundqvist, Flykt, & Öhman, 1998). At the beginning of each trial (see Figure 5.1), participant saw "€2" for 1000 ms to signal the amount they could invest or keep. Next, a fixation point appeared in the center of the screen (0'95º) for 100 ms followed by an image of a partner (8.10º x 10.95º) for 1000 ms. Participants had to decide whether or not to cooperate by pressing the key Z or M, counterbalanced across participants. If participants decided not to cooperate, they would keep the initial €2 and the partner would receive nothing. If participants decided to cooperate, the partner ostensibly received the initial amount multiplied by 5 (i.e., €10). The partner could then decide to either give €5 to the participant or to keep all of the money. At the end of the trial, feedback about the partner's decision was presented for 1700 ms.

Results and Discussion

Baseline. Before analyzing behaviors during the learning phase in which specific associations between facial expressions and cooperation were manipulated, we verified if participants spontaneously cooperated more with partners with happy compared to angry expressions. Specifically, we conducted a 2 (Emotion: Angry vs. Happy) x 2 (Number of Partners: 4 vs. 32) mixed-design Analysis of Variance (ANOVA) on cooperation responses during the baseline phase. This analysis revealed a main effect of emotion, $F(1, 30) = 20.50$, $p < .001$, $\eta_p^2 = .41$, 95% CI = [.13, .59]. As expected, participants cooperated more with partners when they displayed a happy ($M = .63$, $SD = .15$) than an angry ($M = .46$, $SD = .15$) expression. This effect was not qualified by the number of partners, as the two-way interaction was not significant, $F(1, 30) = .33$, $p = .568$, $\eta_p^2 = .01$, 95% CI = [.00, .17].

Learning phase. To examine whether participants playing with 4 partners would learn over time that partner emotions predicted cooperative behavior more than participants playing with 32 partners, participants' cooperation rates in the learning phase were subjected to a 2 (Emotion: Angry vs. Happy) x 5 (Block: 2-6) x 2 (Number of Partners: 4 vs. 32) mixed-design ANOVA. The Emotion x Block interaction was significant, $F(4, 120) = 6.54$, , $p < .001$, $\eta_p^2 = .18$, 95% CI = [.05, .28]. Although participants cooperated more across blocks with partners when they portrayed an angry expression, as predicted, this linear effect was not significant, $F(1, 30) = 2.55$, $p = .12$, η_p^2

= .08, as shown in Figure 6.1. However, when partners portrayed a happy expression, participants linearly decreased their cooperation over the five blocks, $F(1, 30) = 11.59$, $p = .002$, $\eta_p^2 = .28$, 95% CI [.05, .49]. Together these results suggest that participants were able to learn the counter-intuitive contingencies between partners' emotional expressions and cooperation over time. Although learning was most clearly reflected in decreasing cooperation with happy partners over time, the pattern of findings related to increasing cooperation with angry partners was in the predicted direction, but not significant.

A marginal Emotion x Number of Partners interaction was also found, $F(1, 30) = 3.23$, $p = .082$, $\eta_p^2 = .10$, 95% CI = [.00, .31]. Simple effects analyses demonstrated that when playing with 32 partners, participants cooperated more with partners when they portrayed an angry ($M = .63$, $SD = .11$) than happy ($M = .40$, $SD = .16$) expressions, $F(1, 15) = 38.01$, $p < .001$, $\eta_p^2 = .72$, 95% CI = [.38, .83]. However, when playing with 4 partners, participants did not differ in their cooperation with partners with angry ($M = .55$, $SD = .23$) and happy ($M = .49$, $SD = .17$) expressions, $F(1, 15) = .67$, $p = .425$, $\eta_p^2 = .04$, 95% CI = [.00, .31].

Notably, participants playing with 32 partners were able to learn and respond to the contingencies between affective information and cooperation. Participants playing with only 4 partners, however, appear to have learned these contingencies to a lesser degree or to not base their responses on these contingencies. It is possible that when interacting with only 4 partners, participants did not automatically respond to emotional cues to predict their behavior but instead focused on their partners' individual actions. This focus on the specific target rather than the affective information would have led them to a less appropriate strategy of cooperation than participants with 32 partners. Alternatively, because 32 partners and their individual behaviors may represent too much information too process, participants may have categorized these targets according to easily accessible visible cues such as emotional expressions. Given that emotions are often automatically processed because of their clear implications for behavior (Tracy & Robins, 2008), when unable to individuate targets, participants may have relied on simply processing emotions and their contingencies.

Although in general all participants over time learned to cooperate less with happy faces and more with angry faces, participants interacting with 32 partners consistently responded more according to these contingencies than participants playing with 4 partners.

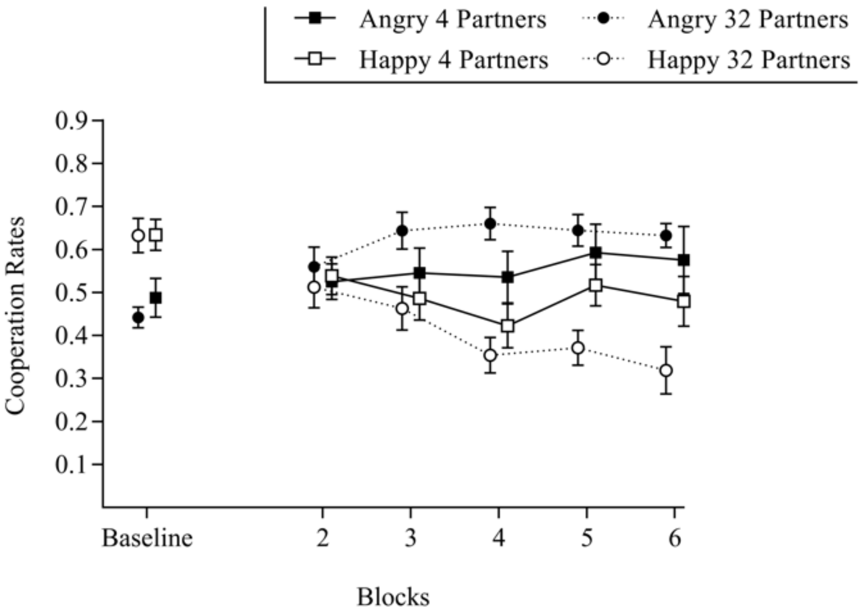

Figure 6.1. Cooperation rates as a function of partner's emotion (happy vs. angry) and the number of partners (4 vs. 32) in Experiment 1. Error bars represent the standard error of the mean.

Experiment 2

Although Experiment 1 provides initial evidence suggesting that people may not respond according to emotional contingencies when there is a limited number of targets and they are able to individuate them, it is important to note that the primary interaction was only marginal. The primary goal of Experiment 2 was therefore to replicate the findings of the first study.

Furthermore, in Experiment 1, participants were presented with partner stimuli that was comprised of male and female adults. Given that people often categorize others based on age, race, and gender (Stolier & Freeman, 2016), it is possible that participants may have attended to whether their partner was male or female when deciding whether to cooperate in the trust game. This focus may have impaired learning, especially in the 4 partner condition, in two different ways. First, when presented with only 4 partners, participants may have been able to attend to a variety of attributes, including gender, and that focus may have diverted attention away from the relevant and predictive affective information. Second, presenting both male and female faces may have activated gender stereotypes that influenced perceptions of emotions, or impacted learning of counter-stereotypical associations of emotions and cooperative behaviors (i.e., men expressing happiness and women expressing anger) (Plant, Hyde, Keltner, & Devine, 2000). To circumvent this issue, in Experiment 2, we focused on female partners.

Method

Participants. We used the same sample size as in Experiment 1 and collected data from 32 undergraduates (13 men, mean age of 21.65 years). As in Experiment 1, participants were rewarded with an economic compensation proportional to their performance in the task (€4.88 on average) at the end of the experiment.

Stimuli, materials and procedure. The stimuli, materials, and procedure were similar to Experiment 1 with the exception that only photographs of women (32 images) from the KDEF database (Lundqvist et al., 1998) were included.

Results and Discussion

Baseline. To initially analyze the data from the baseline, a 2 (Emotion: Angry vs. Happy) x 2 (Number of Partners: 4 vs. 32) mixed-design ANOVA was conducted on cooperation rates. The main effect of emotion was again significant, $F(1, 30) = 14.01$, $p < .001$, $\eta_p^2 = .32$, 95% CI = [.07, .52]. As in Experiment 1, participants cooperated more with partners when they portrayed a happy expression ($M = .61$, $SD = .11$) than when they portrayed an angry expression ($M = .50$, $SD = .11$). As in Experiment 1, the Emotion x Number of Partners interaction was not significant, $F(1, 30) = .39$, $p = .539$, $\eta_p^2 = .01$, 95% CI = [.00, .17].

Learning phase. To examine whether playing with 4 compared to 32 partners increased learning of the emotion-behavior contingencies, cooperation rates in the learning phase were subjected to a 2 (Emotion: Angry vs. Happy) x 5 (Block: 2-6) x 2 (Number of Partners: 4 vs. 32) mixed-design ANOVA. The Emotion x Block interaction was significant, $F(4, 120) = 12.88$, $p < .001$, $\eta_p^2 = .30$, 95% CI = [.15, .40]. Across blocks, participants' cooperation linearly increased when partners portrayed angry expressions, $F(1, 30) = 9.90$, $p = .004$, $\eta_p^2 = .25$, and linearly decreased when they portrayed happy expressions, $F(1, 30) = 33.40$, $p < .001$, $\eta_p^2 = .53$, see Figure 6.2.

Moreover, the Emotion x Number of Partners interaction was now also significant, $F(1, 30) = 16.89$, $p < .001$, $\eta_p^2 = .36$, 95% CI = [.10, .55], and not qualified by blocks $F(4, 120) = 1.89$, $p = .117$, $\eta_p^2 = .06$. Although when playing with 4 partners, participants cooperated more with partners when they portrayed an angry expression ($M = .58$, $SD = .15$) than when they portrayed a happy ($M = .43$, $SD = .15$) expression, $F(1, 15) = 5.28$, $p < .036$, $\eta_p^2 = .26$, 95% CI = [.00, .53], this difference was much larger when participants played with 32 partners, $F(1, 15) = 121.75$, $p < .001$, $\eta_p^2 = .89$, 95% CI = [.72, .93] (angry

M = .77, *SD* = .10; happy *M* = .30, *SD* = .12), as shown in Figure 6.2. Importantly, however, this difference was again opposite to our prediction and now clearly significant.

Despite focusing on female targets, these data replicated the results in Experiment 1 that interacting with 32 partners in the trust game led to better learning of emotion-partner cooperation contingencies than playing with 4 partners. The difference in learning between participants who ostensibly played the trust game with 4 compared to 32 partners indicates that their cooperation decisions were differentially influenced by the affective information from their partners' faces. While both groups in Experiment 2 were spontaneously influenced by the contingencies related to emotional cues, the context impacted learning. Because affective information and its predictive value remained the same across groups, the differences in learning between participants playing with 4 vs. 32 partners suggest that they used different learning strategies. Notably, when playing with 4 partners, participants may have been better able to attend to the individual identity of each partner because of fewer cognitive demands (Gilbert & Hixon, 1991) and therefore may have been less impacted by the emotional expressions. However, when playing with 32 partners, participants may not have been able to individuate each partner because of the increased cognitive demands, and therefore, relied on the emotional expressions to predict the cooperation behaviors of their partners. In short, the cognitive demands of processing 32 compared to 4 partners may have resulted in a reduced ability to individuate each target and focus on their particular behaviors, and in a greater reliance on categorical cues such as affective information to inform their cooperation tendencies.

It is interesting to note that participants playing with 32 partners differentiated between happy and angry partners from the first block of the learning phase, $F(1, 15) = 37.46$, $p < .001$, $\eta_p^2 = .71$. These results suggest that participants in this groups never attended to individual identities or only attended to in the very early stages of learning. This question is explored in Experiment 3 by making both individual identities and emotional expressions predictive of partners' behaviors.

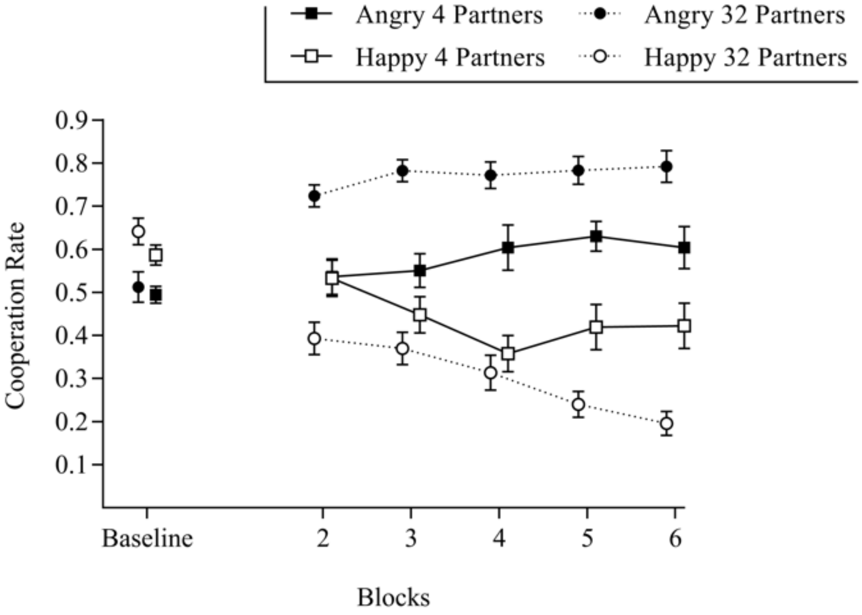

Figure 6.2. Cooperation rates as a function of partner emotion (happy vs. angry) and number of partners (4 vs. 32) in Experiment 2. Error bars represent the standard error of the mean.

Experiment 3

The goal of Experiment 3 was to investigate learning strategies when both individual identities and emotions predict partners' behaviors. Therefore, each partner expressed only one emotion, *either* happiness or anger, in the entire task. Because attention to either identities or emotions would be a successful strategy to learn behavioral contingencies, we expected participants' overall learning to increase compared to previous studies. Importantly, with this procedure, the impact of individual identities information on learning should be reflected in the difference in learning between participants playing with 4 vs. 32 partners. If participants primarily attend to individual identities, they should show better learning when playing with 4 compared to 32 partners, as individual learning is much easier in the former condition. However, if participants primarily attend to emotion, participants playing with 32 partners should show better learning than participants playing with 4 partners, as observed in Experiments 1 and 2. Or at least, there should be no difference in learning between participants playing with 4 vs. 32 partners as the predictive value of affective information was the same in the two experimental groups.

To better understand whether individuation or categorization strategy are used during learning, we further introduced a final transfer phase in which participants interacted with 4 partners from the learning phase and 32 new individuals portraying either angry or happy expressions. All partners in this phase cooperated on 50% of the trials regardless of their expression. This manipulation allowed us to verify whether participants maintained the same strategy of cooperation with partners in the learning phase when emotion was no longer predictive of behavior. If participants cooperated with all new partners regardless of their expression, that would indicate that they individuated partners and did not categorize them according to their emotional expression. In contrast, if participants cooperated more with new partners with angry compared to happy expressions, that would indicate that they categorized new individuals based on their emotional expression rather than responding to their individual identities.

As in Experiments 1 and 2, we expected participants to attend to individual identities in the learning phase, such that learning would be now larger when playing with 4 compared to 32 partners. Furthermore, in the transfer phase, we expected only participants playing with 32 partners to transfer knowledge from the learning phase to new individuals (i.e., to cooperate more with new partners with angry compared to happy expressions), as playing with a large group of partners (i.e., 32 partners) is more likely to prompt a resources-saving categorization strategy. The hypotheses and analyses of this experiment were preregistered before data collection in Open Science Framework (osf.io/b6wh3).

Method

Participants. Following our pre-registration plan, forty undergraduates (20 men, mean age of 22.65 years) participated in the experiment. A sensitivity power analysis assuming an alpha criterion of 0.05 and a power criterion of 0.80 revealed that with our sample of 40 participants, the smallest effect size that could have been detected for the critical Emotion x Number of Partners interaction was $f = .31$. As in previous experiments, participants were rewarded with an economic compensation proportional to their performance in the task (€4.88 on average) at the end of the experiment.

Stimuli and materials. Stimuli and materials were similar to Experiment 1, except that more target stimuli were needed for the transfer phase (64 images in total, half women), all taken from the KDEF database to represent partners (Lundqvist et al., 1998).

Procedure. The number of partners with whom participants played the trust game was again manipulated as a between-participants factor. Moreover, the task was divided in three phases: the baseline, the learning phase and the transfer phase. The baseline was similar to previous studies, except that each partner displayed the same emotional expression in the entire task. For example, in the group playing with 4 partners, 2 partners (one man and one woman) always displayed happy emotional expression while the other 2 partners (a different man and a different woman) always displayed angry emotional expression. The learning phase consisted of 6 blocks of 64 trials in which we introduced our counterintuitive association between partners' emotional expression and their reciprocation rates. Because in this experiment individual identities were relevant for learning, this phase was longer than in the previous experiments to ensure that all partners were presented the same number of times (96 in the group who played with 4 partners and 12 in the group who played with 32 partners). Finally, a transfer phase was also included which consisted of one block of 64 trials. In this phase, 4 partners from the learning phase were presented 8 times with the same emotion as in the previous blocks, but cooperated in half of the trials independent of target emotions. Sixteen new partners (half women) not previously presented were shown twice, once being cooperative and once uncooperative. Half of the partners displayed a happy expression while the other half displayed an angry expression. In the group who played with 32 partners, the set of 4 faces taken from the learning phase was counterbalanced across participants, such that across participants, all 32 faces used in the learning phase were also used in the transfer phase.

Results and discussion

Baseline. To examine initial cooperation with emotional expressions, a 2 (Emotion: Angry vs. Happy) x 2 (Number of Partners: 4 vs. 32) mixed-design ANOVA was conducted on cooperation responses during the baseline phase. The main effect of emotion was significant, $F(1, 38) = 9.79$, $p = .003$, $\eta_p^2 = .21$, 95% CI = [.03, .40]. Once again, before contingencies related to emotional expression and cooperation were introduced, participants cooperated more with partners with happy ($M = .60$, $SD = .12$) than angry ($M = .49$, $SD = .17$) expressions.

Surprisingly, and in contrast to the previous experiments, the Emotion x Number of Partners interaction was marginally significant, $F(1, 38) = 3.84$, $p = .057$, $\eta_p^2 = .09$, 95% CI = [.00, .28]. When participants played with 32 partners, they cooperated more with partners with happy ($M = .65$, SD = .12) than angry ($M = .48$, $SD = .18$) expressions, $F(1, 19) = 10.98$, $p = .004$, $\eta_p^2 = .37$, 95% CI = [.05, .59]. However, when participants played with 4 partners, $F(1, 19) = .83$, $p = .373$, $\eta_p^2 = .04$, 95% CI = [.00, .28], they did not differ in their cooperation with partners with happy ($M = .55$, $SD = .09$) and angry (M

= .51, *SD* = .17) expressions, suggesting that participants playing with 32 partners used partners' emotion to predict their behavior while participants playing with 4 partners did not. Alternatively, participants may have used emotions at the beginning of the block, but quickly identified their partners' individual random cooperative behavior, so they finally reached a 50% cooperation rate with all of them. In fact, during the first 32 trials, participants playing with 4 partners did cooperate more with partners with happy (M = .63, SD = .15) than angry (M = .52, SD = .20) emotional expressions, $F(1, 39) = 5.73$, $p = .020$, $\eta_p^2 = .13$.

Learning phase. To examine whether playing with 4 compared to 32 partners increased learning of contingencies between emotional expressions and partner behavior, cooperation rates in the learning phase were subjected to a 2 (Emotion: Angry vs. Happy) x 6 (Block: 2-7) x 2 (Number of Partners: 4 vs. 32) mixed-design ANOVA. The Emotion x Block interaction was significant, $F(5, 190) = 11.62$, $p < .001$, $\eta_p^2 = .23$, 95% CI = [.12, .31], reflecting an increase in learning across blocks. Specifically, participants linearly increased their cooperation across blocks with partners with angry expressions, $F(1, 38) = 9.16$, $p = .004$, $\eta_p^2 = .19$, and linearly decreased it with those with happy expressions, $F(1, 38) = 21.79$, $p < .001$, $\eta_p^2 = .36$, see Figure 6.3.

The Emotion x Number of Partners interaction was also significant, $F(1, 38) = 22.13$, $p < .001$, $\eta_p^2 = .37$, 95% CI = [.13, .54]. Although both groups cooperated more with partners with angry expressions than with partners with happy expressions, this difference was larger when participants played with 4 partners, $F(1, 19) = 68.73$, $p < .001$, $\eta_p^2 = .78$, 95% CI = [.54, .86] (angry M = .76, SD = .14, and happy M = .27, SD = .17), than when they played with 32 partners, $F(1, 19) = 20.35$, $p < .001$, $\eta_p^2 = .52$, 95% CI = [.17, .69] (angry M = .57, SD = .16, and happy M = .41, SD = .13).

In contrast to the previous two experiments, participants who played the trust game with 32 partners were worse, not better, at learning the emotion-partner behavior contingencies, suggesting that participants chose a strategy that is harder to apply with 32 than with 4 partners. Notably, these findings confirmed that all participants at least attempted to predict their partners' behaviors on the basis of their individual identities. Because individuating 4 partners is relatively easy and relevant in Experiment 3, when behavioral contingencies were related to both individual identities and emotional cues, this strategy led to better performance on the trust game. However, when focusing on individual identities was not functional as in Experiments 1 or 2, when behavioral contingencies were only related to emotional cues, this strategy led to worse performance. Moreover, categorizing individuals according to two emotional categories should be easier than individuating 4 different individuals. The fact that participants playing with 32 partners learned to a lesser extent the predictive value of affective information than participants who were individuating 4 partners suggests that attention to individual identities information impaired learning in the group playing with 32 partners.

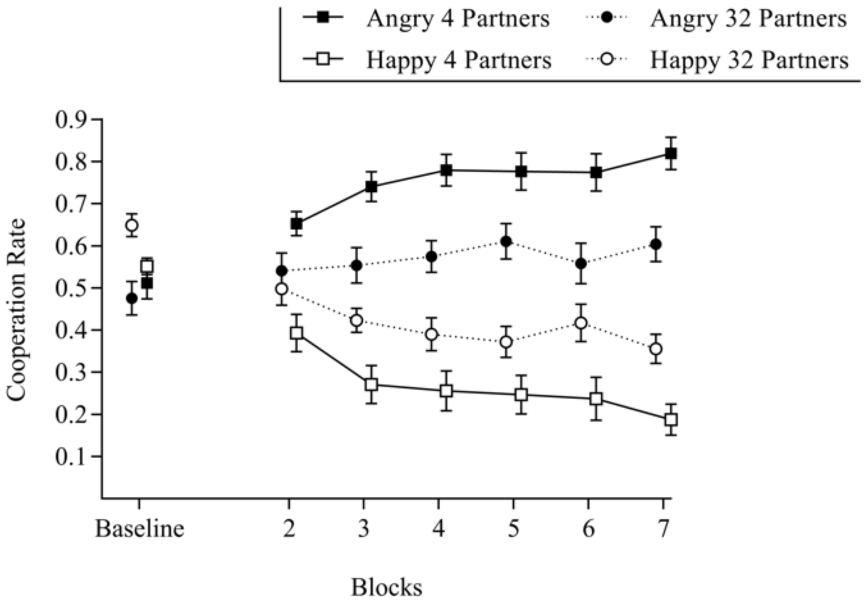

Figure 6.3. Cooperation rates as a function of partner emotion (happy vs. angry) and number of partners (4 vs. 32) in Experiment 3. Error bars represent the standard error of the mean.

Transfer phase. To verify whether participants transferred their knowledge from the learning phase to new individuals with whom they had no prior experience, a 2 (Emotion: Angry vs. Happy) x 2 (Partner: Old vs. New) x (Number of Partners in the Learning Phase: 4 vs. 32) mixed-design Analysis of Variance (ANOVA) was conducted on cooperation rates in the transfer phase. This analysis showed a significant Emotion x Partner x Number of Partners interaction, $F(1, 38) = 16.00$, $p < .001$, $\eta_p^2 = .30$, 95% CI = [.08, .48]. When participants played with 4 partners, the Emotion x Partner interaction was significant, $F(1, 19) = 32.73$, $p < .001$, $\eta_p^2 = .63$, 95% CI = [.30, .77]. Simple effects analyses demonstrated that participants cooperated more with partners with angry (M = .72, SD = .19) than happy (M = .28, SD = .18) expressions, but only for partners who had been previously presented in the learning phase, $F(1, 19) = 42.94$, $p < .001$, $\eta_p^2 = .69$, 95% CI = [.38, .80], and not for new partners, $F(1, 19) = .12$, $p = .731$, $\eta_p^2 < .01$, 95% CI = [.00, .19]. Therefore, participants who played with 4 partners in the learning phase learned accurately their partners' reciprocation tendencies and expected them to maintain the same behavior in the transfer phase. This strong learning and expectation prevented them from realizing that their partners' emotion were no longer predictive of their behavior in the transfer phase. However, for new partners who had not been presented in the learning phase, there was no difference in cooperation between partners with happy (*M* = .50, *SD* = .23) and angry (*M* = .48, *SD* = .28) expressions, $F(1, 19) = .12$, $p = .731$, $\eta_p^2 < .01$, 95% CI = [.00, .19], indicating that participants did not transfer their knowledge about happy and angry partners presented in the learning phase to new individuals.

In contrast, when participants played with 32 partners, the Emotion x Partner interaction was not significant, $F(1, 19) = .02$, $p = .888$, $\eta_p^2 < .01$, 95% CI = [.00, .13]. Participants did not differ in their cooperation with partners with happy (*M* = .51, *SD* = .17) and angry (*M* = .50, *SD* =.22) expressions when these partners had been previously presented in the learning phase, $F(1, 19) = .06$, $p = .818$, $\eta_p^2 < .01$, 95% CI = [.00, .16], or with partners with happy (*M* = .51, *SD* = .22) and angry (*M* = .48, *SD* = .21) expressions when these partners had not previously been presented, $F(1, 19) = .30$, $p = .593$, $\eta_p^2 = .02$, 95% CI = [.00, .24], see Figure 6.4, suggesting that participants playing with 32 partners did not use emotions to predict their partners' behavior at all in the transfer phase.

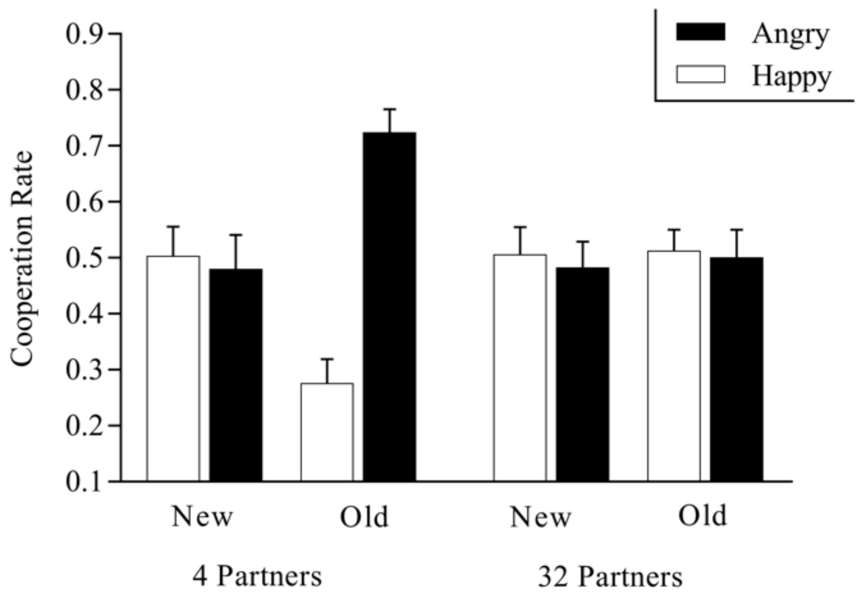

Figure 6.4. Cooperation rates as a function of the type of partner (old partners were presented in the learning phase and new were presented only in the transfer phase) and the number of partners (4 vs. 32). Error bars represent the standard error of the mean.

General Discussion

While it has extensively been argued that categorization is the default strategy in impression formation (Fiske & Neuberg, 1990; Kawakami, Amodio, & Hugenberg, 2017), the present research provides new evidence that in some circumstances individual identities may be favored over categorical cues, even when this strategy impairs performance. Across three experiments, we examined the differentiated use of individual identities and emotional cues in a context in which participants should be highly motivated to individuate their partners to earn economic rewards (Neuberg & Fiske, 1987; Telga et al., 2018). Partners' behaviors were associated with specific emotions: happy faces were paired with non-equitable behaviors and angry faces with equitable behaviors.

We predicted that in the baseline, when neither emotional expression nor individual identities predicted partners' behavior, participants would cooperate more with happy compared to angry faces. The data across three studies supported these predictions, although in Experiment 3, this effect was qualified by the number of partners. In particular, when individual identities and emotions were fully overlapped, and participants played with only 4 partners, and therefore individuation was easy, participants may have noticed that emotions were not predictive of their partners' behavior, and thus chose not to use it.

During the learning phase, we predicted that although participants would be highly motivated to individuate, learning would be affected by either a categorization or an individuation strategy, depending on the attentional resources available to individuate (Fiske & Neuberg, 1990; Macrae, Hewstone, & Griffiths, 1993). However, data from the three experiments seem to suggest that participants tried to individuate their partners in all task settings. In fact, when individual identities were not predictive of partners' behavior (Experiments 1 and 2), participants playing with 4 partners showed worse learning than those playing with 32. It is reasonable to think that, because the 4 individual identities were chosen over categorical information (i.e., emotions) to predict partners' behavior, the overall learning was worse, as participants playing with 4 partners did not abandon the individuation strategy despite its lack of efficiency. Only when, in addition to being non-predictive, individual identities of 32 partners were hard to process, participants responded according to categorical information based on emotion.

When individual identities and emotional expressions were perfectly correlated (Experiment 3), learning was impaired with 32 compared to 4 partners. In this condition, attending to individual identities was helpful to perform the task but required more resources than attending to affective information. The fact that playing with 32 partners resulted in impaired learning compared to playing with 4 partners suggests that both experimental groups attended individual identities, but participants playing with 32 partners were less able to learn because of the large number of individuals. These findings also suggest that although participants playing with 32 partners in Experiments 1 and 2 seemed to respond to categorical emotional cues, they did not completely categorize their partners. In fact, data from Experiment 3 suggested that participants' learning was reduced by paying attention to individual information when playing with 32 partners. Interestingly, the magnitude of this learning is not much different from the magnitude of the learning observed in Experiment 1. Therefore, it is possible than even in Experiment 1, in which attending to partners' individual identities was not functional to perform the task, participants' responses were also impacted by individual information. Because categorizing in 2 categories should be easier than individuating 4 individuals, a clear categorization strategy, not impacted by individual identities, would have been reflected in an even larger learning when playing with 32 partners (2 categories) than the one observed when participants individuated 4 partners in Experiment 3.

This interpretation is consistent with the data observed in the transfer phase, suggesting an individuating approach of new partners, as participants did not respond to their emotional expressions in any of the experimental groups. Participants who played with 4 partners, however, cooperated more with the same partners previously associated with greater cooperation in the learning phase. These findings may be the result of highly consolidated individual learning that persisted in the transfer phase.

Importantly, the data of the present research challenge both impression formation theories and evolutionary accounts of emotions. While the former postulates that the default strategy to process others is the one that requires the least cognitive resources (Fiske & Neuberg, 1990), the latter predicts that affective cues are automatically processed to make quick decisions (Tracy & Robins, 2008). Therefore, both approaches would predict that participants' responses would be largely guided by affective cues. Importantly, beyond cognitive resources, attention to emotional expressions was especially relevant in this task, as it was either the only (Experiments 1 and 2) or the best (Experiment 3) strategy to perform the task. While we made no attempt to mask the partners' emotional expressions, nor their predictive value (i.e., emotional expressions consistently cued partners' behaviors), participants failed to respond to affective cues, unless they were under a high cognitive load and individual identities did not predict their partners' behaviors (Experiments 1 and 2), and never refrained from responding to individual identity information.

Together, these experiments suggest that with the appropriate motivation, people may individuate others and rely less on categorical processing. Notably, individuation may even be preferred, in cases when categorical information is more predictive of behavior. While it has been largely demonstrated that we perceive others categorically even when individuation is objectively possible (e.g., Cañadas, Rodríguez-Bailón, Milliken, & Lupiáñez, 2013; Kawakami et al., 2017), the present findings provide evidence for an alternative possibility: individuation may be largely preferred, even when individual identities do not predict behaviors (Experiments 1 and 2) or under high cognitive load (Experiment 3). Beyond the need of cognitive efficiency in social contexts, some situations may rather promote that all resources are engaged to achieve the proposed goals.

Acknowledgments and funding

We are thankful to all students who participated in the experiments, to Nuria Aguerre for her help in programming the experiments, and to Beatriz Barbero and Cristina Cano for their help in data collection. This research is part of MT's thesis dissertation under the supervision of JL.

This work was supported by the Spanish Ministry of Education, Culture and Sports, with pre-doctoral FPU fellowship FPU14/07106 to MT, the Social Sciences and Humanities Research Council of Canada (SSHRC 435-2013-0992) to KK, and the Spanish Ministry of Economy and Competitiveness, with research projects PSI2014-52764-P and PSI2017-84926-P to JL.

Correspondence concerning this article should be addressed to Maïka Telga, Departamento de Psicología Experimental, Universidad de Granada, Campus de Cartuja s/n, 18 071 Granada, Spain. Email: mtelga@ugr.es

EXPERIMENTAL SERIES 7

The content of this experimental series is in preparation as:

Telga, M., Acosta, A., & Lupiáñez, J. (2019). Differential Impact of Sadness and Anger on Social Perception. [Manuscript in preparation].

DIFFERENTIAL IMPACT OF SADNESS AND ANGER ON SOCIAL PERCEPTION

Abstract

In social perception, social categorization reflects the use of heuristic based information, allowing us to make inferences about others on the basis of information related to their group membership, while individuation involves a deeper analysis and integration of diverse individual features. One of the several factors impacting impression formation processes is the perceiver's emotional state. The present research examines whether two emotions of negative valence, sadness and anger, may prompt different strategies to make inferences about unknown individuals. Specifically, in an adaptation of the trust game, participants learned that two ethnic groups (i.e., Blacks and Whites) displayed opposite patterns of cooperation, one group being equitable and the other one being non-equitable. Next, participants were induced with either anger, sadness, or a neutral emotional state, and subsequently played the trust game with new black and white partners with whom they had no prior experience. We predicted that participants induced with anger would use a heuristic-based strategy to predict their new partners' behaviors, cooperating with them according to the previous categorical learning. In contrast, participants induced with sadness were expected to engage in more detail-oriented analyses, attempting to individuate their new partners independently of the categorical associations learned. The data supported these hypotheses, suggesting that negative emotions may impact differentially our perception of and behavior with others.

Keywords: impression formation, heuristics, categorization, individuation, emotion

Being able to make accurate and quick decisions in social contexts is crucial for harmonious social interactions. In many daily situations, we may need to make inferences about a person's state of mind to behave accordingly, even at zero acquaintance. In such contexts, many bodily and facial cues can inform our judgments (Kawakami, Amodio, & Hugenberg, 2017; Oosterhof & Todorov, 2008). Notably, research in social psychology has demonstrated that the strategies used to integrate and make sense of all this information lie on a continuum reflecting the extent to which we use category-based vs. individual information (Brewer, 1988; Fiske, Lin, & Neuberg, 1999; Fiske & Neuberg, 1990).

A social categorization strategy consists of using the information diagnostic of a person's group membership to inform our judgments. In fact, information related to gender, ethnicity and age is quickly identified in the first stages of face processing (Bruce & Young, 1986; Stolier & Freeman, 2016), allowing to classify people according to these dimensions effortlessly (Brewer, 2007; Fiske & Neuberg, 1990; Kawakami et al., 2017). Once a person is categorized into a particular social group, inferences about this specific individual can be made based on our experience or knowledge about this social group. On the other end of the continuum, the individuation strategy requires to attend to the unique personal characteristics of a target to inform our judgments. Individuating targets requires integrating piecemeal information to get to a final assessment of the individual. This strategy demands more cognitive resources, therefore we individuate targets only when we have sufficient motivation (Neuberg & Fiske, 1987) and resources (Gilbert & Hixon, 1991; Macrae, Milne, & Bodenhausen, 1994) to do so. Put simply, social categorization is a heuristic processing fulfilling a resources-saving function. Whether a person will prioritize individual or category-based information depends on several personal and contextual factors such as the perceiver's goal and focus of attention, interdependence, power (Blair, 2002; Cañadas, Rodríguez-Bailón, Milliken, & Lupiáñez, 2013; Neuberg & Fiske, 1987), or the perceiver's emotional state.

In fact, discrete emotions from the same valence may trigger different strategies of information processing and decision-making according to their evolutionary functions (Angie, Connelly, Waples, & Kligyte, 2011). For instance, both anger and sadness are elicited in situations of negative valence. However, sadness is elicited when the perceiver feels responsible for the situation, or perceives he or she lacks control over the events. In contrast, anger emerges when the perceiver blames others for what has happened, or believes that something may be done to fix it (Lench, Tibbett, & Bench, 2016; Siemer, Mauss, & Gross, 2007). Thus, sadness and anger result in different coping strategies. While sadness has been linked to detail-oriented and analytic reasoning, as a strategy to avoid the thoughts related to the situation that elicited the feeling of sadness (Smith & Ellsworth, 1985; Wenzlaff, Wegner, & Roper, 1988), anger is associated with strategies allowing us to make quick decisions (Bodenhausen, Sheppard, & Kramer, 1994; Scott, 1980). Consistently with this theorizing, Bodenhausen et al. (1994) showed that participants induced with anger used heuristic cues and stereotyping judgments to a greater extent than participants induced with sadness.

Interestingly, Tiedens and Linton (2001) replicated these data with different emotions and reported that disgust promotes more heuristics-based thinking than fear. The authors suggested that the certainty appraisal of the situation determines whether the perceiver will engage in analytic reasoning or rather rely on heuristics. Emotions associated with certainty such as anger and disgust promote heuristic-based reasoning. In contrast, emotions associated with uncertainty or lack of control, such as sadness or fear, are associated with more analytic reasoning. Importantly, this interpretation also converges in indicating that the appraisal of the situation will determine the coping strategies, which in turn will differentially impact reasoning. However, neither of these studies directly tested these hypotheses in interpersonal settings, as participants were third-party judges who evaluated a situation in which they were not personally involved. The present study explores the impact of sadness and anger on impression formation, in a context in which participants are personally involved and outcome-dependent on making accurate predictions about people, at zero acquaintance.

The aim of the present research was to examine whether incidental emotional states of anger vs. sadness would promote different strategies to make inferences about unfamiliar people in interpersonal settings. To achieve this goal, we conducted an adaptation of the multi-round trust game (Telga, de Lemus, Cañadas, Rodríguez-Bailón, & Lupiáñez, 2018) in which participants had to learn the cooperative behaviors of unfamiliar partners to earn economic rewards. With this procedure, participants were active learners with several opportunities to form an accurate impression about their partners across repeated interactions. Participants played with black and white partners across three phases. In a baseline, all partners were cooperative in half of the trials, allowing to examine whether participants were biased to spontaneously cooperate more with one of the ethnic groups. In a learning phase, the two ethnic groups were associated with opposite cooperative behaviors. For instance, most black partners were equitable while most white partners were non-equitable. With this manipulation, participants may adopt a categorization strategy to make predictions about their partners, cooperating with them according to their ethnicity. Within each ethnic group, we also ensured that 25% of the partners were inconsistent with respect to the group behavior. Following the same example, a small proportion of black partners were non-equitable, and a small proportion of white partners were equitable. Therefore, participants may also adopt an individuation strategy, cooperating with inconsistent partners according to their individual behavior instead of the group behavior. By comparing participants' cooperation strategies with consistent vs. inconsistent partners, we can determine whether they used an individuation or a categorization strategy. If participants categorized their partners, they should adopt the same cooperation behavior with consistent and inconsistent individuals from the same ethnic group, according to the group behavior. Alternatively, if participants individuated their partners, they should revert their cooperation behavior with inconsistent individuals as compared with consistent ones. That is, they should display opposite patterns of cooperation for consistent vs. inconsistent partners belonging to the same ethnic group. Although previous studies using the trust game showed that participants mostly use an individuation strategy when playing with 8 partners (Telga et al., 2018), we expected a greater reliance on categorical information in the present experiment because of the higher cognitive cost of individuating 32 compared to 8 partners. Once participants have learned categorical, and to some extent individual information, we could verify how a specific emotional state differentially impacts the expression of the categorical information.

After the learning phase, participants were induced with either angry, sad or neutral emotional state. Next, in a transfer phase, participants played with new black and white partners with whom they had no prior experience, all being equitable in half of the trials. This phase was crucial to determine whether participants relied on heuristics or individuating strategies to predict their new partners' behaviors. If they used a categorization strategy, they should use their knowledge from the learning phase to categorize new partners, cooperating more with partners belonging to the ethnic group that was equitable in the learning phase. If they used an individuation strategy, they should try to learn the cooperative behavior of their new partners, independently of their ethnicity.

In the baseline, we expected to replicate previous research using the trust game in which white participants spontaneously cooperated more with black than with white partners (Telga et al., 2018; Tortosa, Lupiáñez, & Ruz, 2013). In the learning phase, we expected participants to accurately learn the group behavior, cooperating with consistent partners accordingly. We also expected participants to learn the cooperative behaviors of inconsistent individuals, although to a lesser extent. Finally, in the transfer phase, we expected participants induced with anger to use a categorization strategy, while participants induced with sadness were expected to individuate their new partners in line with Bodenhausen's (1994) work.

Method

Participants. Ninety volunteers took part in the study in exchange for economical reward proportional to their accuracy in the task (€5.70 on average). A sensitivity power analysis revealed that with this sample, the smallest effect size that could have been detected for the critical Induction x Group Behavior interaction in the transfer phase was $f = .20$. All participants reported normal or corrected to normal vision. Written informed consent was obtained from all participants. The study was part of a larger project approved by the local university ethical committee (175/CEIH/2017).

Apparatus, stimuli and materials. E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) was used for stimuli presentation and data collection. For the trust game, 64 pictures of 32 black (16 men and 16 women) and 32 white (16 men and 16 women) people were extracted from the Chicago Database (Ma, Correll, & Wittenbrink, 2015) to represent the partners. All partners were presented against a white background with a neutral emotional expression.

At the beginning of each trial, participants saw the euro symbol "€" for 190 ms, followed by a fixation point for 500 ms in the center of the screen. Next, the picture of the partner of this trial appeared for 1500 ms (5.68º x 7.77º) and participants had to decide whether or not to cooperate with him or her by pressing '1' to cooperate and '0' not to cooperate. If participants decided not to cooperate, they would keep the initial euro and the partner of this trial would receive nothing. If participants decided to cooperate, the partner would receive €5 and in turn decide to reciprocate giving back €2.50 to the participant, or to keep the whole money for him or herself. After making their decision, participants received visual feedback on their final outcomes in a single display presented during 1000 ms. The feedback displays included "You have cooperated and your partner has also cooperated. You have €2.50", "You have cooperated and your partner has not cooperated. You have €0", and "You have not cooperated. You have €1".

For the emotional induction, we used 20 pictures extracted from the International Affective Pictures System (IAPS, Lang, Bradley, & Cuthbert, 2005). The same 10 pictures of negative valence were used in the groups induced with sadness and anger, and 10 different pictures were used in the neutral induction condition. The pictures were displayed during 9 seconds with statements supposedly describing the scene from the picture to reinforce and specify the emotion inducted. For instance, one of the pictures depicted the corpse of a young man surrounded by 2 soldiers. In the group induced with anger, the description was "*He could have let him live, but his cruelty has no limit. He killed him intentionally. He thought it was fun to see the others' reaction*". In contrast, in the group induced with sadness, the description was "*He lost his little brother during the war. It was an irreparable misfortune. They were very united. His loss filled him with pain*". The pictures were presented together with a musical piece related to the affective state induced. In the group inducted with anger, the piece used was *Sacrificial Dance* extracted from *Rite of Spring* from Igor Stravinsky (Stravinsky, 1913), in the group inducted with sadness, the piece used was *Adagio* extracted from *Piano Concerto No 1, BWV 1052* from Johann Sebastian Bach (Bach, 1734), and finally, in the control group the piece used was a background music reproducing sounds from the forest (FreeAudioMusic, n.d.).

Two questionnaires were used to verify the effectiveness of the emotional induction before the induction, after the induction, and at the end of the experiment. First, the Escala de Valoración del Estado de Ánimo (EVEA, Sanz, Gutiérrez, & García-Vera, 2014) is a 16-item scale measuring transitory emotional states and providing scores from 0 to 4 on anger, sadness, happiness and fear. Second, the Self-Assessment Mannequin (SAM, Bradley & Lang, 1994) evaluates two basic dimensions of emotional states: activation and valence on a scale stretching from 1 to 9. At the end of the experiment, participants also filled the State-Trait Anxiety Inventory (STAI, Spielberg, Gorsuch, & Lushene, 2008), the Beck Depression Inventory II (BDI II, Sanz, Navarro, & Vázquez, 2003) and the State-Trait Anger Expression Inventory 2 (STAXI-II, Spielberger, 1999) to verify possible individual differences on anxiety, anger and depression traits, existing prior to the emotional induction. Finally, participants also filled a Social Desirability Scale (SDS, Ferrando & Chico, 2000) to ensure that they were not aware of the purposes of the study.

Procedure. The study was divided in four phases: baseline, learning, emotional induction and transfer, as shown in Figure 7.1. In the baseline phase, participants played the trust game with 16 white and 16 black partners (50% of women in each ethnic group), all of them being equitable on 50% of the trials. This phase allowed to verify whereas participants were biased to cooperate more with one of the two ethnic groups before learning. Each partner was presented twice, once being cooperative and once being not cooperative, resulting in 64 trials. In the learning phase, the two ethnic groups were associated with opposite behaviors such that one group was mostly equitable while the other group was mostly non-equitable. For instance, most Blacks were equitable on 75% of the trials while most Whites were non-equitable on 75% of the trials. Moreover, within each ethnic group, 25% of the partners were inconsistent with respect to their group, that is, they displayed a pattern of cooperation opposite to the group behavior. Following the same example, 4 Blacks were non-equitable on 75% of the trials and 4 Whites were equitable on 75% of the trials. The ethnic group associated with equitable or non-equitable behavior and the inconsistent partners within each ethnic group were counterbalanced across participants. With this procedure, participants' cooperation with inconsistent partners is informative of their learning strategies. If participants used a categorization strategy, they should cooperate across blocks in the same way with both consistent and inconsistent partners according to the group behavior. In contrast, if participants used an individuation strategy, they should display opposite cooperative behaviors across blocks with consistent and inconsistent partners, according to their individual behavior. Each partner was presented 16 times, resulting in 512 trials with auto-administered breaks every 64 trials. After the learning phase participants filled the EVEA and SAM questionnaires as a measure of their pre-induction affective state. Next, they were exposed to the induction procedure lasting around 2:30 minutes and filled again the EVEA and SAM questionnaires for a post-induction measure of their affective state. After the induction, participants performed the transfer phase in which they played the trust game with 32 new partners, 16 Blacks and 16 Whites with half of women in each ethnic group. Each partner was presented twice, once being cooperative, and once being uncooperative, resulting in 64 trials. This phase allowed to verify whether participants' strategy of cooperation with new partners. If participants categorized their new partners according to their knowledge from the learning phase, they should cooperate more with partners belonging to the group that was equitable in the learning phase, and less with the group that was non-equitable in the learning phase. In contrast, if participants individuated the new partners, they should try to start over learning their individual behaviors and cooperate with them independently of their ethnicity. At the end of the experiment, participants filled again the EVEA and SAM questionnaires, together with the STAI, the STAXI-2, the BDI 2 and the SDS. The experiment lasted around 80 minutes.

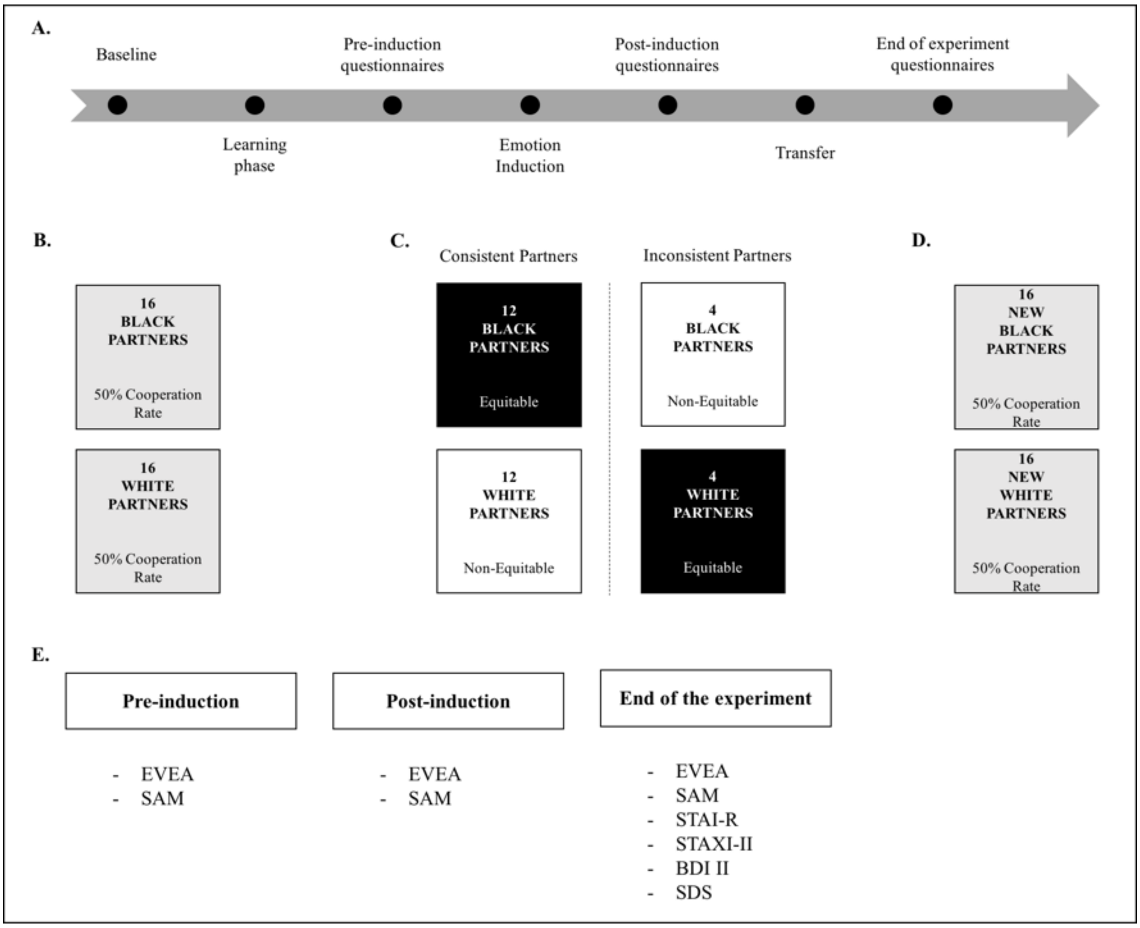

Figure 7.1. The general procedure of the experiment is represented on panel (A), as well as an example of the specific associations established in the (B) baseline, (C) learning and (D) transfer phases of the trust game. The questionnaires given to the participants in the pre- and post-induction, and at the end of the experiment, are described on panel (E).

Results

Because of technical issues, 13 participants with missing data were excluded from the analyses leaving in 77 participants for the analyses.

Baseline. A mixed-design Analysis of Variance (ANOVA) was conducted on cooperation rates in the baseline with partner ethnicity (black vs. white) as a within-participants variable and induction (anger vs. neutral vs. sadness) as a between-participants factor. We found a main effect of partner ethnicity, $F(1, 74) = 7.48$, $p = .008$, $\eta_p^2 = .09$, indicating higher cooperation with black ($M = .69$, $SD = .15$) than with white ($M = .64$, $SD = .15$) partners. As expected, there was no difference between the three experimental groups before the emotional induction. Indeed, neither the main effect of induction, $F(2, 74) = .23$, $p = .79$, $\eta_p^2 = .01$, nor the interaction Partner Ethnicity x Induction, $F(2, 74) = .60$, $p = .55$, $\eta_p^2 = .02$, were significant.

Learning. Data from the 512 trials of the learning phase were collapsed in 4 blocks of 128 trials. Therefore, cooperation rates during the learning phase were subjected to a mixed-design ANOVA with group behavior (equitable vs. non-equitable), consistency (consistent vs. inconsistent) and blocks (2-5) as within-participants variables, and induction (anger vs. neutral vs. sadness) as a between-participants variable. The critical Group Behavior x Consistency interaction was significant, $F(1, 70) = 59.14$, $p < .001$, $\eta_p^2 = .46$, and was qualified by the block variable as the three-way interaction was also significant, $F(3, 210) = 16.11$, $p < .001$, $\eta_p^2 = .19$. Surprisingly, the Group Behavior x Consistency x Block x Induction was also significant, $F(6, 210) = 2.25$, $p = .04$, $\eta_p^2 = .06$, likely because across blocks, the group under sadness induction differed less between equitable and non-equitable inconsistent individuals than the two other groups, as shown in Figure 7.2. The fact that Group Behavior x Consistency x Induction interaction was not significant, $F(2, 79) = 2.35$, $p = .10$, $\eta_p^2 = .06$, suggested that participants' overall learning was similar in the three experimental groups, but differed across the blocks. Therefore, we decided to verify that the three experimental groups did not significantly differ in their learning right before the induction, that is, in the last block of learning. We conducted the same analysis on cooperation rates in Block 5 and found again a significant Group Behavior x Consistency interaction, $F(1, 74) = 36.21$, $p < .001$, $\eta_p^2 = .34$. Importantly, this interaction was not qualified by the induction variable as the three-way interaction was not significant, $F(2, 71) = 1.50$, $p = .282$, $\eta_p^2 = .04$. Simple effects analyses showed that in the last block of learning, participants accurately cooperated more with equitable ($M = .72$, $SD = .16$) than non-equitable ($M = .42$, $SD = .23$) partners when they were consistent, $F(1, 74) = 85.51$, $p < .001$, $\eta_p^2 = .54$. In contrast, with inconsistent partners, participants did not significantly differ in their cooperation between partners belonging to an equitable ($M = .56$, $SD = .23$) and non-equitable ($M = .59$, $SD = .20$) group, $F(1, 79) = .60$, $p = .441$, $\eta_p^2 < .01$. This indicated that inconsistent partners were neither individuated nor totally categorized according to their group behavior. As none of the strategies was clearly favored over the other one, this leaves us with an ideal situation to explore the effect of the emotional induction on social learning strategies.

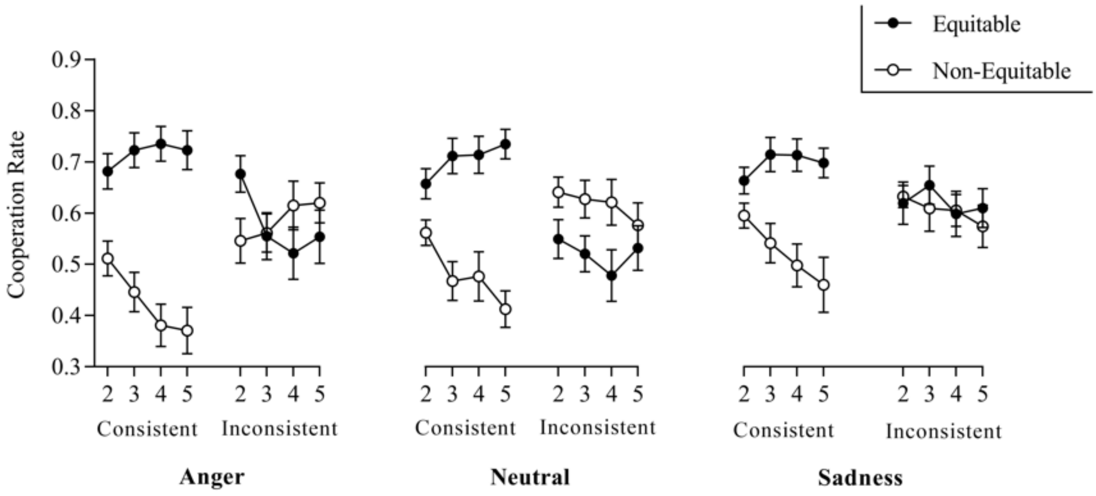

Figure 7.2. Cooperation rates in the learning phase with equitable and non-equitable partners across blocks as a function of their consistency, for each experimental group. Error bars represent the standard error of the mean.

Emotional induction. Before analyzing the effects of the induction procedure, we aimed at verifying that the three experimental groups did not differ on anger anxiety and depression traits. Separate univariate ANOVAs were conducted on the scores from the BDI (depression), STAXI-2 (anger) and STAI (anxiety) questionnaires with the induction group as a between-participants factor. We observed that the groups did not differ on the aforementioned measures, larger $F(2, 74) = 1.48$, $p = 234$, $\eta_p^2 = 04$, for the BDI scores.

To verify the effect of the induction procedure, we conducted a mixed-design ANOVA on scores at the EVEA questionnaire with time (pre-induction vs. post-induction vs. final) and emotion (anger vs. sadness) as within-participants variables and induction (anger vs. neutral vs. sadness) as a between-participants factor. This analysis revealed a significant Time x Emotion x Induction interaction, $F(4, 148) = 16.49$, $p < .001$, $\eta_p^2 = .31$.

On the measures of anger, the Time x Induction interaction was significant, $F(4, 148) = 18.04$, $p < .001$, $\eta_p^2 = .33$, and remained significant after introducing the SDS scores as a co-variable in the analysis, $F(4, 148) = 17.49$, $p < .001$, $\eta_p^2 = .32$, confirming that participants' anger scores were not led by social desirability effects. As shown in Figure 7.3, the three experimental groups did not significantly differ on their pre-induction scores of anger, $F(2, 74) = .16$, $p = .856$, $\eta_p^2 < 01$. In contrast, on the post-induction measures, the group induced with anger reported higher scores on anger than the neutral induction group, $F(1, 51) = 31.56$, $p < .001$, $\eta_p^2 = .38$, and the group induced with sadness, although in this case the difference was in the predicted direction, but not significant, $F(1, 50) = 2.17$, $p = .147$, $\eta_p^2 = .04$. The group induced with sadness also reported higher anger scores than the neutral group, $F(1, 47) = 19.20$, $p < .001$, $\eta_p^2 = .329$. At the end of the experiment, the three experimental groups did not significantly differ on their anger scores, $F(2, 74) = 1.83$, $p = .168$, $\eta_p^2 = .07$.

In the analysis of the sadness scores, the Time x Induction interaction was also significant, $F(4, 148) = 14.63$, $p < .001$, $\eta_p^2 = .28$, and remained significant after introducing the SDS scores as a co-variable in the analysis, $F(4, 148) = 14.12$, $p < .001$, $\eta_p^2 = .28$, again confirming that participants' sadness scores were not led by social desirability effects. The three experimental groups did not differ on their pre-induction sadness scores, $F(2, 74) = 2.43$, $p = .095$, $\eta_p^2 = .06$. In contrast, on the post-induction measures, the group induced with sadness reported higher sadness scores than the neutral group, $F(1, 47) = 42.31$, $p < .001$, $\eta_p^2 < .47$, and the group induced with anger, $F(1, 50) = 5.95$, $p = .018$, $\eta_p^2 = .11$. The group induced with anger also reported higher sadness scores than the neutral group, $F(1, 51) = 5.30$, $p = .025$, $\eta_p^2 = .09$. At the end of the experiment, the three experimental groups did not differ on their scores on sadness, $F(2, 74) = 1.67$, $p = .195$, $\eta_p^2 = .04$. Overall, although the groups induced with anger and sadness increased their scores on both anger and sadness, the increase on anger tended to be larger in the group induced with anger, and the increase in sadness was larger in the group induced with sadness.

Valence scores obtained from the SAM questionnaire were subjected to a mixed-design ANOVA with time (pre-induction vs. post-induction vs. final) as a within-participants variable and induction (anger vs. neutral vs. sadness) as a between-participants factor. This analysis revealed a significant Time x Induction interaction, $F(4, 148) = 14.56$, $p < .001$, $\eta_p^2 = .28$. In the pre-induction measures, the three experimental groups did not significantly differ on their valence scores, $F(2, 74) = .27$, $p = .767$, $\eta_p^2 < .01$. In contrast, in the post-induction measures, the main effect of induction was significant, $F(2, 74) = 13.19$, $p < .001$, $\eta_p^2 = .26$. The group induced with anger reported more negative valence scores after the induction procedure than the neutral induction group, $F(1, 51) = 21.28$, $p < .001$, $\eta_p^2 < .29$. The group induced with sadness also reported more negative valence scores than the neutral induction group, $F(1, 47) = 20.13$, $p < .001$, $\eta_p^2 = .30$. There was no difference between the group induced with sadness and the group induced with anger on their valence scores, $F(1, 50) = .04$, $p = .845$, $\eta_p^2 < .01$. At the end of the experiment, the three experimental groups did not differ on their valence scores, $F(1, 50) = 1.72$, $p = .187$, $\eta_p^2 < .04$.

The same analyses were conducted on the arousal scores obtained from the SAM questionnaire. Again, the Time x Induction interaction was significant, $F(4, 148) = 4.96$, $p < .001$, $\eta_p^2 = .12$. Before the induction procedure, the three experimental groups did not differ on their arousal scores, $F(2, 74) = .31$, $p = .733$, $\eta_p^2 < .01$. In the post-induction measure, the main effect of induction was significant, $F(2, 74) = 7.31$, $p < .001$, $\eta_p^2 = .17$. The group induced with anger reported higher arousal scores than the neutral induction group, $F(1, 51) = 13.74$, $p < .001$, $\eta_p^2 = .21$. The group induced with sadness also reported arousal scores higher than the neutral induction group, $F(1, 47) = 4.00$, $p = .051$, $\eta_p^2 = .08$. Finally, the group induced with anger reported arousal scores higher than the group induced with sadness, although this difference was only marginal, $F(1, 50) = 3.35$, $p = .073$, $\eta_p^2 = .06$. At the end of the experiment, the main effect of induction was no longer significant, $F(2, 74) = 1.14$, $p = .327$, $\eta_p^2 = .03$.

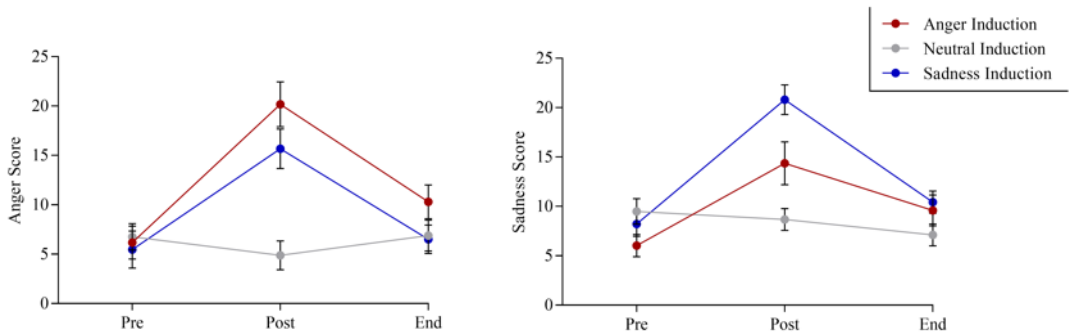

Figure 7.3. Scores on anger and sadness as a function of the moment of the measure for the three experimental groups. Error bars represent the standard error of the mean.

Table 7.1. Means (and standard deviations) of the scores in valence and arousal (from the SAM questionnaire), and anger and sadness (from the EVEA questionnaire) in the three experimental groups before and after the induction procedure, and at the end of the experiment.

		Anger	Neutral	Sadness
Pre-induction	Valence	6.39 (1.55)	6.16 (1.38)	6.46 (1.62)
	Arousal	4.39 (1.99)	4.32 (1.95)	4.00 (1.64)
	Anger	5.96 (8.59)	6.76 (6.67)	5.46 (9.23)
	Sadness	5.82 (5.83)	9.48 (6.53)	8.21 (6.13)
Post-induction	Valence	3.86 (1.92)	6.12 (1.62)	3.96 (1.76)
	Arousal	6.14 (2.19)	4.16 (1.63)	5.13 (1.75)
	Anger	20.11 (11.66)	4.88 (7.29)	15.67 (9.81)
	Sadness	14.32 (11.05)	8.68 (5.57)	20.79 (7.37)
End	Valence	5.14 (1.41)	5.88 (1.59)	5.58 (1.38)
	Arousal	5.07 (1.92)	4.52 (1.53)	4.46 (1.35)
	Anger	10.29 (8.70)	6.88 (7.87)	6.50 (7.11)
	Sadness	9.50 (7.97)	7.12 (5.55)	10.42 (5.61)

Transfer. Cooperation rates in the transfer phase were subjected to a mixed-design ANOVA with group behavior (equitable vs. non-equitable) as a within-participants variable and induction (anger vs. neutral vs. sadness) as a between-participants factor. The main effect of group behavior was significant, $F(1, 74) = 10.73$, $p = .002$, $\eta_p^2 = .13$, indicating that participants cooperated more with partners belonging to the group that was equitable in the learning phase ($M = .68$, $SD = .18$) than with partners belonging to the group that was non-equitable in the learning phase ($M = .61$, $SD = .21$). As shown in Figure 7.4, the pattern of categorization was different for the three groups, although the expected Group Behavior x Induction interaction was not significant, $F(2, 74) = 1.49$, $p = .232$, $\eta_p^2 = .04$. Nevertheless, we decided to examine the effect of the group behavior variable in each experimental group, to directly test our hypotheses. As expected, the group induced with anger cooperated more with partners from the equitable group ($M = .68$, $SD = .18$) than with partners from the non-equitable group ($M = .57$, $SD = .22$), $F(1, 27) = 6.55$, $p < .016$, $\eta_p^2 = .20$. This effect was also significant in the group under neutral induction, $F(1, 24) = 6.74$, $p < .016$, $\eta_p^2 = .22$, as participants cooperated more with partners from the equitable group ($M = .70$, $SD = .15$) than with partners from the non-equitable group (M = .63, SD = .19). Interestingly, the group induced with sadness showed no such difference, as they did not significantly differ in their cooperation between partners from the equitable (M = .64, SD = .23) and the non-equitable (M = .62, SD = .21) group, $F(1, 23) = .415$, $p < .526$, $\eta_p^2 = .02$.

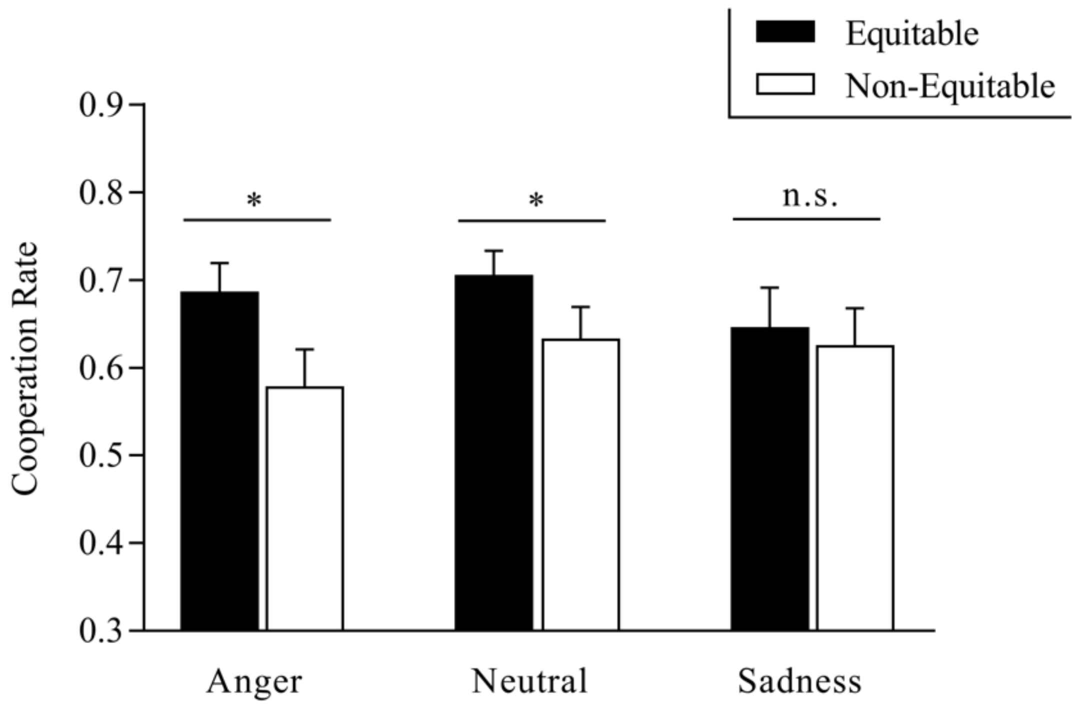

Figure 7.4. Cooperation rate in the transfer phase with partners belonging to the group that was equitable in the learning phase and to the group that was not equitable in the learning phase, for the three induction groups. Error bars represent the standard error of the mean.

Our induction paradigm was not completely specific, and both anger and sadness increased after the two emotional inductions, as compared to the neutral induction. This result is not surprising, as changes in more than one discrete emotion are commonly observed when eliciting emotions of negative valence (Lench, Flores, & Bench, 2011). However, we decided to further examine the differences in the transfer phase depending on the emotional induction (the neutral group was not considered for this analysis). Therefore, we computed an induction index allowing us to discriminate between participants who were induced more anger than sadness, and conversely, participants who were induced more sadness than anger. This index was computed in two steps: first we subtracted the pre-induction from the post-induction scores in anger and sadness respectively, to quantify the change induced in each emotion. Then, we subtracted the change in anger from the change in sadness, resulting in positive values for participants induced with more sadness than anger, and negative values for participants induced with more anger than sadness. To have a sense of the categorical transfer from the learning phase to new individuals, we computed a categorization index by subtracting cooperation rate with new partners from the equitable group from cooperation rate with new partners from the non-equitable group. The more participants categorize new partners, the higher the categorization index score. A bivariate correlation coefficient was computed to assess the relationship between the induction index and the categorization index in the transfer phase. There was a marginal negative correlation between the two variables, $r(77) = -.21$, $p = .072$, indicating that the more participants were induced with sadness, the less they categorized. Note that the same analysis conducted only with participants who categorized their partners in the learning phase (i.e., participants who cooperated more with partners from the equitable vs. non-equitable in the inconsistent condition in the fifth block), and therefore the ones who might be prone to categorize in the transfer phase, this correlation was significant, $r(43) = -.35$, $p = .022$.

Discussion

The present study explored social perception in interpersonal settings. First, we presented participants with black and white unfamiliar partners to test whether they would spontaneously cooperate more with one of the ethnic groups. Second, we manipulated the partners' cooperative behavior such that participants could use both individual and categorical information to predict their behaviors. After an emotional induction phase, we tested whether participants induced with anger would express more categorical strategies than those induced with sadness. Several conclusions may be drawn from the different phases of the experiments.

First, results from the baseline indicated that white participants spontaneously cooperated more with black than with white partners. Although these data may seem surprising because of their inconsistency with the well documented ingroup favoritism (i.e., the tendency to favor people from one's own social group over people from a different group, Tajfel & Turner, 1979), they replicated previous research investigating white perceivers' cooperation strategies with black targets. In fact, Tortosa et al. (2013) and Telga et al. (2018) also observed that white participants cooperated more with black than with white partners in a trust game, when ethnicity did not predict the partners' behaviors. In these studies, the apparent more positive attitude towards black individuals was not associated with a more positive attitude toward Blacks at the implicit level.

Conversely, those participants either showed an implicit racial bias favoring Whites over Blacks (Tortosa et al., 2013) or a category-based learning for black but not white partners (Telga et al., 2018). These discrepancies led the authors to interpret white participants' inclination to cooperate more with black partners as an attempt to be perceived equalitarian and not to show prejudice-related biases (Maddux, Barden, Brewer, & Petty, 2005), consistently with research showing that explicit and implicit attitudes toward Blacks can be widely dissociated (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Kawakami, Dunn, Karmali, & Dovidio, 2009).

In the learning phase, we observed that participants learned about the contingencies established at the group level, as they accurately cooperated with consistent individuals according to the group behavior. In contrast, learning about inconsistent individuals was poorer, as participants did not differ in their cooperation between partners who were individually equitable, and those who were not. If participants had categorized their partners, they would have cooperated with inconsistent individuals in the exact same way as with consistent individuals, according to the group behavior. Conversely, if they had adopted an individuation strategy, they would have displayed opposite cooperation patterns with consistent and inconsistent individuals, cooperating more with inconsistent partners belonging to the non-equitable group, and less with inconsistent partners belonging to the equitable group, as happened in previous studies when cooperating with only 8 individuals (Telga et al., 2018). However, the observed pattern of data suggests that participants fell somewhere between the two strategies, noticing the inconsistency of those individuals, but lacking either motivation or resources to get a full assessment of their individual behaviors. This interpretation perfectly fits in the continuum model described by Fiske and Neuberg (1990). After the initial categorization, participants motivated by the economic inter-dependence (Neuberg & Fiske, 1987) may have attended individual attributes. Because the individual attributes of inconsistent individuals did not match the categorical information, these partners needed to be re-categorized or individuated. These processes were not successful, broadly resulting in a poor learning about inconsistent partners. Alternatively, considering individual differences, it is possible that some participants might have taken mainly a categorization strategy, whereas others might have taken mainly an individuation strategy.

Even though the three experimental groups did not differ in their learning in the last block of the learning phase, an unexpected four-way interaction indicated that before the induction, the three experimental groups' learning varied across blocks. Notably, these differences cannot explain the pattern of data observed in the transfer phase. In fact, the group induced with anger and the neutral induction group are those who showed a better learning about inconsistent individuals, as shown in Figure 7.2. Therefore, if any, those groups would be the ones more likely to show an individuation strategy in the transfer phase. However, and in line with our hypothesis, we observed the opposite pattern of data in the transfer phase, as participants induced with sadness tended to use heuristics, and therefore categorize individuals, to a lesser extent than those induced with anger.

This pattern, albeit weak, is consistent with previous research associating sadness with deep reasoning and attention to details and anger with the need to make quick decisions (Bodenhausen et al., 1994). One limitation that may explain why the observed pattern of results is not very strong is that the induction procedure of the present research, although differential, was not completely specific. Participants from the groups induced with anger and sadness reported increased scores on both anger and sadness emotions. Also, although participants from the anger and sadness induction groups reported lower valence scores in the post- compared to pre-induction measures, consistently with the negative valence of both emotions, they did not clearly differ on their post-induction measures of arousal. Because anger is associated with higher arousal levels than sadness (Clark, Milberg, & Erber, 1984), the lack of difference between the two experimental groups in terms of arousal is rather surprising. This issue directly impacts our hypothesis, as higher arousal levels (as commonly observed in anger compared to sadness) have been related to more category-related stereotypic judgments (Bodenhausen, 1993). Therefore, it is important that future studies employ an induction procedure allowing for clearer differences between the experimental groups of interest, both in terms of discrete emotion induction but also regarding the arousal dimension underlying the elicitation of these emotions.

Despite the data are consistent with our hypotheses, they should be replicated in future studies addressing the aforementioned limitations. It may also be interesting to manipulate when to introduce the emotional induction, to test its impact on different stages of the impressions formation processes. If participants undergo the emotional induction before the baseline, we could verify its impact on naïve participants' decisions at zero acquaintance. Alternatively, if they are induced after the baseline but before the learning phase, we could explore how a specific emotional state may bias learning strategies towards individual vs. categorical cues. A unique experimental design manipulating the moment when participants are induced with a specific emotional state would inform of whether its impact is larger on the formation of first impressions, the acquisition of learning or the expression of this learning with new individuals.

Despite its limitations, these data also extends previous research, by showing that the greater use of heuristics under anger is not necessarily specific to making predictions about outgroup members (Bodenhausen et al., 1994), but also occurs when making predictions about ingroup members. Overall, this study provides promising data on the impact of emotional states on social perception in interpersonal settings.

Acknowledgments and funding

We are thankful to all the volunteers who took part in the study, and to Leticia Galdón Pascual and Raquel Lezama Cantero for their help in data collection.

This work was supported by the Spanish Ministry of Education, Culture and Sports, with pre-doctoral FPU fellowship FPU14/07106 to MT, and the Spanish Ministry of Economy and Competitiveness, with research projects PSI2014-52764-P and PSI2017-84926-P to JL. This research is part of MT's thesis dissertation under the supervision of JL.

Correspondence concerning this article should be addressed to Maïka Telga, Departamento de Psicología Experimental, Universidad de Granada, Campus de Cartuja s/n, 18 071 Granada, Spain. Email: mtelga@ugr.es

CHAPTER 8. General Discussion

Social interactions are at the heart of humans' life. The benefits of harmonious relationships have been extensively demonstrated in numerous domains related to health and well-being (Abbey, Andrews, & Halman, 1991; Dyck & Holtzman, 2013; Entwistle, Carter, Cribb, & McCaffery, 2010; Lynch et al., 2008). Daily social interactions are an essential part of our life since birth, so we are equipped with different cognitive tools allowing us to make sense of the social world. Notably, understanding others' affective states, expectations and goals is essential for satisfactory relationships.

Research in social psychology has demonstrated that social categorization and individuation processes are the main strategies to form impressions of others (Brewer, 1988). These strategies are not exclusive but lie on a continuum reflecting to what extent we attend others' individual or categorical attributes (Fiske, Lin, & Neuberg, 1999; Fiske & Neuberg, 1990). While social categorization consists of using information related to group membership to predict others' behaviors, individuation requires a costlier integration of their unique individual attributes. Several factors related to both the social context and the perceiver influence the impression formation processes (Macrae, Milne, & Bodenhausen, 1994; Neuberg & Fiske, 1987).

The present work aimed at investigating the factors impacting inferences about others at zero acquaintance and across repeated interactions. To achieve this goal, we adapted the trust game by creating a task setting in which participants had to make predictions about the cooperative behavior of unfamiliar partners across several trials. Specifically, we provided participants with category- and individual-based information, and explored to what extent social categorization and individuation processes impacted their judgments. Across seven experimental series, several factors were explored. Notably, we examined the main dimensions of social categorization (i.e., gender, ethnicity and age) in Experimental Series 1 and 2. We analyzed the effect of power on social learning strategies in Experimental Series 3. We directly manipulated the principal perceivers-related attributes impacting social perception (i.e., motivation and cognitive resources) in Experimental Series 4 and 5. Finally, in Experimental Series 6 and 7, we explored the relationship between emotions and social perception.

Decades of research have established that social categorization is the default strategy to make sense of others because of its resources-saving function (Macrae & Bodenhausen, 2001; Macrae et al., 1994). In fact, past and present theorizing on social perception would predict that people mostly adopt a categorization strategy, that may be furthered to individuation processes under specific circumstances related to cognitive ease and high motivation (Cañadas, Rodríguez-Bailón, Milliken, & Lupiáñez, 2013; Cloutier, Mason, & Macrae, 2005; Fiske & Neuberg, 1990). Nonetheless, the present research rather suggests that in some contexts, goal achievement largely prevails not only over cognitive economy but even over cognitive efficiency.

1. Attention to categorical features. When do we categorize others?

Being adapted in a social world means not only to be able to make accurate predictions about others but also to make them quickly and under complex circumstances such as time pressure or multi-tasking. In these situations, social categorization is the cognitive tool allowing heuristic-based fast decisions (Bodenhausen, Kang, & Peery, 2012; Macrae & Bodenhausen, 2000). The use of categorical information to predict others has been extensively reported (Kawakami, Amodio, & Hugenberg, 2017; Leyens, Yzerbyt, & Schadron, 1994; Macrae & Bodenhausen, 2001; McGarty, Yzerbyt, & Spears, 2002; Quinn & Macrae, 2005) even in situations in which targets possess traits inconsistent with their group membership (Fiske et al., 1999). Category-related information guides first interactions (Freeman & Ambady, 2011; Le Pelley et al., 2010), and categorical judgments may be further improved, for instance, across repeated interactions.

1.1. Categorization during learning

Despite the widely reported predominance of categorical thinking in impression formation, the first striking result of the present work is that across fourteen experiments, we never found a complete pattern of categorization in learning about human partners, which would have been reflected in the use of categorical information to predict inconsistent partners' behaviors, as reported by Cañadas, Rodríguez-Bailón, and Lupiáñez (2015). The only instance in which we observed categorical learning was when participants made predictions about non-social stimuli (Experimental Series 5), as participants' learning about inconsistent artificial races and paintings was mostly driven by the associations established at the categorical level. These results suggest that motivational factors encouraged individuation of human targets across repeated interactions.

Some hint of categorization was observed in a much subtler way, when ethnicity was manipulated. In fact, in Experimental Series 1, white participants showed an outgroup homogeneity effect (Park & Rothbart, 1982), being impacted by categorical information to a greater extent when learning about inconsistent black compared to inconsistent white partners. This pattern seems to be specific to ethnic categories as it was not replicated with gender (Experimental Series 1 and 3) or age (Experimental Series 2). The outgroup homogeneity effect observed manipulating ethnicity dimension may be the result of participants' reduced motivation to individuate black partners. Alternatively, it is also possible that participants' poorer learning about ethnic outgroup members was determined by a poorer capacity to discriminate between these individuals (Tanaka, Kiefer, & Bukach, 2004), and that individuation was at its best taking into account participants' discrimination ability.

However, we consider that this interpretation may not fully account for the pattern of data observed in Experimental Series 1 for two reasons. First, differences in discrimination between ingroup and outgroup faces may be simply explained by differential attention to these faces, as reflected in the other-race effect (ORE, Hugenberg, Young, Bernstein, & Sacco, 2010) described in Chapter 1. Therefore, if the outgroup homogeneity effect is based on differences in face processing between ingroup and outgroup members, it has to be considered that this difference may be determined itself by motivational factors. As argued by Hills, Pake, Dempsey, and Lewis (2018), what is often interpreted as an effect of exposure may actually be a result of motivation, as exposure itself (i.e., and the associated expectations of future interactions) may motivate people to fully attend targets' individual features. In this sense, participants from Experimental Series 1 may have low expectation of interacting with a scarce population of black individuals (i.e., unless in very specific settings such as this experiment), and therefore, may have lacked the motivation to explore black partners' individual characteristics.

Second, and more importantly, data collected in a heterogeneous ethnic context support the motivational account of the outgroup homogeneity effect in the trust game. In fact, Telga, Kawakami, and Lupiáñez (2019) observed similar effects in an unpublished study collected in Toronto, where the multi-ethnic population allows interactions with black and white individuals on a daily basis. Interestingly, in this context in which exposure to black and white people should be similar, both black and white ethnic groups showed an outgroup homogeneity effect across repeated interactions. Specifically, black participants learned less about inconsistent white compared to inconsistent black partners. Conversely, white participants learned less about inconsistent black compared to inconsistent white counterparts. Altogether, these findings suggest that the reduced exposure to black individuals in the context in which Experimental Series 1 was collected is not sufficient to explain the impaired learning about inconsistent outgroup members, making the motivational accounts of the outgroup homogeneity effect quite convincing.

Differences between ethnicity on the one hand, and gender and age on the other hand, may be understood in terms of salience. In contrast to ethnicity, interactions with age or gender outgroups occur on a daily basis, even within the family nucleus. However, such intimate interactions with ethnic outgroup members are far less likely, especially in an ethnically homogeneous context. Therefore, it is possible that using an infrequent outgroup broadly made social categories more salient as compared to gender and age manipulations (see, for instance, Fiske, 1980; Mullen, Brown, & Smith, 1992)

The use of categorical information was also observed in Experimental Series 6, when emotional expressions were used as predictive categories of partners' behaviors. Importantly, the procedure used in this experimental series differs from the one employed in the aforementioned experiments in that categorical knowledge was not to be applied to inconsistent partners. In fact, participants played with partners whose emotional expression predicted their cooperative behavior, being equitable when they portrayed anger, and non-equitable when they portrayed happiness. Therefore, in this task setting with only category-consistent partners, a categorization strategy would not have hindered but increased learning, as individual identities were irrelevant to predict partners' behaviors. However, we observed that participants playing with 4 partners used categorical information to a lesser extent (if any) than participants playing with 32 partners, the former showing, in consequence, an impaired learning. Hence, it is interesting that categorical information was only attended in the complete unfeasibility of making predictions from individual identities. Namely, emotional categories were really attended when individual identities were irrelevant to perform the task *and* participants were playing with a larger number of partners that made very difficult, if not impossible, to individually process all of them. Altogether, these data suggest that participants tried to individuate their partners and finally gave up, allowing a categorization strategy when individuation was impossible. Oddly enough, this interpretation goes in the opposite direction of theorizing arguing that people first categorize, and finally adopt an individuation strategy under specific circumstances (Fiske & Neuberg, 1990).

1.2. Categorization at zero acquaintance

We also evaluated strategies of impression formation at zero acquaintance in two phases: the baseline and the transfer phase. The baseline was the first phase of the studies, such that in this phase, the expression of categorical knowledge should be impacted by participants' real-life beliefs (e.g., stereotypes) and attitudes (e.g., ingroup favoritism). However, the transfer phase was performed after the learning phase, in which we established specific associations between social categories and cooperative behaviors. Therefore, in the transfer phase, we expected categorization processes to be influenced by the artificial associations established during learning, although real-life beliefs may also, at least partially, impact categorization in the transfer phase. Because during learning participants could also learn that all category members do not adopt the same cooperative behaviors, we broadly expected a larger expression of categorical knowledge in the baseline than in the transfer, and the data supported this general prediction.

In fact, we did observe categorization in the baseline on ethnicity and gender dimensions (Experimental Series 1, 5 and 7), but not age (Experimental Series 2). Interestingly, the use of social categories in the baseline did not seem to reflect stereotypes application, but rather a motivated ingroup or outgroup favoritism elicited by different specific motives. Notably, white participants cooperated more with black than with white partners (Experimental Series 1 and 7), a pattern clearly inconsistent with ethnic stereotypes associating black individuals with threat (Quillian & Pager, 2001). Moreover, this pattern does not fit either with the ingroup favoritism (Turner, Brown, & Tajfel, 1979) as white participants chose to cooperate more with outgroup than ingroup partners. Therefore, the most plausible interpretation is related to participants' motivation not to be perceived as prejudiced (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997). In fact, racist attitude is strongly condemned, at least at the explicit level (Kawakami, Dunn, Karmali, & Dovidio, 2009), and people often use different strategies to meet the requirements of social desirability (Lowery, Hardin, & Sinclair, 2001). In that sense, being cooperative with black people at first, in spite of ethnic stereotypes, may have seemed an adequate strategy to distance oneself from prejudiced biases, resulting in artificially high cooperation rates with black partners in the baseline phase of the trust game (see Tortosa, Lupiáñez, & Ruz, 2013 for a similar argument). This interpretation is particularly convincing taking into account the scarcity of the black population in the context where these experiments were conducted, as participants may have believed that the purpose of this study was to investigate ethnicity-related prejudice. Supporting this argument, in the unpublished experiment mentioned in the previous section, we observed that in a more heterogeneous ethnic context such as Toronto, participants do not show any ethnic bias in the baseline of the trust game (Telga et al., 2019).

As for gender, female participants cooperated more with female than with male partners (Experimental Series 1 and 5), which has been related to ingroup favoritism (instead of gender stereotypes) in Experimental Series 3. Specifically, in a series of meta-analyses comparing male and female participants' cooperation decisions at zero acquaintance, we concluded that women were more prone to cooperate with female partners, while men did not differ in their cooperation between male and female partners. This difference speaks against the possibility that participants' cooperation decisions were driven by gender stereotypes as, if so, men should have also been biased to cooperate more with women than with men. Therefore, the use of categorical information in the baseline by means of a more positive attitude toward ingroup compared to outgroup members is likely driven by female participants' values and their desire to enhance their social identity in a context in which they may have perceived themselves as low-status participants (Ellemers & Haslam, 2012; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). Importantly, for both ethnicity and gender, the use of category-related information seems to be driven by motivational concerns rather than cognitive economy, even promoting responses inconsistent with stereotypes. The cognitive effort necessary to inhibit stereotypes and produce counter-stereotypical responses (Macrae & Bodenhausen, 2000) in the case of ethnicity completely challenges the idea of using social categorization for its resources-saving function.

Results from the transfer phase, in which participants were tested on their use of heuristics after learning specific associations between social categories and behaviors, confirmed the impact of motivation on categorical judgments at zero acquaintance. In fact, new individuals were categorized on the basis of previous learning only in three instances: when the stimuli were not social (Experimental Series 5), when participants felt powerful over gender outgroup members (Experimental Series 3), and when participants were induced with anger (Experimental Series 7). Interestingly, all these effects are better explained in terms of motivation than in terms of cognitive cost. In fact, in Experimental Series 5, we observed that when motivation related to personal value was low (i.e., when playing with non-social stimuli), categorization occurs in the transfer phase both when those stimuli were easy (i.e., paintings) and hard (i.e., artificial races) to discriminate. In Experimental Series 3, we observed that male participants categorized women in the transfer phase although they were perfectly able to discriminate them and to treat them as individuals in the learning phase. Finally, the comparison between inductions of anger, sadness and neutral emotional states in social perception suggests that categorization occurring under anger is motivated by the appraisal of the situation, prompting quick decisions for the restoration of control (Bodenhausen, 1993; Bodenhausen, Sheppard, & Kramer, 1994), but does not respond to cognitive resources constraints (Experimental Series 7).

When participants did not use the associations established in the learning phase to make predictions about new individuals in the transfer phase, they either showed the aforementioned motivated ingroup favoritism (Experimental Series 4 and 5), or treated new individuals as such, without any categorical bias (Experimental Series 1, 2, 3, 6 and 7). In fact, in Experimental Series 4 and 5 we observed that in the transfer phase, women cooperated more with female partners than with their male counterparts. In the transfer phase of Experimental Series 1, 2, 3, and 6, participants did not differ in their cooperation with the two social categories manipulated.

Overall, the use of categorical information at zero acquaintance reflected participants' (lack of) motivation to individuate, either triggered by social norms, social identity concerns, power or appraisal, but did not seem to depend on the cognitive resources needed to achieve an individual assessment of targets. In fact, in Experimental Series 4 and 6 in which we could directly compare participants' cooperation decisions as a function of the cognitive resources needed to individuate, the groups under higher cognitive load (i.e., the groups playing with 32 and 64 partners in Experimental Series 4, and the group playing with 32 partners in Experimental Series 6) never used the categorical knowledge acquired in the learning phase to make decisions about their new partners at zero-acquaintance. These results echo the findings from the learning phase, as participants' learning patterns consistently indicate a high motivation to individuate human targets, only using categorical clues when the outgroup was particularly salient (Experimental Series 1) or when participants could not do otherwise (Experimental Series 5 and 6). Importantly, when the outcome-dependency was removed in the memory tests, in which participants were not rewarded according to their performance, the impact of categorical thinking was salient in all conditions, including those in which participants playing with only 8 partners were perfectly able to individuate them (Experimental Series 4).

2. Attention to individual features. When do we individuate others?

Accurate predictions about others' behaviors require more analytical processing with a piece-meal integration of a diversity of individual features (Fiske & Neuberg, 1990). This strategy focuses on the uniqueness of individuals, and hence requires to process them beyond category-related knowledge, to inhibit stereotypic information, and to update our knowledge as we receive new information (Macrae & Bodenhausen, 2000). Because of the cognitive cost of such a strategy as compared to social categorization (Macrae, Hewstone, & Griffiths, 1993; Macrae et al., 1994), people fail to individuate in numerous circumstances, with important implications for intergroup and interpersonal relationships (Kawakami et al., 2017). However, in the present research, individuation seemed to be the default strategy to process others, even when persistence to individuate resulted in flawed decisions and impaired learning.

Indeed, attention to individual identities instead of social categories have been found in numerous instances of the present dissertation, across repeated interactions but also at zero acquaintance, as long as participants were outcome-dependent on making accurate predictions about their partners (i.e., in the trust game but not in the memory tests). In first interactions, individuation was favored over age (Experimental Series 2) and in one occasion gender (Experimental Series 3) categorization, to make inferences about unfamiliar partners. In fact, in these experimental series, participants did not use social categorization to decide whether or not to cooperate with their partners in the baseline. Moreover, after learning, categorical features were generally ignored (with evidence in Experimental Series 1, 2, 3, 6 and 7). Similarly, patterns of individuation were consistently found for all human stimuli in the learning phase, in line with the outcome-dependency argument.

Furthermore, individuation attempts were also found far beyond expectations, when individuation was not functional to perform the task, or even when it hindered learning. For instance, efforts to individuate were found when participants played with a large number of partners, up to 64 different individuals, and the evident cost of individuating such a large number of people did not prompt a resources-saving categorization strategy. As a consequence, when participants played with 64 partners, learning about both consistent and inconsistent partners was hindered compared to the groups playing with a smaller number of partners with lesser cognitive demands (Experimental Series 4). In a possibly even more surprising case, we also observed that attention was not withdrawn from individual identities to categorical information even when the former were absolutely not predictive of people's behaviors. This is what happened in Experimental Series 6 when participants played with a few partners whose behaviors were manipulated according to their emotional expressions (and independently of their individual identities in Experiments 1 and 2). It is only when, in addition to individual identities not being predictive of partners' behaviors, participants were dealing with a large number of partners, that categorical learning emerged. However, this categorical learning was still impaired by some attention to individual identities. If participants playing with a large number of partners had exclusively focused on the two predictive emotional categories, they would have learned more than participants who were individuating 4 partners, which was not observed. These data also challenged the classical argument of automatic emotion processing (Tracy & Robins, 2008), as in this context, participants failed to respond to their partner's emotional expression. Altogether, these findings suggest that motivation to individuate was high, and neither the cognitive demands of individuating, nor the evident alternative categorization strategy discouraged participants from individuating.

Economic rewards are an efficient way to motivate people to overcome their category-based biases (Kawakami et al., 2014). This argument is supported by the larger influence of category-related features in non-economic tasks such as the memory tests that participants performed in Experimental Series 4 and 5, as compared to their individual-based strategies in the trust game. The particular settings of the trust game, intrinsically related to monetary rewards and individual learning, has likely been the core motivation for individual-based decision making. One may argue that such motivation to individuate is rather unlikely in real life, which limits the generalizability of the current findings. However, and importantly, the relevance of the results of the present dissertation does not lie in *why* people may or may not individuate at any cost, but in *how* they act under circumstances promoting individuation. The present research suggests that people engage all the cognitive resources needed to achieve the proposed goal, overcoming previous biases and easy-to-process categorical cues. They might also miss important category-related information that could have been easier to process and use. And more importantly, they do not abandon the individuation strategy. Perceivers possess a remarkable ability to make individuated judgments, even in cognitively demanding situations. Social policies making individuation "worth the effort" may be the missing motivating agent for perceivers to fully exploit their capacity to individuate.

3. Implications and further directions

Decades of research in social psychology have extensively described the positive and negative sides of categorical thinking. Some researchers have referred to categorical thinking as an "elegant mental tool" (Quinn & Macrae, 2005, p. 467), with emphasis on the associated cognitive efficiency (Macrae & Bodenhausen, 2000; Macrae & Bodenhausen, 2001), while less benevolent theorizers have compared perceivers to "cognitive misers" (Fiske & Taylor, 1984) or "mental sluggards" (Gilbert & Hixon, 1991, p. 509) for relying on category-based information. A more conciliatory approach grants a substantial weight to motivational factors, considering perceiver as "motivated tacticians" (Fiske & Taylor, 1991), who choose among a variety of processing strategies the one that best fits their current goal. However, unlike previous research indicating that the motivated tactician is limited by goals incompatibility (Ruscher, Fiske, & Schnake, 2000) or cognitive resources (Pendry & Macrae, 1994), one key finding from the seven experimental series of the present dissertation is that, when people are highly motivated to make accurate predictions about unknown individuals, they engage all the cognitive resources needed to achieve this goal. Although we do not aim at contributing to labeling perceivers according to their strategies of social perception, we arguably believe that the present dissertation supports the idea that with the appropriate setting, motivated thinking overcomes categorical thinking in any possible way.

The impact of motivation on behaviors has been largely acknowledged on both social and non-social processes (Botvinick & Braver, 2015; Locke & Latham, 2013; Todd Maddox & Markman, 2010). In social contexts, motivation for accuracy results in deeper processing, active search for information, enhanced recall and consideration of multiple interpretations (Molden & Higgins, 2012). However, motivated processes have traditionally been considered within the constraints of cognitive resources (Fiske & Neuberg, 1990). The present research challenges this view by suggesting that instead of behaving within the constraints of limited capacity cognition, motivated perceivers may fully exploit this capacity in unforeseeable ways.

This argument echoes recent theorizing in non-social fields exploring the determinants of effortful cognitive processes. Notably, Botvinick and Braver (2015) integrated research from different fields to understand the determinants of the allocation of cognitive control, namely, the cognitive processes allowing goal driven behaviors and distractors inhibition. Observing that people do not always perform at the maximum of their possibilities, the authors sought to understand what does refrain people to allocate the required control to maximize their performance. They concluded that allocation of control is better explained by reward-based models considering the maximization of payoffs associated with performance in a task, rather than by resource-models considering the resources consumption of cognitive control processes. In fact, recent theorizing challenges the traditional view that depletion effects arise from a lack of resources, suggesting instead a motivational perspective. This approach finds support in neuroscientific evidence showing that depletion is not associated with a reduced cerebral activity, but with changes in brain activity, notably in rewards network (Botvinick, Huffstetler, & McGuire, 2009; Kurniawan, Guitart-Masip, Dayan, & Dolan, 2013). Accordingly, the mere unavailability of attentional resources does not explain the poor performance observed in depletion settings. Instead, the subjective cost of control allocation yields a cost-benefit analysis, whose outcome determines whether or not “it is worth” applying control. Social decisions may be driven by analogous processes. Data from the present dissertation suggest that categorization is not favored when participants with depleted attentional resources cannot do otherwise (in these contexts, they often try to keep individuating anyways), but that effortful individuation processes are applied when participants believe that the result will be worth it. This interpretation is consistent with evidence that incentives promote effortful processes, both in the present research and in cognitive control settings (e.g., Padmala & Pessoa, 2011).

Being motivation central in accounting for the results of the present dissertation, it is important to consider that Fiske and Neuberg (1990) highlighted three main motivating agents in impressions formation: target, personal values and third-party. Although motivation related to the targets have been fairly covered in the present work, as argued below, motivation associated with the presence of a third-party or perceivers’ personal values has not been fully explored, and need to be taken in consideration.

First, target-related motivation includes situations in which a perceiver is outcome-dependent on a target to achieve a particular goal (Neuberg & Fiske, 1987). For instance, participants were outcome dependent on the targets’ accurate evaluation to earn monetary rewards in the trust game, while it was not the case in the subsequent memory test in Experimental Series 4 and 5. In a difference instance, power differences are also intrinsically related to outcome-dependency in relation to a target, as powerholders generally control the resources of powerless individuals. Although we did not directly manipulate power in Experimental Series 3, the differences in status between men and women may have led them to appraise differently their power in the trust game. Men may have perceived that they controlled the partners' outcomes, as they were able to decide in the first stage of the trust game whether or not to send the initial money received. Women may have focused instead on their vulnerability once having sent the money to their partners. The fact that men are generally more confident in economic decisions making than women (Bordalo, Coffman, Gennaioli, & Shleifer, 2019) may have enhanced these differences. Importantly, we observed that low target-related motivation promoted categorization as men (but not women) categorized their partners in the transfer phase of Experimental Series 3. Conversely, high target-related motivation promoted individuation, as performance was impacted by individuation in the trust game, but by categorization in the memory tests.

Second, motivation related to personal values was manipulated in Experimental Series 5, by comparing the perception of humans versus non-social stimuli. In fact, because of epistemic motives related to one's desire to understand the social world, and the expectations to interact with humans (but not with artificial races or paintings) outside the lab, we expected human stimuli to be intrinsically more important for participants, and therefore to prompt more individuation than non-social stimuli. The data supported this hypothesis. However, we do not have a measure of participants' personal values when playing with humans in this and the other experimental series. Although we initially included explicit measures of sexism and racism in Experimental Series 1, they were not informative as participants systematically responded with extreme rates. However, and alternatively, measuring their implicit bias toward the social groups manipulated, or controlling their strategy-oriented motivation may be informative of the influence of personal values as motivating agent. In fact, Higgins and Molden (2003) argued that beyond outcome-dependency, people may be motivated regarding the strategies used to inform their judgments, favoring "righteous means" or manners that "feel right" for reaching their conclusions. Therefore, participants who value the use of an ethically appropriate strategy to predict their partners' behaviors may perform differently from participants who do not exhibit such motivation. In fact, this strategy-oriented motivation has been shown to impact information processing by means of counterfactual thinking or preference for accurate over fast information processing.

Third, aside from outcome-dependency on target and personal values, perceivers may also be motivated by an interdependency between them and a third-party. Because we readily attribute cognitive and affective states to others, an ability known as theory of mind (Frith & Frith, 1999), motivation may arise from the fear to be negatively evaluated by or the desire of approval of a third-party. Generally, people are more prone to individuate targets when they are accountable to others (Tetlock, 1983). This possibility was explored in Experimental Series 1 by comparing participants' performance with a black compared to a white experimenter. The logic was that participants should be more likely concerned by a negative evaluation from an outgroup as compared to an ingroup experimenter, and therefore should individuate to a greater extent with the former. The data did not support this hypothesis as participants broadly showed the same pattern of learning with a black and a white experimenter, qualified by an outgroup homogeneity effect. However, this question could have been more systematically explored by providing participants with different instructions, emphasizing that they are being evaluated or they will have to explain their learning strategies at the end of the task in the critical experimental group. Although we did not thoroughly explore the different motivating agents on social performance, we may assume that all of them yield important consequences for the strategic use of individuation or categorization, which may be further explored in future research.

Overall, although the actual limits of our cognition are reasonable constraints to predict judgments accuracy, we believe that according to the overall pattern of results observed in this dissertation, motivational factors outweigh cognitive efficiency. By reason of the cognitive economy argument, categorical reasoning has been admitted as a prominent and efficient cognitive strategy to organize the social world. However, findings indicating that perceivers may also fail to use categorical information under high motivation suggest that impression formation processes are more about how *willing* are perceivers to employ the attentional resources needed to individuate, than how *capable* they are. More importantly, the fact that category-related information may be ignored even at the cost of learning indicates that unfortunate social categorization may be drastically reduced by motivated thinking, and encourages motivation-oriented interventions in reducing social biases related to categorical thinking, including stereotypes and prejudice.

CHAPTER 9. References

Abbey, A., Andrews, F. M., & Halman, L. J. (1991). The Importance of Social Relationships for Infertile Couples' Well-Being. In A. L. Stanton & C. Dunkel-Schetter (Eds.), *Infertility: Perspectives from Stress and Coping Research* (pp. 61-86). Boston, MA: Springer US.

Adams Jr, R. B., Ambady, N., Macrae, C. N., & Kleck, R. E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and Emotion, 30*(2), 179-188. doi:10.1007/s11031-006-9020-2

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature, 393*(6684), 470-474. doi:10.1038/30982

Ahmed, A. M. (2007). Group identity, social distance and intergroup bias. *Journal of Economic Psychology, 28*(3), 324-337. doi:10.1016/j.joep.2007.01.007

Alguacil, S., Madrid, E., Espín, A. M., & Ruz, M. (2017). Facial identity and emotional expression as predictors during economic decisions. *Cognitive, Affective and Behavioral Neuroscience, 17*(2), 315-329. doi:10.3758/s13415-016-0481-9

Alguacil, S., Tudela, P., & Ruz, M. (2015). Ignoring facial emotion expressions does not eliminate their influence on cooperation decisions. *Psicológica, 36*(2), 309-335.

Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.

Ames, D. R., & Johar, G. V. (2009). I'll know what you're like when i see how you feel: how and when affective displays influence behavior-based impressions. *Psychol Sci, 20*(5), 586-593. doi:10.1111/j.1467-9280.2009.02330.x

Angie, A. D., Connelly, S., Waples, E. P., & Kligyte, V. (2011). The influence of discrete emotions on judgement and decision-making: A meta-analytic review. *Cognition and Emotion, 25*(8), 1393-1422. doi:10.1080/02699931.2010.550751

Arrow, K. J. (1974). *The limits of organization*. New York: Norton.

Bach, J. S. (1734). Piano Concerto No 1, BWV 1052.

Bailey, P. E., & Leon, T. (2019). A systematic review and meta-analysis of age-related differences in trust. *Psychology and Aging, 34*(5), 674-685. doi:10.1037/pag0000368

Bailey, P. E., Slessor, G., Rieger, M., Rendell, P. G., Moustafa, A. A., & Ruffman, T. (2015). Trust and trustworthiness in young and older adults. *Psychology and Aging, 30*(4), 977-986. doi:10.1037/a0039736

Balliet, D., & Van Lange, P. A. (2013). Trust, conflict, and cooperation: a meta-analysis. *Psychol Bull, 139*(5), 1090-1112. doi:10.1037/a0030939

Balliet, D., Wu, J., & De Dreu, C. K. (2014). Ingroup favoritism in cooperation: a meta-analysis. *Psychol Bull, 140*(6), 1556-1581. doi:10.1037/a0037737

Barnard, C. I. (1968). *The functions of the executive*. Cambridge: Harvard University Press.

Barragan-Jason, G., Cauchoix, M., & Barbeau, E. J. (2015). The neural speed of familiar face recognition. *Neuropsychologia, 75*, 390-401. doi:10.1016/j.neuropsychologia.2015.06.017

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122-142. doi:10.1006/game.1995.1027

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychol Sci, 18*(8), 706-712. doi:10.1111/j.1467-9280.2007.01964.x

Bijlstra, G., Holland, R. W., Dotsch, R., & Wigboldus, D. H. J. (2018). Stereotypes and prejudice affect the recognition of emotional body postures. *Emotion*. doi:10.1037/emo0000438

Billig, M., & Tajfel, H. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology, 3*(1), 27-52. doi:10.1002/ejsp.2420030103

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*(3), 242-261. doi:10.1207/S15327957PSPR0603_8

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of afrocentric facial features in criminal sentencing. *Psychol Sci, 15*(10), 674-679. doi:10.1111/j.0956-7976.2004.00739.x

Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and Afrocentric facial features in social judgments. *Journal of Personality and Social Psychology, 87*(6), 763-778. doi:10.1037/0022-3514.87.6.763

Blair, I. V., Judd, C. M., Sadler, M. S., & Jenkins, C. (2002). The role of Afrocentric features in person perception: judging by features and categories. *Journal of Personality and Social Psychology, 83*(1), 5-25. doi:10.1037/0022-3514.83.1.5

Blue, P. R., Hu, J., & Zhou, X. (2018). Higher status honesty is worth more: The effect of social status on honesty evaluation. *Front Psychol, 9*(MAR). doi:10.3389/fpsyg.2018.00350

Bodenhausen, G. V. (1993). Emotions, arousal, and stereotypic judgments: A heuristic model of affect and stereotyping. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, Cognition and Stereotyping* (pp. 13-37). San Diego: Academic Press.

Bodenhausen, G. V., Kang, S. K., & Peery, D. (2012). Social categorization and the perception of social groups. London: SAGE Publications Ltd. doi:10.4135/9781446247631

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. *Advances in Social Cognition, 11*, 1-52.

Bodenhausen, G. V., Sheppard, L. A., & Kramer, G. P. (1994). Negative affect and social judgment: The differential impact of anger and sadness. *European Journal of Social Psychology, 24*(1), 45-62. doi:10.1002/ejsp.2420240104

Bodenhausen, G. V., Todd, A. R., & Becker, A. P. (2006). Categorizing the Social World: Affect, Motivation, and Self-Regulation *Psychology of Learning and Motivation* (Vol. 47, pp. 123-155): Academic Press.

Bogliacino, F., Jiménez Lozano, L., & Reyes, D. (2018). Socioeconomic stratification and stereotyping: lab-in-the-field evidence from Colombia. *International Review of Economics, 65*(1), 77-118. doi:10.1007/s12232-017-0285-4

Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2019). Beliefs about Gender. *American Economic Review, 109*(3), 739-773. doi:10.1257/aer.20170007

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: from behavior to neural mechanism. *Annu Rev Psychol, 66*, 83-113. doi:10.1146/annurev-psych-010814-015044

Botvinick, M. M., Huffstetler, S., & McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cogn Affect Behav Neurosci, 9*(1), 16-27. doi:10.3758/cabn.9.1.16

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59. doi:10.1016/0005-7916(94)90063-9

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychol Bull, 86*(2), 307-324. doi:10.1037/0033-2909.86.2.307

Brewer, M. B. (1981). Ethnocentrism and its role in interpersonal trust. In M. B. Brewer & B. E. Collins (Eds.), *Scientific Inquiry and the Social Sciences* (pp. 345-359). San Francisco: Jossey-Bass.

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer, Jr (Eds.), *A dual process model of impression formation* (pp. 1-36). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Brewer, M. B. (2001). Ingroup identification and intergroup conflict: When does ingroup love become outgroup hate? *Social identity, intergroup conflict, and conflict reduction.* (pp. 17-41). New York, NY, US: Oxford University Press.

Brewer, M. B. (2007). The importance of being we: human nature and intergroup relations. *Am Psychol, 62*(8), 726-738. doi:10.1037/0003-066x.62.8.728

Brewer, M. B. (2008). Depersonalized trust and ingroup cooperation *Rationality and social responsibility: Essays in honor of Robyn Mason Dawes.* (pp. 215-232). New York, NY, US: Psychology Press.

Brown, E., & Perrett, D. I. (1993). What gives a face its gender? *Perception, 22*(7), 829-840. doi:10.1068/p220829

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*(3), 305-327. doi: 10.1111/j.2044-8295.1986.tb02199.x

Buchan, N. R., Croson, R. T. A., & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the Investment Game. *Journal of Economic Behavior & Organization, 68*(3), 466-476. doi:10.1016/j.jebo.2007.10.006

Bukowski, M., Moya, M., de Lemus, S., & Szmajke, A. (2009). Selective stereotype activation: The joint impact of interpersonal goals and task context. *European Journal of Social Psychology, 39*(2), 317-324. doi:10.1002/ejsp.585

Burkhardt, A., Blaha, L. M., Jurs, B. S., Rhodes, G., Jeffery, L., Wyatte, D., . . . Busey, T. (2010). Adaptation modulates the electrophysiological substrates of perceived facial distortion: support for opponent coding. *Neuropsychologia, 48*(13), 3743-3756. doi:10.1016/j.neuropsychologia.2010.08.016

Cañadas, E., Lupiáñez, J., Kawakami, K., Niedenthal, P. M., & Rodríguez-Bailón, R. (2016). Perceiving emotions: Cueing social categorization processes and attentional control through facial expressions. *Cognition and Emotion, 30*(6), 1149-1163. doi:10.1080/02699931.2015.1052781

Cañadas, E., Rodríguez-Bailón, R., & Lupiáñez, J. (2015). The effect of Social Categorization on Trust Decisions in a Trust Game Paradigm. *Front Psychol, 6*. doi:10.3389/fpsyg.2015.01568

Cañadas, E., Rodríguez-Bailón, R., Milliken, B., & Lupiáñez, J. (2013). Social categories as a context for the allocation of attentional control. *Journal of Experimental Psychology: General, 142*(3), 934-943. doi:10.1037/a0029794

Carragher, D. J., Thomas, N. A., & Nicholls, M. E. R. (2018). Is trustworthiness lateralized in the face? Evidence from a trust game. *Laterality, 23*(1), 20-38. doi:10.1080/1357650X.2017.1298120

Charness, G., & Villeval, M.-C. (2009). Cooperation and competition in intergenerational experiments in the field and the laboratory. *The American Economic Review, 99*(3), 956-978. doi:10.1257/aer.99.3.956

Chen, Y., & Sun, Y. (2003). Age differences in financial decision-making: Using simple heuristics. *Educational Gerontology, 29*(7), 627-635. doi:10.1080/713844418

Chua, K. W., Richler, J. J., & Gauthier, I. (2014). Becoming a Lunari or Taiyo expert: learned attention to parts drives holistic processing of faces. *J Exp Psychol Hum Percept Perform, 40*(3), 1174-1182. doi:10.1037/a0035895

Chua, K. W., Richler, J. J., & Gauthier, I. (2015). Holistic processing from learned attention to parts. *J Exp Psychol Gen, 144*(4), 723-729. doi:10.1037/xge0000063

Clark, M. S., Milberg, S., & Erber, R. (1984). Effects of arousal on judgments of others' emotions. *Journal of Personality and Social Psychology, 46*(3), 551-560. doi:10.1037/0022-3514.46.3.551

Cloutier, J., Mason, M. F., & Macrae, C. N. (2005). The perceptual determinants of person construal: Reopening the social-cognitive toolbox. *Journal of Personality and Social Psychology, 88*(6), 885-894. doi:10.1037/0022-3514.88.6.885

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83*(6), 1314-1329.

Cuddy, A. J., Fiske, S. T., Kwan, V. S., Glick, P., Demoulin, S., Leyens, J. P., . . . Ziegler, R. (2009). Stereotype content model across cultures: towards universal similarities and some differences. *British Journal of Social Psychology, 48*(Pt 1), 1-33. doi:10.1348/014466608x314935

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2004). When professionals become mothers, warmth doesn't cut the ice. *Journal of Social Issues, 60*(4), 701-718. doi:10.1111/j.0022-4537.2004.00381.x

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map *Advances in Experimental Social Psychology* (Vol. 40, pp. 61-149): Academic Press.

Cuddy, A. J. C., Norton, M. I., & Fiske, S. T. (2005). This old stereotype: The pervasiveness and persistence of the elderly stereotype. *Journal of Social Issues, 61*(2), 267-285. doi:10.1111/j.1540-4560.2005.00405.x

Darley, J. M., Fleming, J. H., Hilton, J. L., & Swann, W. B. (1988). Dispelling negative expectancies: The impact of interaction goals and target characteristics on the expectancy confirmation process. *Journal of Experimental Social Psychology, 24*(1), 19-36. doi:10.1016/0022-1031(88)90042-X de Lemus, S., Spears, R., & Moya, M. (2012). The power of a smile to move you: complementary submissiveness in women's posture as a function of gender salience and facial expression. *Pers Soc Psychol Bull, 38*(11), 1480-1494. doi:10.1177/0146167212454178

DeLozier, S., & Rhodes, M. G. (2015). The impact of value-directed remembering on the own-race bias. *Acta Psychologica, 154*, 62-68. doi:10.1016/j.actpsy.2014.11.009

Dépret, E., & Fiske, S. T. (1993). Social Cognition and Power: Some Cognitive Consequences of Social Structure as a Source of Control Deprivation. In G. Weary, F. Gleicher, & K. L. Marsh (Eds.), *Control Motivation and Social Cognition* (pp. 176-202). New York, NY: Springer New York.

Deutsch, M. (1973). *The resolution of conflict: Constructive and destructive processes*. New Haven: Yale University Press.

Díaz-Gutiérrez, P., Alguacil, S., & Ruz, M. (2017). Bias and control in social decision-making. In A. Ibañez, L. Sedeño, & A. García (Eds.), *Neuroscience and social science: the missing link*. Switzerland: Springer.

Dong, Y., Liu, Y., Jia, Y., Li, Y., & Li, C. (2018). Effects of facial expression and facial gender on judgment of trustworthiness: The modulating effect of cooperative and competitive settings. *Front Psychol, 9*(2022). doi:10.3389/fpsyg.2018.02022

Dovidio, J. F., Glick, P., & Rudman, L. A. (2005). *On the nature of prejudice: Fifty years after Allport*. Malden: Blackwell Publishing.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*(5), 510-540. doi:10.1006/jesp.1997.1331

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews, 6*(5), 178-190. doi:10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8

Dunbar, R. I. M., & Shultz, S. (2007). Evolution in the Social Brain. *Science, 317*(5843), 1344. doi:10.1126/science.1145463

Dunning, D., Anderson, J. E., Schlosser, T., Ehlebracht, D., & Fetchenhauer, D. (2014). Trust at zero acquaintance: more a matter of respect than expectation of reward. *Journal of Personality and Social Psychology, 107*(1), 122-141. doi:10.1037/a0036673

Dunning, D., & Fetchenhauer, D. (2010) Trust as an expressive rather than an instrumental act. *Vol. 27. Advances in Group Processes* (pp. 97-127).

Dunning, D., Fetchenhauer, D., & Schlösser, T. M. (2012). Trust as a social and emotional act: Noneconomic considerations in trust behavior. *Journal of Economic Psychology, 33*(3), 686-694. doi:10.1016/j.joep.2011.09.005

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23*(3), 316-326. doi:10.1177/0146167297233009

Dyck, K. T. H., & Holtzman, S. (2013). Understanding humor styles and well-being: The importance of social relationships and gender. *Personality and Individual Differences, 55*(1), 53-58. doi:10.1016/j.paid.2013.01.023

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing Black: Race, Crime, and Visual Processing. *Journal of Personality and Social Psychology, 87*(6), 876-893. doi:10.1037/0022-3514.87.6.876

Eimer, M., & Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia, 45*(1), 15-31. doi:10.1016/j.neuropsychologia.2006.04.022

Ekman, P. (2003). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. New York, NY, US: Times Books/Henry Holt and Co.

Ekman, P., & Cordaro, D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review, 3*(4), 364-370.

Ellemers, N. (2012). The group self. *Science, 336*(6083), 848-852. doi:10.1126/science.1220987

Ellemers, N., & Haslam, S. A. (2012). Social Identity Theory. In P. A. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), Handbook of Theories of Social Psychology: Volume 2. London: SAGE Publications Ltd. doi:10.4135/9781446249222

Engfors, L. M., Jeffery, L., Gignac, G. E., & Palermo, R. (2017). Individual differences in adaptive norm-based coding and holistic coding are associated yet each contributes uniquely to unfamiliar face recognition ability. *J Exp Psychol Hum Percept Perform, 43*(2), 281-293. doi:10.1037/xhp0000265

Entwistle, V. A., Carter, S. M., Cribb, A., & McCaffery, K. (2010). Supporting Patient Autonomy: The Importance of Clinician-patient Relationships. *Journal of General Internal Medicine, 25*(7), 741-745. doi:10.1007/s11606-010-1292-2

Ewing, L., Leach, K., Pellicano, E., Jeffery, L., & Rhodes, G. (2013). Reduced Face Aftereffects in Autism Are Not Due to Poor Attention. *PLoS One, 8*(11), e81353. doi:10.1371/journal.pone.0081353

Ewing, L., Pellicano, E., & Rhodes, G. (2013). Atypical updating of face representations with experience in children with autism. *Dev Sci, 16*(1), 116-123. doi:10.1111/desc.12007

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proc Natl Acad Sci U S A, 99*(14), 9602-9605. doi:10.1073/pnas.152159999

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods, 39*(2), 175-191.

Ferrando, P. J., & Chico, E. (2000). Adaptación y análisis psicométrico de la escala de deseabilidad social de Marlowe y Crowne. *Psicothema, 12*(3), 383-389.

Fine, H. C., Shing, Y. L., & Naveh-Benjamin, M. (2018). Effects of changes in schematic support and of item repetition on age-related associative memory deficits: Theoretically-driven empirical attempts to reduce older adults' high false alarm rate. *Psychology and Aging, 33*(1), 57-73. doi:10.1037/pag0000211

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology, 38*(6), 889-906. doi:10.1037/0022-3514.38.6.889

Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist, 48*(6), 621-628. doi:10.1037/0003-066X.48.6.621

Fiske, S. T. (2010). Interpersonal stratification: Status, power, and subordination *Handbook of social psychology, Vol. 2, 5th ed.* (pp. 941-982). Hoboken, NJ, US: John Wiley & Sons Inc.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878-902. doi:10.1037/0022-3514.82.6.878

Fiske, S. T., & Dépret, E. (1996). Control, interdependence and power: Understanding social cognition in its social context. *European Review of Social Psychology, 7*(1), 31-61. doi:10.1080/14792779443000094

Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 231-254). New York, NY, US: Guilford Press.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In P. Z. Mark (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 1-74): Academic Press.

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition (1st ed.)*. New York: Mcgraw-Hill Book.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition (2nd ed.)*. New York: Mcgraw-Hill.

Folstein, M. F., Robins, L. N., & Helzer, J. E. (1983). The Mini-Mental State Examination. *JAMA Psychiatry, 40*(7), 812-812. doi:10.1001/archpsyc.1983.01790060110016

FreeAudioMusic. (n.d.). Ambiente de bosque. Retrieved from https://www.freeaudiolibrary.com/es/efectos-gratuitos-ambientes/

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychol Rev, 118*(2), 247-279. doi:10.1037/a0022327

Freeman, J. B., Schiller, D., Rule, N. O., & Ambady, N. (2010). The neural origins of superficial and individuated judgments about ingroup and outgroup members. *Hum Brain Mapp, 31*(1), 150-159. doi:10.1002/hbm.20852

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. San Diego, CA, US: Academic Press.

Frith, C. D., & Frith, U. (1999). Interacting Minds--A Biological Basis. *Science, 286*(5445), 1692. doi:10.1126/science.286.5445.1692

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*(4), 509-517. doi:10.1037/0022-3514.60.4.509

Goffaux, V., & Rossion, B. (2006). Faces are "spatial"--holistic face perception is supported by low spatial frequencies. *J Exp Psychol Hum Percept Perform, 32*(4), 1023-1039. doi:10.1037/0096-1523.32.4.1023

Goldstein, A. G., Chance, J. E., & Gilbert, B. (1984). Facial stereotypes of good guys and bad guys: A replication and extension. *Bulletin of the Psychonomic Society, 22*(6), 549-552. doi:10.3758/BF03333904

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology, 45*(2), 410-414. doi:10.1016/j.jesp.2008.11.004

Goodwin, S. A., Gubin, A., Fiske, S. T., & Yzerbyt, V. Y. (2000). Power can bias impression processes: Stereotyping subordinates by default and by design. *Group Processes & Intergroup Relations, 3*(3), 227-256. doi:10.1177/1368430200003003001

Govorun, O., & Payne, B. K. (2006). Ego—depletion and prejudice: Separating automatic and controlled components. *Social Cognition, 24*(2), 111-136. doi:10.1521/soco.2006.24.2.111

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464-1480.

Guinote, A. (2007a). Power affects basic cognition: Increased attentional inhibition and flexibility. *Journal of Experimental Social Psychology, 43*(5), 685-697. doi:10.1016/j.jesp.2006.06.008

Guinote, A. (2007b). Power and Goal Pursuit. *Personality and Social Psychology Bulletin, 33*(8), 1076-1087. doi:10.1177/0146167207301011

Guinote, A., & Phillips, A. (2010). Power can increase stereotyping: Evidence from managers and subordinates in the hotel industry. *Social Psychology, 41*(1), 3-9. doi:10.1027/1864-9335/a000002

Hartstone, M., & Augoustinos, M. (1995). The minimal group paradigm: Categorization into two versus three groups. *European Journal of Social Psychology, 25*(2), 179-193. doi:10.1002/ejsp.2420250205

Haslam, S. A., Oakes, P. J., & Turner, J. C. (1996). Social identity, self-categorization, and the perceived homogeneity of ingroups and outgroups: The interaction between social motivation and cognition. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition, Vol. 3: The interpersonal context* (pp. 182-222). New York, NY, US: Guilford Press.

Hassin, R., & Trope, Y. (2000). Facing faces: studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology, 78*(5), 837-852. doi:10.1037//0022-3514.78.5.837

Hechler, S., Neyer, F. J., & Kessler, T. (2016). The infamous among us: Enhanced reputational memory for uncooperative ingroup members. *Cognition, 157*, 1-13. doi:10.1016/j.cognition.2016.08.001

Heisz, J. J., Pottruff, M. M., & Shore, D. I. (2013). Females Scan More Than Males: A Potential Mechanism for Sex Differences in Recognition Memory. *Psychol Sci, 24*(7), 1157-1163. doi:10.1177/0956797612468281

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33*(1), 98-106. doi:10.3758/BF03195300

Henrich, N., & Henrich, J. (2007). Why humans cooperate : a cultural and evolutionary explanation. In J. P. Henrich (Ed.). Oxford: Oxford University Press.

Higgins, E. T., & Molden, D. C. (2003). How strategies for making judgments and decisions affect cognition: Motivated cognition revisited *Foundations of Social Cognition: A Festschrift in Honor of Robert S. Wyer, Jr.* (pp. 211-235). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Hills, P. J., Pake, J. M., Dempsey, J. R., & Lewis, M. B. (2018). Exploring the contribution of motivation and experience in the postpubescent own-gender bias in face recognition. *Journal of Experimental Psychology: Human Perception and Performance, 44*(9), 1426-1446. doi:10.1037/xhp0000533

Hodsoll, S., Viding, E., & Lavie, N. (2011). Attentional capture by irrelevant emotional distractor faces. *Emotion, 11*(2), 346-353. doi:10.1037/a0022771

Hugenberg, K., & Wilson, J. P. (2013). Faces are central to social cognition *The Oxford handbook of social cognition.* (pp. 167-193). New York, NY, US: Oxford University Press.

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review, 117*(4), 1168-1187. doi:10.1037/a0020463

Ibañez, A., Gleichgerrcht, E., Hurtado, E., Gonzalez, R., Haye, A., & Manes, F. F. (2010). Early Neural Markers of Implicit Attitudes: N170 Modulated by Intergroup and Evaluative Contexts in IAT. *Front Hum Neurosci, 4*, 188. doi:10.3389/fnhum.2010.00188

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*(4), 616-626. doi:10.1037/0022-3514.85.4.616

Jacques, C., & Rossion, B. (2009). The initial representation of individual faces in the right occipito-temporal cortex is holistic: electrophysiological evidence from the composite face illusion. *J Vis, 9*(6), 8.1-16. doi:10.1167/9.6.8

Janik, S. W., Wellens, A. R., Goldberg, M. L., & Dell'Osso, L. F. (1978). Eyes as the center of focus in the visual examination of human faces. *Percept Mot Skills, 47*(3 Pt 1), 857-858. doi:10.2466/pms.1978.47.3.857

Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(1), 236-259. doi:10.1037/0278-7393.25.1.236

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology, 32*(5), 865-889. doi:10.1016/j.joep.2011.05.007

Judd, C. M., & Park, B. (1988). Out-group homogeneity: Judgments of variability at the individual and group levels. *Journal of Personality and Social Psychology, 54*(5), 778-788. doi:10.1037/0022-3514.54.5.778

Kalokerinos, E. K., von Hippel, W., & Henry, J. D. (2017). Social cognition and aging. In N. A. Pachana (Ed.), *Encyclopedia of Geropsychology* (pp. 2168-2174). Singapore: Springer Singapore.

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrativefFramework for understanding the causes and consequences of social categorization. In M. O. James (Ed.), *Advances in Experimental Social Psychology* (Vol. Volume 55, pp. 1-80). San Diego, CA, US: Elsevier Academic Press.

Kawakami, K., Dunn, E., Karmali, F., & Dovidio, J. F. (2009). Mispredicting affective and behavioral responses to racism. *Science, 323*(5911), 276.

Kawakami, K., Williams, A., Sidhu, D., Choma, B. L., Rodriguez-Bailon, R., Canadas, E., . . . Hugenberg, K. (2014). An eye for the I: Preferential attention to the eyes of ingroup members. *Journal of Personality and Social Psychology, 107*(1), 1-20. doi:10.1037/a0036838

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, Approach, and Inhibition. *Psychological Review, 110*(2), 265-284. doi:10.1037/0033-295X.110.2.265

Kim, S., Goldstein, D., Hasher, L., & Zacks, R. T. (2005). Framing effects in younger and older adults. *The Journals of Gerontology: Series B, 60*(4), P215-P218. doi:10.1093/geronb/60.4.P215

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science, 308*(5718), 78-83. doi:10.1126/science.1108062

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch Gen Psychiatry, 59*(9), 809-816. doi:10.1001/archpsyc.59.9.809

Kurniawan, I. T., Guitart-Masip, M., Dayan, P., & Dolan, R. J. (2013). Effort and valuation in the brain: The effects of anticipation and execution. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 33*(14), 6160-6169. doi:10.1523/jneurosci.4777-12.2013

Kutzner, F., & Fiedler, K. (2017). Stereotypes as Pseudocontingencies. *European Review of Social Psychology, 28*(1), 1-49. doi:10.1080/10463283.2016.1260238

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS) : affective ratings of pictures and instruction manual*. Gainesville, Fla.: NIMH, Center for the Study of Emotion & Attention.

Le Pelley, M. E., Reimers, S. J., Calvini, G., Spears, R., Beesley, T., & Murphy, R. A. (2010). Stereotype formation: biased by association. *J Exp Psychol Gen, 139*(1), 138-161. doi:10.1037/a0018210

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitations. *Psychol Bull, 137*(5), 834-855. doi:10.1037/a0024244

Lench, H. C., Tibbett, T. P., & Bench, S. W. (2016). Exploring the toolkit of emotion: What do sadness and anger do for us? *Social and Personality Psychology Compass, 10*(1), 11-25. doi:10.1111/spc3.12229

Leopold, D. A., Rhodes, G., Muller, K. M., & Jeffery, L. (2005). The dynamics of visual adaptation to faces. *Proc Biol Sci, 272*(1566), 897-904. doi:10.1098/rspb.2004.3022

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72*(2), 275-287. doi:10.1037/0022-3514.72.2.275

Lerner, J. S., Goldberg, J. H., & Tetlock, P. E. (1998). Sober Second Thought: The Effects of Accountability, Anger, and Authoritarianism on Attributions of Responsibility. *Personality and Social Psychology Bulletin, 24*(6), 563-574. doi:10.1177/0146167298246001

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(6), 1364-1382. doi:10.1037/0278-7393.22.6.1364

Levin, D. T. (2000). Race as a visual feature: using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *J Exp Psychol Gen, 129*(4), 559-574.

Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust and distrust: New relationships and realities. *The Academy of Management Review, 23*(3), 438-458. doi:10.2307/259288

Leyens, J.-P., Yzerbyt, V., & Schadron, G. (1994). *Stereotypes and social cognition*. Thousand Oaks, CA, US: Sage Publications, Inc.

Li, T., Liu, X., Pan, J., & Zhou, G. (2017). The interactive effect of facial appearance and behavior statement on trust belief and trust behavior. *Personality and Individual Differences, 117*, 60-65. doi:10.1016/j.paid.2017.05.038

Liang, J., Chen, Y.-H., Yan, W.-J., Qu, F., & Fu, X. (2018). Effects of task-irrelevant emotional information on deception. *Cognition and Emotion, 32*(6), 1265-1274. doi:10.1080/02699931.2017.1404967

Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). The origins of social categorization. *Trends Cogn Sci, 21*(7), 556-568. doi:10.1016/j.tics.2017.04.004

Linville, P. W., Fischer, G. W., & Salovey, P. (1989). Perceived distributions of the characteristics of in-group and out-group members: empirical evidence and a computer simulation. *Journal of Personality and Social Psychology, 57*(2), 165-188.

Linville, P. W., Salovey, P., & Fischer, G. W. (1986). Stereotyping and perceived distributions of social characteristics: An application to ingroup-outgroup perception. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 165-208). San Diego, CA, US: Academic Press.

Locke, E. A., & Latham, G. P. (2013). *New developments in goal setting and task performance*. New York, NY, US: Routledge/Taylor & Francis Group.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81*(5), 842-855. doi:10.1037/0022-3514.81.5.842

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, 91*, 630.

Lynch, E. B., Butt, Z., Heinemann, A., Victorson, D., Nowinski, C. J., Perez, L., & Cella, D. (2008). A Qualitative Study of Quality of Life After Stroke: The Importance of Social Relationships. *Journal of Rehabilitation Medicine, 40*(7), 518-523. doi:10.2340/16501977-0203

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behav Res Methods, 47*(4), 1122-1135. doi:10.3758/s13428-014-0532-5

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annu Rev Psychol, 51*(1), 93-120. doi:10.1146/annurev.psych.51.1.93

Macrae, C. N., & Bodenhausen, G. V. (2001). Social cognition: categorical person perception. *Br J Psychol, 92*(Pt 1), 239-255.

Macrae, C. N., Hewstone, M., & Griffiths, R. J. (1993). Processing load and memory for stereotype-based information. *European Journal of Social Psychology, 23*(1), 77-87. doi:10.1002/ejsp.2420230107

Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychol Sci, 13*(5), 460-464. doi:10.1111/1467-9280.00481

Macrae, C. N., & Martin, D. (2007). A boy primed Sue: feature-based processing and person construal. *European Journal of Social Psychology, 37*(5), 793-805. doi:10.1002/ejsp.406

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66*(1), 37-47. doi:10.1037/0022-3514.66.1.37

Maddux, W. W., Barden, J., Brewer, M. B., & Petty, R. E. (2005). Saying no to negativity: The effects of context and motivation to control prejudice on automatic evaluative responses. *Journal of Experimental Social Psychology, 41*(1), 19-35. doi:10.1016/j.jesp.2004.05.002

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion, 5*(1), 119-124. doi:10.1037/1528-3542.5.1.119

Martin, D., & Macrae, C. N. (2007). A face with a cue: exploring the inevitability of person categorization. *European Journal of Social Psychology, 37*(5), 806-816. doi:10.1002/ejsp.445

Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends Cogn Sci, 6*(6), 255-260.

McGarty, C., Yzerbyt, V. Y., & Spears, R. (2002). *Stereotypes as Explanations: The Formation of Meaningful Beliefs about Social Groups*. Cambridge: Cambridge University Press.

McKelvie, S. J. (1976). The role of eyes and mouth in the memory of a face. *The American Journal of Psychology, 89*(2), 311-323. doi:10.2307/1421414

McKone, E. (2008). Configural processing and face viewpoint. *J Exp Psychol Hum Percept Perform, 34*(2), 310-327. doi:10.1037/0096-1523.34.2.310

McKone, E. (2009). Integrating holistic processing and face-space approaches to the coding of facial identity. *Journal of Vision, 9*(8), 539-539. doi:10.1167/9.8.539

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*(1), 3. doi:10838250

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers, 36*(4), 630-633.

Molden, D. C., & Higgins, E. T. (2012). Motivated thinking *The Oxford handbook of thinking and reasoning.* (pp. 390-409). New York, NY, US: Oxford University Press.

Montepare, J. M., & Zebrowitz, L. A. (1998). Person Perception Comes of Age: The Salience and Significance of Age in Social Judgments. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 30, pp. 93-161): Academic Press.

Moran, J. M., Jolly, E., & Mitchell, J. P. (2012). Social-cognitive deficits in normal aging. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 32*(16), 5553-5561. doi:10.1523/JNEUROSCI.5511-11.2012

Mouchetant-Rostaing, Y., & Giard, M. H. (2003). Electrophysiological correlates of age and gender perception on human faces. *J Cogn Neurosci, 15*(6), 900-910. doi:10.1162/089892903322370816

Mouchetant-Rostaing, Y., Giard, M. H., Bentin, S., Aguera, P. E., & Pernier, J. (2000). Neurophysiological correlates of face gender processing in humans. *Eur J Neurosci, 12*(1), 303-310. doi:10.1046/j.1460-9568.2000.00888.x

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22*(2), 103-122. doi:10.1002/ejsp.2420220202

Mussel, P., Göritz, A. S., & Hewig, J. (2013). The value of a smile: Facial expression affects ultimatum-game responses. *Judgment and Decision Making, 8*(3), 381-385.

Neuberg, S. L. (1989). The goal of forming accurate impressions during social interactions: attenuating the impact of negative expectancies. *Journal of Personality and Social Psychology, 56*(3), 374-386. doi:10.1037//0022-3514.56.3.374

Neuberg, S. L., & Fiske, S. T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology, 53*(3), 431-444. doi:10.1037/0022-3514.53.3.431

North, M. S., Todorov, A., & Osherson, D. N. (2010). Inferring the preferences of others from spontaneous, low-emotional facial expressions. *Journal of Experimental Social Psychology, 46*(6), 1109-1113. doi:10.1016/j.jesp.2010.05.021

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: from measures to mechanisms. *Trends Cogn Sci, 15*(4), 152-159. doi:10.1016/j.tics.2011.01.005

Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol Rev, 108*(3), 483-522. doi:10.1037/0033-295x.108.3.483

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biol Psychol, 77*(3), 247-265. doi:10.1016/j.biopsycho.2007.11.006

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*(32), 11087. doi:10.1073/pnas.0805664105

Orbell, J., Dawes, R., & Schwartz-Shea, P. (1994). Trust, social categories, and individuals: The case of gender. *Motivation and Emotion, 18*(2), 109-128. doi:10.1007/BF02249396

Otten, S., & Moskowitz, G. B. (2000). Evidence for Implicit Evaluative In-Group Bias: Affect-Biased Spontaneous Trait Inference in a Minimal Group Paradigm. *Journal of Experimental Social Psychology, 36*(1), 77-89. doi:10.1006/jesp.1999.1399

Overbeck, J. R., & Park, B. (2006). Powerful perceivers, powerless objects: Flexibility of powerholders' social attention. *Organ Behav Hum Decis Process, 99*(2), 227-243. doi:10.1016/j.obhdp.2005.10.003

Padmala, S., & Pessoa, L. (2011). Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *J Cogn Neurosci, 23*(11), 3419-3432. doi:10.1162/jocn_a_00011

Palermo, R., Rivolta, D., Wilson, C. E., & Jeffery, L. (2011). Adaptive face space coding in congenital prosopagnosia: typical figural aftereffects but abnormal identity aftereffects. *Neuropsychologia, 49*(14), 3801-3812. doi:10.1016/j.neuropsychologia.2011.09.039

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia, 49*(5), 1226-1235. doi:10.1016/j.neuropsychologia.2011.02.021

Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology, 42*(6), 1051-1068. doi:10.1037/0022-3514.42.6.1051

Parkinson, B. (2005). Do facial movements express emotions or communicate motives? *Pers Soc Psychol Rev, 9*(4), 278-311. doi:10.1207/s15327957pspr0904_1

Paxton, P. (2001). Trust: A sociological theory. *Social Forces, 79*(3), 1187-1188. doi:10.1353/sof.2001.0022

Pendry, L. (1998). When the mind is otherwise engaged: resource depletion and social stereotyping. *European Journal of Social Psychology, 28*(2), 293-299. doi:10.1002/(SICI)1099-0992(199803/04)28:2<293::AID-EJSP869>3.0.CO;2-7

Pendry, L. F., & Macrae, C. N. (1994). Stereotypes and mental life: The case of the motivated but thwarted tactician. *Journal of Experimental Social Psychology, 30*(4), 303-325. doi:10.1006/jesp.1994.1015

Pendry, L. F., & Macrae, C. N. (1996). What the Disinterested Perceiver Overlooks: Goal-Directed Social Categorization. *Personality and Social Psychology Bulletin, 22*(3), 249-256. doi:10.1177/0146167296223003

Pendry, L. F., & Macrae, C. N. (1999). Cognitive load and person memory: The role of perceived group variability. *European Journal of Social Psychology, 29*(7), 925-942. doi:10.1002/(SICI)1099-0992(199911)29:7<925::AID-EJSP973>3.0.CO;2-O

Phaf, R. H., Mohr, S. E., Rotteveel, M., & Wicherts, J. M. (2014). Approach, avoidance, and affect: a meta-analysis of approach-avoidance tendencies in manual reaction time tasks. *Front Psychol, 5*, 378. doi:10.3389/fpsyg.2014.00378

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *J Cogn Neurosci, 12*(5), 729-738.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*(3), 811-832. doi:10.1037/0022-3514.75.3.811

Plant, E. A., Hyde, J. S., Keltner, D., & Devine, P. G. (2000). The Gender Stereotyping of Emotions. *Psychology of Women Quarterly, 24*(1), 81-92. doi:10.1111/j.1471-6402.2000.tb01024.x

Plöderl, M. (2014). Bayesian Advice for Gaydar-based Picking Up: Commentary on Lyons, Lynch, Brewer, and Bruno (2013). *Archives of Sexual Behavior, 43*(1), 7-9. doi:10.1007/s10508-013-0178-x

Porcheron, A., Mauger, E., & Russell, R. (2013). Aspects of facial contrast decrease with age and are cues for age perception. *PLoS One, 8*(3), e57985-e57985. doi:10.1371/journal.pone.0057985

Poundstone, W. (1992). *Prisoner's Dilemma: John Von Neumann, Game Theory and the Puzzle of the Bomb*: Doubleday.

Price, J., & Murray, R. G. (2012). The region of proximal learning heuristic and adult age differences in self-regulated learning. *Psychology and Aging, 27*(4), 1120-1129. doi:10.1037/a0029860

Quadflieg, S., Turk, D. J., Waiter, G. D., Mitchell, J. P., Jenkins, A. C., & Macrae, C. N. (2009). Exploring the neural correlates of social stereotyping. *J Cogn Neurosci, 21*(8), 1560-1570. doi:10.1162/jocn.2009.21091

Quattrone, G. A., & Jones, E. E. (1980). The perception of variability within in-groups and out-groups: Implications for the law of small numbers. *Journal of Personality and Social Psychology, 38*(1), 141-152. doi:10.1037/0022-3514.38.1.141

Quillian, L., & Pager, D. (2001). Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime. *American Journal of Sociology, 107*(3), 717-767. doi:10.1086/338938

Quinn, K. A., & Macrae, C. N. (2005). Categorizing others: the dynamics of person construal. *Journal of Personallity and Social Psychology, 88*(3), 467-479. doi:10.1037/0022-3514.88.3.467

Radvansky, G. A., Copeland, D. E., & von Hippel, W. (2010). Stereotype activation, inhibition, and aging. *Journal of Experimental Social Psychology, 46*(1), 51-60. doi:10.1016/j.jesp.2009.09.010

Re, D. E., & Rule, N. (2017). Distinctive Facial Cues Predict Leadership Rank and Selection. *Personality and Social Psychology Bulletin, 43*(9), 1311-1322. doi:10.1177/0146167217712989

Reuter-Lorenz, P. A., & Park, D. C. (2014). How does it STAC up? Revisiting the scaffolding theory of aging and cognition. *Neuropsychology Review, 24*(3), 355-370. doi:10.1007/s11065-014-9270-9

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research, 46*(18), 2977-2987. doi:10.1016/j.visres.2006.03.002

Rhodes, G., Jeffery, L., Taylor, L., Hayward, W. G., & Ewing, L. (2014). Individual differences in adaptive coding of face identity are linked to individual differences in face recognition ability. *J Exp Psychol Hum Percept Perform, 40*(3), 897-903. doi:10.1037/a0035939

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychol Sci, 22*(4), 464-471. doi:10.1177/0956797611401753

Richler, J. J., & Gauthier, I. (2014). A Meta-Analysis and Review of Holistic Face Processing. *Psychol Bull, 140*(5), 1281-1302. doi:10.1037/a0037004

Richler, J. J., Mack, M. L., Gauthier, I., & Palmeri, T. J. (2009). Holistic processing of faces happens at a glance. *Vision Research, 49*(23), 2856-2861. doi:10.1016/j.visres.2009.08.025

Ridout, N., Astell, A., Reid, I., Glen, T., & O'Carroll, R. (2003). Memory bias for emotional facial expressions in major depression. *Cogn Emot, 17*(1), 101-122. doi:10.1080/02699930302272

Rivers, A. M., Sherman, J. W., Rees, H. R., Reichardt, R., & Klauer, K. C. (2019). On the roles of stereotype activation and application in diminishing implicit bias. *Pers Soc Psychol Bull*, 146167219853842. doi:10.1177/0146167219853842

Rodriguez-Bailon, R., Moya, M., & Yzerbyt, V. (2006). When power is undeserved: its effects on perception and social judgements. *Psicothema, 18*(2), 194-199.

Rodríguez-Bailón, R., Moya, M., & Yzerbyt, V. (2000). Why do superiors attend to negative stereotypic information about their subordinates? Effects of power legitimacy on social perception. *European Journal of Social Psychology, 30*(5), 651-671. doi:10.1002/1099-0992(200009/10)

Romano, A., Balliet, D., Yamagishi, T., & Liu, J. H. (2017). Parochial trust and cooperation across 17 societies. *Proceedings of the National Academy of Sciences, 114*(48), 12702.

Rule, N. O. (2017). Perceptions of sexual orientation from minimal cues. *Arch Sex Behav, 46*(1), 129-139. doi:10.1007/s10508-016-0779-2

Rule, N. O., & Ambady, N. (2008). The Face of Success: Inferences from Chief Executive Officers' Appearance Predict Company Profits. *Psychol Sci, 19*(2), 109-111.

Rule, N. O., Ambady, N., Adams, R. B., Jr., & Macrae, C. N. (2007). Us and them: memory advantages in perceptually ambiguous groups. *Psychonomic Bulletin & Review, 14*(4), 687-692.

Rule, N. O., Garrett, J. V., & Ambady, N. (2010). On the perception of religious group membership from faces. *PLoS One, 5*(12), e14241. doi:10.1371/journal.pone.0014241

Ruscher, J. B., & Fiske, S. T. (1990). Interpersonal competition can cause individuating processes. *Journal of Personality and Social Psychology, 58*(5), 832-843. doi:10.1037//0022-3514.58.5.832

Ruscher, J. B., Fiske, S. T., & Schnake, S. B. (2000). The motivated tactician's juggling act: compatible vs. incompatible impression goals. *Br J Soc Psychol, 39 (Pt 2)*, 241-256.

Ruz, M., Madrid, E., & Tudela, P. (2013). Interactions between perceived emotions and executive attention in an interpersonal game. *Soc Cogn Affect Neurosci, 8*(7), 838-844. doi:10.1093/scan/nss080

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion, 9*(2), 260-264. doi:10.1037/a0014681

Salthouse, T. A. (2010). Selective review of cognitive aging. *J Int Neuropsychol Soc, 16*(5), 754-760. doi:10.1017/s1355617710000706

Sanz, J., Gutiérrez, S., & García-Vera, M. P. (2014). Propiedades psicométricas de la Escala de Valoración del Estado de Ánimo (EVEA): Una revisión. [Psychometric properties of the Scale for Mood Assessment (EVEA): A review.]. *Ansiedady Estrés, 20*(1), 27-49.

Sanz, J., Navarro, M. E., & Vázquez, C. (2003). Adaptación española del inventario para la depresión de beck--II(BDI-II): 1. Propiedades psicométricas en estudiantes universitarios. [Spanish adaptation of the Beck Depression Inventory-II (BDI-II): 1. Psychometric properties with university students.]. *Análisis y Modificación de Conducta, 29*(124), 239-288.

Sattler, J. M. (1970). Racial "experimenter effects" in experimentation, testing, interviewing, and psychotherapy. *Psychol Bull, 73*(2), 137-160. doi:10.1037/h0028511

Schlösser, T., Fetchenhauer, D., & Dunning, D. (2016). Trust against all odds? Emotional dynamics in trust behavior. *Decision, 3*(3), 216-230. doi:10.1037/dec0000048

Schmid Mast, M., Khademi, M., & Palese, T. (2020). Power and social information processing. *Current Opinion in Psychology, 33*, 42-46. doi:10.1016/j.copsyc.2019.06.017

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburg: Psychology Software Tools Inc.

Schniter, E., & Shields, T. W. (2014). Ageism, honesty, and trust. *Journal of Behavioral and Experimental Economics, 51*, 19-29. doi:10.1016/j.socec.2014.03.006

Scott, J. P. (1980). The function of emotions in behavioral systems: A systems theory analysis. In R. Plutchik & H. Kellerman (Eds.), *Theories of Emotion* (pp. 35-56): Academic Press.

Scott, L. S. (2006). Featural and configural face processing in adults and infants: a behavioral and electrophysiological investigation. *Perception, 35*(8), 1107-1128. doi:10.1068/p5493

Sen, A. (2003). Missing women--revisited. *BMJ (Clinical research ed.), 327*(7427), 1297-1298. doi:10.1136/bmj.327.7427.1297

Shen, H. (2013). Inequality quantified: Mind the gender gap. *Nature, 495*(7439), 22-24. doi:10.1038/495022a

Sherman, J. W., Macrae, C. N., & Bodenhausen, G. V. (2000). Attention and stereotyping: Cognitive constraints on the construction of meaningful social impressions. *European Review of Social Psychology, 11*(1), 145-175. doi:10.1080/14792772043000022

Siemer, M., Mauss, I., & Gross, J. J. (2007). Same situation--different emotions: how appraisals shape our emotions. *Emotion, 7*(3), 592-600. doi:10.1037/1528-3542.7.3.592

Slone, A. E., Brigham, J. C., & Meissner, C. A. (2000). Social and Cognitive Factors Affecting the Own-Race Bias in Whites. *Basic and Applied Social Psychology, 22*(2), 71-84. doi:10.1207/S15324834BASP2202_1

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48*(4), 813-838. doi:10.1037/0022-3514.48.4.813

Solnick, S. J., & Schweitzer, M. E. (1999). The influence of physical attractiveness and gender on Ultimatum Game decisions. *Organ Behav Hum Decis Process, 79*(3), 199-215. doi:10.1006/obhd.1999.2843

Sommers, S. R. (2006). On racial diversity and group decision making: Identifying multiple effects of racial composition on jury deliberations. *Journal of Personality and Social Psychology, 90*(4), 597-612. doi:10.1037/0022-3514.90.4.597

Spears, R., Oakes, P. J., Ellemers, N., & Haslam, S. A. (1997). *The social psychology of stereotyping and group life*. Malden: Blackwell Publishing.

Spielberg, C. D., Gorsuch, R. L., & Lushene, R. E. (2008). *STAI : cuestionario de ansiedad estado-rasgo : Manual*. Madrid: TEA.

Spielberger, C. D. (1999). *STAXI-2 : State-Trait Anger Expression Inventory-2 : professional manual*. Odessa, FL (P.O. Box 998, Odessa 33556): Psychological Assessment Resources.

Sporer, S. L. (2001). The cross-race effect: Beyond recognition of faces in the laboratory. *Psychology, Public Policy, and Law, 7*(1), 170-200. doi:10.1037/1076-8971.7.1.170

Stahl, J., Wiese, H., & Schweinberger, S. R. (2008). Expertise and own-race bias in face processing: an event-related potential study. *Neuroreport, 19*(5), 583-587. doi:10.1097/WNR.0b013e3282f97b4d

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences, 108*(19), 7710. doi:10.1073/pnas.1014345108

Stolier, R. M., & Freeman, J. B. (2016). The neuroscience of social vision. In J. Cloutier (Ed.), *Neuroimaging Personality, Social Cognition, and Character* (pp. 139-157). San Diego: Academic Press.

Stravinsky, I. (1913). Rite of Spring.

Stroessner, S. J. (1996). Social Categorization by Race or Sex: Effects of Perceived Non-Normalcy on Response Times. *Social Cognition, 14*(3), 247-276. doi:10.1521/soco.1996.14.3.247

Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., . . . Rhodes, G. (2010). Face recognition impairments despite normal holistic processing and face space coding: evidence from a case of developmental prosopagnosia. *Cogn Neuropsychol, 27*(8), 636-664. doi:10.1080/02643294.2011.613372

Sweeny, T. D., Grabowecky, M., Suzuki, S., & Paller, K. A. (2009). Long-lasting effects of subliminal affective priming from facial expressions. *Consciousness and cognition, 18*(4), 929-938. doi:10.1016/j.concog.2009.07.011

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues, 25*(4), 79-97. doi:10.1111/j.1540-4560.1969.tb00620.x

Tajfel, H. (1970). Experiments in Intergroup Discrimination. *Scientific American, 223*(5), 96-103.

Tajfel, H. (1978). The achievement of group differentiation. In H. Tajfel (Ed.), *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations* (pp. 77-98). London: Academic Press.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*(2), 149-178. doi:10.1002/ejsp.2420010202

Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Monterey, Calif: Brooks/Cole Pub. Co.

Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgement. *Br J Psychol, 54*, 101-114. doi:10.1111/j.2044-8295.1963.tb00865.x

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. *Cognition, 93*(1), B1-9. doi:10.1016/j.cognition.2003.09.011

Tanis, M., & Postmes, T. (2005). A social identity approach to trust: interpersonal perception, group membership and trusting behaviour. *European Journal of Social Psychology, 35*(3), 413-424. doi:10.1002/ejsp.256

Telga, M., Alcalá, J. A., & Lupiáñez, J. (2019). *Motivated social and non-social categorization in decision making and learning*. [Manuscript in preparation].

Telga, M., Cantiani, A., & Lupiáñez, J. (2019). *Individuation at any cost: Motivation to individuate prevents social categorization in cognitively demanding contexts*. [Manuscript in preparation].

Telga, M., de Lemus, S., Cañadas, E., Rodríguez-Bailón, R., & Lupiáñez, J. (2018). Category-based learning about deviant outgroup members hinders performance in trust decision making. *Front Psychol, 9*(1008). doi:10.3389/fpsyg.2018.01008

Telga, M., Kawakami, K., & Lupiáñez, J. (2019). *Ethnic outgroup homogeneity effect in trust decision-making*. [Manuscript in preparation].

Telga, M., & Lupiáñez, J. (2019). *Differences between men and women in social perception: An effect of social power?* [Under review].

Telga, M., & Lupiáñez, J. (2019). *Use of heuristics in social learning by younger and older adults*. [Under review].

Tetlock, P. E. (1983). Accountability and complexity of thought. *Journal of Personality and Social Psychology, 45*(1), 74-83. doi:10.1037/0022-3514.45.1.74

Teunisse, J. P., & de Gelder, B. (2003). Face processing in adolescents with autistic disorder: the inversion and composite effects. *Brain Cogn, 52*(3), 285-294.

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: complementarity in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology, 84*(3), 558-568.

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology, 81*(6), 973-988. doi:10.1037/0022-3514.81.6.973

Todd Maddox, W., & Markman, A. B. (2010). The Motivation–Cognition Interface in Learning and Decision Making. *Current Directions in Psychological Science, 19*(2), 106-110. doi:10.1177/0963721410364008

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition, 27*(6), 813-833. doi:10.1521/soco.2009.27.6.813

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Science, 12*(12), 455-460. doi:10.1016/j.tics.2008.10.001

Tomasello, M. (2009). *Why we cooperate*. Cambridge, MA: MIT Press.

Tortosa, M. I., Lupiáñez, J., & Ruz, M. (2013). Race, emotion and trust: An ERP study. *Brain Research, 1494*(0), 44-55. doi:10.1016/j.brainres.2012.11.037

Tortosa, M. I., Strizhko, T., Capizzi, M., & Ruz, M. (2013). Interpersonal effects of emotion in a multi-round Trust Game. *Psicológica, 24*(2), 179-198.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res, 168*(3), 242-249. doi:10.1016/j.psychres.2008.05.006

Tracy, J. L., & Robins, R. W. (2008). The automaticity of emotion recognition. *Emotion, 8*(1), 81-95. doi:10.1037/1528-3542.8.1.81

Tskhay, K. O., & Rule, N. O. (2013). Accuracy in categorizing perceptually ambiguous groups: a review and meta-analysis. *Pers Soc Psychol Rev, 17*(1), 72-86. doi:10.1177/1088868312461308

Tskhay, K. O., Wilson, J. P., & Rule, N. O. (2016). People Use Psychological Cues to Detect Physical Disease From Faces. *Personality and Social Psychology Bulletin, 42*(10), 1309-1320. doi:10.1177/0146167216656357

Turner, J. C., Brown, R. J., & Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology, 9*(2), 187-204. doi:10.1002/ejsp.2420090207

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA: Basil Blackwell.

Uleman, J. S., & Kressel, L. M. (2013). A Brief History of Theory and Research on Impression Formation. In D. Carlston (Ed.), *The Oxford Handbook of Social Cognition*: Oxford University Press.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Q J Exp Psychol A, 43*(2), 161-204.

van 't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition, 108*(3), 796-803. doi:10.1016/j.cognition.2008.07.002

von Hippel, W., & Henry, J. D. (2012). Social cognitive aging. In S. T. Fiske & C. N. Macrae (Eds.), The SAGE handbook of social cognition. London: SAGE Publications Ltd. doi:10.4135/9781446247631

Wenzlaff, R. M., Wegner, D. M., & Roper, D. W. (1988). Depression and mental control: the resurgence of unwanted negative thoughts. *Journal of Personality and Social Psychology, 55*(6), 882-892.

Willis, J., & Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychol Sci, 17*(7), 592-598. doi:10.1111/j.1467-9280.2006.01750.x

Wilson, R. K., & Eckel, C. C. (2011). Trust and Social Exchange. In J. N. Druckman, D. P. Green, J. H. Kuklinski, & A. Lupia (Eds.), *Cambridge Handbook of Experimental Political Science:* (pp. 243-257). Cambridge: Cambridge University Press.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: unwanted influences on judgments and evaluations. *Psychol Bull, 116*(1), 117-142. doi:10.1037/0033-2909.116.1.117

Wilson, W., & Kayatani, M. (1968). Intergroup attitudes and strategies in games between opponents of the same or of a different race. *Journal of Personality and Social Psychology, 9*(1), 24-30.

Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica, 114*(1), 101-114. doi:10.1016/S0001-6918(03)00052-0

Wright, D. B., & Stroud, J. N. (2002). Age differences in lineup identification accuracy: People are better with their own age. *Law and Human Behavior, 26*(6), 641-654. doi:10.1023/A:1020981501383

Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *The Journal of Neuroscience, 29*(26), 8525.

Yamagishi, T., Jin, N., & Kiyonari, T. (1999). *Bounded Generalized Reciprocity: Ingroup Boasting and Ingroup Favoritism* (Vol. 16).

Zebrowitz, L. A., White, B., & Wieneke, K. (2008). Mere exposure and racial prejudice: Exposure to other-race faces increases liking for strangers of that race. *Social Cognition, 26*(3), 259-275. doi:10.1521/soco.2008.26.3.259